Asking About Data

Exploring Different Realities of Data via the
Social Data Flow Network Methodology

Brian Ballsun-Stanton
The University of New South Wales

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: Ballsun-Stanton

First name: Brian   Other name/s:

Abbreviation for degree as given in the University calendar: PhD

School: Humanities   Faculty: Arts and Social Sciences

Title: Asking About Data Exploring Different Realities of Data via the Social Data Flow Network Methodology

Abstract

What is data? That question is the fundamental investigation of this dissertation. I have developed a methodology from social-scientific processes to explore how different people understand the concept of data, rather than to rely on my own philosophical intuitions or thought experiments about the "nature" of data. The evidence I have gathered as to different individuals' constructions of data can be used to inform further inquiry of data and the design of information systems.

My research demonstrates that people have different constructions of data. The methodology of the SDFN, created for this dissertation, has proven able to probe those understandings. The SDFN, loosely based on a DFD and combined with ideas from SNA, provides a way of discovering practical definitions of hard-to-operationalize terms like data. The process of repeatedly categorizing various items as data allows the methodology to explore how participants actually use the term, rather than relying on theoretical dictionary-based definitions.

Analysis of the interviews found three different constructions of data: data as communications, a container for meaning; data as subjective observations, sense-impressions filtered by knowledge; and data as objective facts, measurements revealing the relationships of reality.

FOR OFFICE USE ONLY   Date of completion of requirements for Award:

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed   …………………………………………….................

Date   …………………………………………….................

Asking About Data

Exploring Different Realities of Data via the
Social Data Flow Network Methodology

By: Brian Ballsun-Stanton

A thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF
NEW SOUTH WALES

SYDNEY · AUSTRALIA

School of Humanities

April 2012

Contents

1 Acknowledgements

I thank the employees of BlueScope Steel in Australia for participating in my survey and answering difficult questions about how they thought about the world. Their participation has made my research possible.

I appreciate the members of the Pentagon's INTELST mailing list for their participation in the survey and Kevin DiVico for putting me into contact with them and his assistance in other matters.

I extend a significant deal of gratitude to my present and past advisers, Dr. Anthony Corones and Dr. Deborah Bunker. Without their expert guidance and assistance, I would not have had a dissertation of this quality.

I sincerely thank John Rennie for his expert editing assistance, lending a final polish to this work.

Lucy: "Coffee on Sunday?"

Mikhail: I can intone only: "Ia! Ia! Cthulhu fhtagn! Ph'nglui mglw'nafh Cthulhu R'lyeh wgah-nagl fhtaga -"

To the rest of my friends: thank you for keeping me sane-ish during this three year ordeal. Without their time, friendship, and support, I would never have completed this work.

1.1 Figures

The title background is a heavily photoshopped version of Jim Sanborn's Kryptos statue. The image was provided to the Wikimedia Commons under a Creative Commons CC-BY-SA 3.0 license. As this work is also licensed under a CC-BY-SA-NC license, it satisfies all requirements. Original image available at: http://en.wikipedia.org/wiki/File:Kryptos01_1.jpg

The other external image I use, figure 4.1, is also CC licensed, used by permission of the author, Daniel P. Lee. http://www.flickr.com/photos/yankeeincanada/3658159896/

All other images were produced by myself for this dissertation.

2 Abbreviations

DFD	Data Flow Diagram
HCI	Human-Computer Interaction
IST	Information Systems & Technology
SDFN	Social Data Flow Network
SNA	Social Network Analysis
UoD	Universe of Discourse

3 Abstract

What is data? That question is the fundamental investigation of this dissertation. I have developed a methodology from social-scientific processes to explore how different people understand the concept of data, rather than to rely on my own philosophical intuitions or thought experiments about the "nature" of data. The evidence I have gathered as to different individuals' constructions of data can be used to inform further inquiry of data and the design of information systems.

My research demonstrates that people have different constructions of data. The methodology of the Social Data Flow Network, created for this dissertation, has proven able to probe those understandings. The Social Data Flow Network, loosely based on a Data Flow Diagram and combined with ideas from Social Network Analysis, provides a way of discovering practical definitions of hard-to-operationalize terms like *data.* The process of repeatedly categorizing various items as data allows the methodology to explore how participants actually use the term, rather than relying on theoretical dictionary-based definitions.

Analysis of the interviews found three different constructions of data: data as communications, a container for meaning; data as subjective observations, sense-impressions filtered by knowledge; and data as objective facts, measurements revealing the relationships of reality[*].

---

[*] For a longer summary of this research, look at Appendix D. The peer-reviewed paper on page 397 was presented at the IEEE 5th International Conference on Computer Sciences and Convergence Information Technology in Seoul, Korea during the process of writing the thesis.

4 Introduction

In Information Systems & Technology studies (IST), I have noticed that practitioners use and understand the term "data" differently than the people they are helping. The purpose of this research is to explore the different conceptions of data that may exist beyond the domain of IST and demonstrate a methodology that allows practitioners to access the conceptions of data present in their workplace.

Exploring a conception of data is fundamentally a philosophical problem. A person's conception of data stems from the affordances they attach to it, their belief in its underlying qualities, and their differentiation between data and non-data. However, this philosophical problem cannot be solved through intuition alone: a methodology is necessary to extract a person's conception of data.

These individual conceptions can then be formalised as "philosophies of data." By 'philosophies' we mean answers to the questions like, 'What is data?', 'What is data for?', 'How do I know the data is reliable?', and 'What are the properties of data?' While individuals may not "have philosophies," understanding that individuals engage philosophically with their conceptions of data allows the creation of a tool to probe those philosophical conceptions of data in a workplace. By probing conceptions, the IST practitioner effectively uncovers de facto philosophies of data in individuals.

This research, however, does not propose to uncover fundamental philosophies of data, only some common conceptions of data that may exist in workplaces. These different conceptions of data can produce frustration, error, and miscommunication if people with different conceptions interact unknowingly. Conceptions of data include context, reliability, constraints as to its nature (can it be a description, must it be a

number), the means of collection, and the means of manipulation.

I have created a methodology called the Social Data Flow Network (SDFN). This interview technique has elicited people's conceptions of data[*], demonstrating three different conceptions within a particular industrial research workplace. A survey developed from the SDFN technique hints that there may be different conceptions of data present in the intelligence analysis community and the IST practitioner community.

It is my hope that IST practitioners can use the SDFN I have developed to make better interfaces and databases: through the understanding of a client's expectations of data, the system can provide natural interaction methods that conform to the client's expectations of what data is and is not. The SDFN might also be used within an organization to reduce miscommunication and error: the explicit definition of one particular conception of data for a workplace.

4.1   Methodological Summary

The primary result of this thesis is the methodology of the Social Data Flow Network. The SDFN uses repeated categorization to explore how individuals group informational or communicative flows into categories. By eliciting categorizes that focus on data, information, and knowledge, the participants use the categorization to operationalize their epistemological understanding of data: they indicate what is and is not data and how it becomes information and knowledge. This elicitation helps both the interviewer and the participant to discover their own situational conceptualization of data.

The repeated categorization allows participants to generate and resolve cognitive dissonance situated around the differences between their theoretical definitions of data

---

[*]   their de facto philosophical approaches towards knowing that something is or is not data

and their practical uses and categorizes of data. In interviews, participants demonstrated a refined understanding of their own conceptions of data at the end of the interview, catalyzed through their participation in the SDFN.

The SDFN involves the articulation of roles as entities, descriptions of content flows between those entities and the categorization of those flows as data, information, knowledge, or other. Participants iterate over a task domain defined at the start of the interview, discussing all the entities and flows between those entities involved in the task. The interview concludes with an opportunity for the participant to self reflect on their "philosophy" of data, discussing what they categorize data as and how it becomes information and/or knowledge.

A scenario based survey, inspired by the SDFN was also trialled with less satisfactory results. While the survey did demonstrate that intelligence officers, IST professionals, and other industrial research employees did have different conceptions of data, it did not do so with any statistical rigor nor with the depth of discussion that the interviews provided.

The SDFN combines two concepts for a novel purpose. It is a graph[*] that combines

---

[*] A graph, strictly speaking, is any diagram that contains edges and nodes. A node is the component of a graph that is a point. The point can be labeled or unlabeled. The node is the element of the graph that is a representation of a thing. Sometimes the thing being represented is a computer or a person, or a place, but in any event the node represents a noun. Edges on the other hand are the relationships or connections between nodes. An edge represents a "flow" of action or stuff between nodes. Edges traditionally have served as network links, roads, phone lines, and simple representations of adjacency. A graph is a non-topological method of representing the relationships between entities through edges and nodes.

Edges can be directed: they show a flow or relationship from one node to another. The direction on the edge indicates the direction of relationship. For example, consider Alice and Bob. To represent Alice sending a letter to Bob, we would make both of them nodes and draw a directed edge from Alice to Bob indicating the one way flow of the letter. By adding the concept of directionality to edges, a causal element is introduced to the representation-specifically, that the originating node causes a relationship to the recipient node. This addition of causality then precipitates the idea of connectiveness.

A node may or may not be reachable by other nodes. A graph or subgraph where every node can be reached from every other node is called a strongly connected graph. A graph where that's not true is weakly

the idea of the social network with that of the data flow diagram. In social network analysis, it is possible to represent interactions between people, a social network, through graphs. Each node on a graph represents a person and each edge represents some sort of connection between people, as a function of the interactions of interest to the researcher [1].

The Data Flow Diagram [2 and 3] contributes its diagrams to the SDFN. A DFD originally was designed for structured programming. The document produced by the DFD would combine the delineation of a universe of discourse via the context diagram with the highly precise definition of flows into and out of that diagram. A Universe of Discourse (UoD) [4] is the term used for defining the topic under consideration. Everything within the UoD is relevant and must be modeled. Everything outside the UoD is irrelevant. Interestingly, as the DFD was repurposed for business modeling, the UoD remained the same: it is still asking, "What bit of reality do we care about right now?"

The DFD would then be refined through a process of "zooming in" on that context diagram to expose the transformations required to produce the outputs from the inputs. Each additional level would seek to conserve inputs and outputs, and thereby produce a diagram that could be mapped to the functions and variables necessary for a structured program.

The DFD contributes great ideas to the SDFN. It contributes the idea that data *is* something that can be modeled. The conception of data embodied by the DFD is that the modeler can translate reality into data-as-bits and that data could be described through text. All actions in the data flow diagram are considered either flows or transformation. Data flows from sources through transformations, and out into sinks. The sources and

---

connected. When we apply the idea of strongly connected graphs to social networks, we can identify small groups by identifying strongly connected subgraphs within a larger, weakly connected graph.

sinks are entities outside the scope of the diagram. By decomposing these transformations into ever simpler and more detailed sets of sub-transformations, modelers could design an entire software system intended to process and transform data. The modeler acts as translator: taking the described reality by the client and forcing it into a computerized mold. Repurposing the methodology of the DFD by subtracting the modeler's translation suggests that it might be possible to use my method to probe and document a client's subjective reality.

The DFD also contributes an iterative structure for the definition of reality. The iterative techniques explore the UoD in order of increasing specificity from the vague context diagram describing the universe of discourse to highly detailed sub-sub-sub (etc.) transformations required deep in the diagram. By starting with broad generalizations, the DFD insured that the client was thinking about the whole task and did not immediately become fixated on any one aspect. With the DFD iterating across each declared "transformation" and decomposing it, the details of each transformation were both evoked and then situated in the scaffolding of the broader context. The requirement to conserve inputs and outputs eliminated any question of missing aspects of the diagram or other design-based blind alleys. The idea of iterative exploration and definition is extremely valuable to the SDFN.

The Social Network Graph provides the concept of a social network[*] to the SDFN. The Social Network Graph also contributes a novel idea about the *scope* of edges. Edges in the DFD were simple *flows* of data, representing the movement of trivial signs. In the social network graph, edges can be individual communications, orders, relationships,

---

[*] A social network graph is a mapping of a person's relationships with other people into non-topological graph format. Each relationship is a directed edge; each person, a node. The social network graph is used in many different fields: communications, social media, and sociology are some of them. In many ways, the idea of the social network graph is strongly related to the ideas of actor-network theory [5].

and objects. The huge diversity of edge types suggested by a social network graph, when combined with the DFD, ruins the DFD for its original purpose: the modeling of software systems. However, they also suggest different possible models that can be applied to the DFD format.

In communicative analysis, social network graphs are used for linguistic analysis[*]. It is possible to explore the control structures of a group by noting, with an edge, who is talking to whom. By exploring the frequency and directionality of those notes, analysts gain insights into the power and influence roles of social networks. As such, the "thought leaders" of the small group can be identified.

Moreover, by graphing flows of communication, it is possible to identify small groups within larger groups, as these small groups will communicate strongly between each other and vaguely to nodes outside. In other circles, this behavior is known as siloing [8]. One design intent of the SDFN is to confer the ability to identify siloing. By rendering flows between members of an organization, it should be possible to identify strongly connected sub-graphs, which suggest communicative silos within that organization.

The social network graph contribution alters the diagramming rules of the DFD. Social network entities can be any actor that participates in a communication. The SDFN is a diagram exploring flows of data between actors, instead of flows between transformations. By creating a web of affiliation [9] between these entities, it should be possible to describe the communicative realities that an individual perceives. It should therefore be possible to explore how they understand the nature of data by exploring how they

---

[*] figure 4.1 provides a trivial example of linguistic analysis as applied to a set of twitter replies during a conference. The different line weights are used to denote quantity of communications along a radially distributes set of nodes. Other approaches can be far more complex, looking at patterns beyond simple frequency [6 and 7].
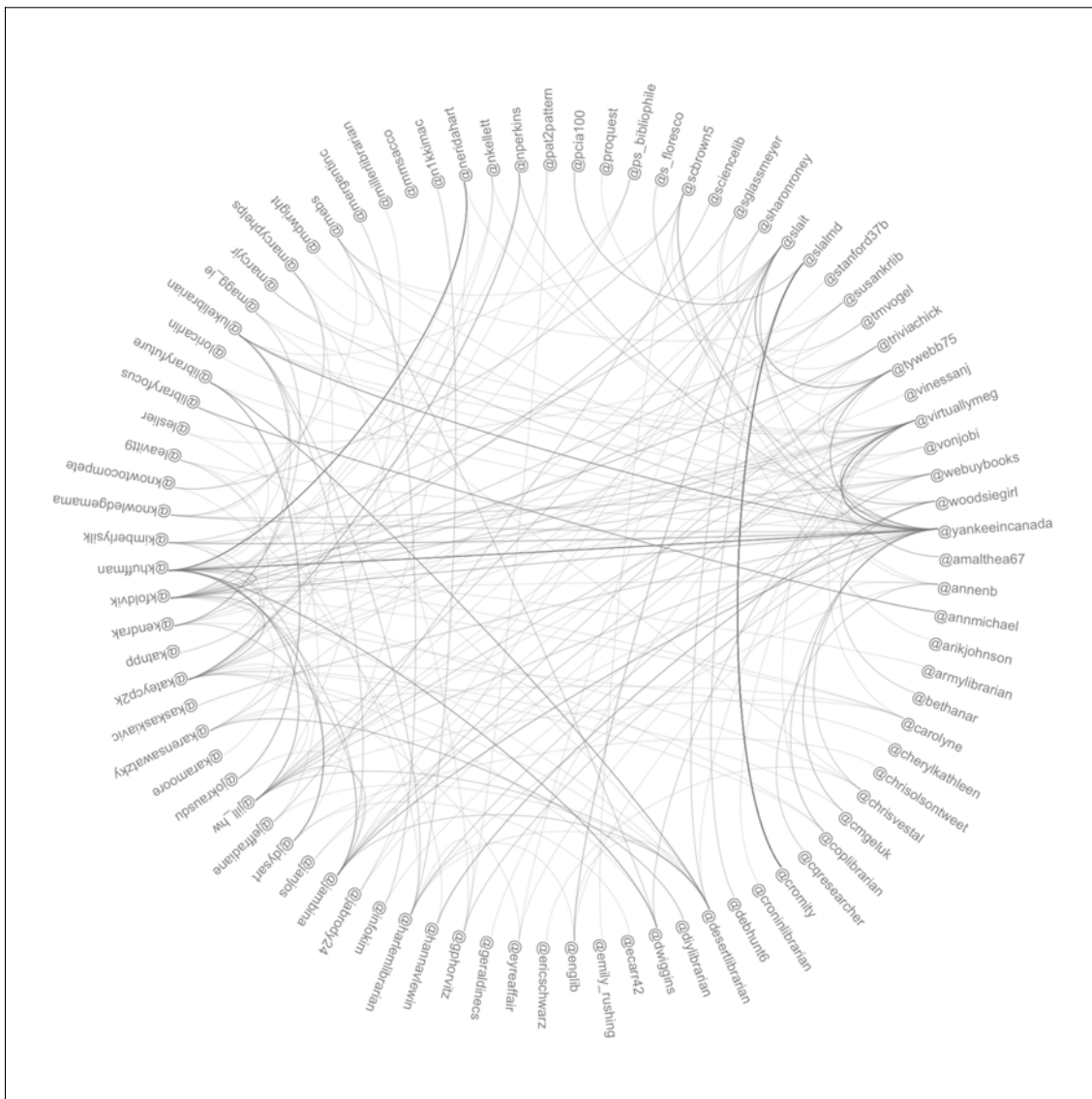
Figure 4.1:   Social network graph of #sla2009 tweet replies to June 19, 2009 "The thicker the line, the more times you sent an @reply to that person. The more lines you have, the more @replies to different people you sent. If you don't appear on the graph, but know that you sent out @replies, it's because the person you sent your @reply to never sent out an @reply and so that person won't appear on the graph and unfortunately, you can't either! Interestingly, a few people only sent replies to themselves, so they do appear on the graph as a line that goes back to themselves." -Image used with permission, created by: Daniel P. Lee, MLIS.

describe its movement from entity to entity in the SDFN.

Despite the terminology of actors, and the use of a social network, my research

does not yet incorporate actor-network theory [10]. While Latour's work offers many useful ideas for understanding the world, it still imposes a framework from which biases may be imparted. Therefore, while I do not use actor-network theory here, it may be useful in later research exploring the implications of held philosophies on Latour's work.

The SDFN does not try to be explanatory, comprehensive, or objective. The point of the SDFN is to reveal part of how the participant understands a concept, not to build upon that understanding nor transform it into a model for a computer system. Consequently, no design provisions in the methodology allow two or more peoples' categories to be reconciled. More work will be necessary before the SDFN can be used directly as a design methodology.

4.2   Analysis and Results

These questions of interest are posted to the reader to keep in mind in the results section. My personal analysis, presented after the "raw data," uses these questions of interest as framing devices for my reflections on the individual interviews.

4.2.1   Questions of Interest and the Methodology of Analysis

My "hypotheses" are described as questions of interest to reflect the rapid iterative nature of abductive explorations. They provide research directions that act as broad guides to the formation of a universe of discourse for future research rather than predictive statements about reality.

The intent of the questions is to frame analysis and guide it towards useful and interesting areas. We need to consider how the evidence relates to these questions of interest.

Each interview, after transcription, was subjected to recursive analysis for my personal reflections on the interviews. I summarized six to ten lines of each interview in a one-line summary. Then between three and six summaries were summarized, filtering for statements about the user's conception of data. Although self-transcription transmits personal bias, two significant factors prevent a traditional double-blind study. An untested methodology is no place for the mass utilization of volunteer interviewers. The limited scope allowed me to retain control of the interview process and to provide for the best possible interviews for each participant while retaining the basics of the SDFN. Because I conducted each interview, the bias would have already been introduced; providing for pious-sounding human coding would have lent false reliability to something inherently subjective.

My personal reflections are very simple. I have tried to extract each participant's intuitions about data from the recursive analysis.

4.2.1.1 Question of interest 1: do people have different realities of data?

If this research produces nothing else, it must investigate whether people have different conceptions of data. This idea was the central intuition that prompted this research, and its testing will demonstrate whether or not there is anything to my intuition.

As the organizing factor of my analysis, this question of interest will focus my activities. It will justify further research on the nature and subjective constructions of data from my experimental results, or else its demonstrable failure will justify not doing so.

The question "Do people have different realities of data?" defines an overly large universe of discourse, one impossible to study at a useful level of granularity in one research project. The very breadth of the question precludes the determination of any

useful and specific facts about the world besides simple exploration of the assertion that people have different understandings of data. The intent of this research and of this question is to generate interest in the research of the nature of data, how people understand it, and to demonstrate that there are potential areas of philosophy to research.

I want to see if, beyond my intuitive insight, people actually have different conceptions of data or if my perception of different conceptions is an artifact of the requirements-gathering process of designing a database. It is therefore not sufficient to state that people have different understandings of data depending on whether they are dealing with it in a technical or scientific context. We must look for evidence.

This question of interest, in its reach, is not ambitious. It suggests no predictions about peoples' conceptions of data, how they act with different realities of data, or any other fact about the world. Instead, it simply directs us to see if there is anything of interest for further explorations.

4.2.1.2 Question of interest 2: can my methodology probe people's realities of data?

My methodology has a simple job: to assess what people mean when they use the term "data." This question of interest is designed as a sanity check. I am investigating a new idea with an untested methodology. It is vital to consider that the success or failure of Question of interest 1 is directly modulated by the success or failure of Question of interest 2. Therefore, the methodology itself deserves distinct analysis.

The methodology should be of use to more people than just those investigating peoples' conceptions of data. If the methodology is useful and judged to add value to Question of interest 1, analysis of the methodology should indicate whether other people

could use it to investigate matters of interest to them.

Question of interest 2 is asking: do these results make sense? Sense-making is a matter of internal and external consistency. This question should force me to explore whether the SDFN correlates with interview results and whether the types of results make sense relative to the survey.

Beyond consistency, I must also ask: Is it possible to get these results from this methodology? In this case, I need to make sure that I am not reading imaginary meaning in the tea leaves of the results. Because this kind of external self-reflection is difficult, the question must be simplified to: Do the results surprise me? If they do not have elements of surprise, then the probability that I am projecting meaning into them must be strongly considered.

All of these are very self-critical questions, as they must be to explore the impact of an untested methodology. I am trying to consider whether my methodology can present a persuasive story, and if it can, does it?

4.3   Interview Analysis

My interview analysis discovered three different conceptions of data. It would be hard to deny that interviews I and II have data as communication, III and IV have subjective observations (with IX hinting at them) and the rest considering data as objective fact. With these broad differences evident, I feel question of interest 1 has been satisfied.

The observation constructions differ strikingly from the numeric constructions, possibly differing on a fundamental perception of reality. As one interview is trying to render the relationships between matter in the world as numbers (objectivist), another is suggesting that everything emits data and we must filter it. The conflict is records versus measurements versus signs. Does data measure objective reality, record subjective

reality, or merely transmit signs? Numbers are seen as a result of precision most of the time, whereas observations are building their way towards knowledge.

4.3.1 Result 1: Data as communications

Data, in the communicative sense, merely requires signs and things to communicate with those signs. The data can be rendered as bits or marks on paper, but it is seen as a factor of semiotic import rather than as something to be discovered or filtered.

This construction is substantively different from the other two inasmuch as it does not uphold data to be an aspect of reality. Instead, data is produced as a function of human intent. Because this understanding does not concern itself with interactions of the real, there is a far greater difference between this and the other two than between the subjective-objective constructions. However, the passivity of this construction allows it to accept facts produced from either source as something to be encoded, stored, and transmitted. Significant research needs to be done to explore how this construction of data relates to the other two.

4.3.2 Result 2: Data as subjective observations

Data, in the subjective observations, requires contextualization and filtering. Everything emits data as sense impressions[*] that can be captured by us. Thus, to perform sense-making activities, we must filter and contextualize the interesting data so that it can become information.

---

[*] Like the ancient Aristotelian idea of species (particles of sensation). Light was the medium that visual species traveled within.
While this ancient philosophy of image is not hugely useful to us, the same intuitions that led to it could have some parallels with data as subjective observations. This research area could make an interesting bridge between intuitive and experimental philosophies.

Subjective data lends itself more to cyclic hierarchies, where data begets the information and knowledge used to collect more data, reflecting an interestingly constructivist view of knowledge. There is quite a lot here available to future research, and I do not feel sufficiently confident in my sample size to make any assertions as to relationships between data and the various philosophies of knowledge or science, though the subjective nature of observations may tend slightly more towards Latour or Feyerabend.

Of more interest is that this inherently subjective data is constructed from the mind's impressions of the surroundings, rather than revealed through measurement of the surroundings. The understanding of the embodiment of data is a significant difference between the two understandings of data.

4.3.3 Result 3: Data as measured facts

Objective data comes with its own context "baked in." It is, in many ways, rare: it requires positive effort to generate, and higher quality data requires a commensurate increase in effort. Data requires analysis to uncover the extant patterns of reality, and with enough data, knowledge about the singular real can be generated.

Objective data requires that data be a fact, usually a numerical, reproducible representation of reality that conveys an understanding of measurement quality and units. Objective data is not filtered, because it is collected with prior intent and all elements of the "data set" may produce interesting patterns.

Both humans and sensors can reveal objective data, which is embodied in the things being measured. There seems to be no significant link with any of the major philosophies of science. Although my investigations did not explore confirmation, falsifiability, or paradigms, there seems to be a common understanding that data-as-fact accurately

represents the universe within the constraints of measurement. This may be because the participants believed data to be a building block upon which their hypotheses or understanding of the universe could be built.

5   Literature Review

The aim of this thesis is to ultimately facilitate better workplace communication, user interfaces, and database design and management. In order to do that, I borrow heavily from concept elicitation methodologies in order to produce personal constructs of data. These personal constructs of data, rendered in a concept map, allow for explicit exposition of the concept of data in a workplace and thereby reduce miscommunication through self-aware modification to available mental maps of the purpose and role of data.

Concept elicitation methodologies are a subset of knowledge elicitation methods, a tool used in many disciplines to "obtain the information needed to solve problems" [11]. Knowledge elicitation, in the main, is focused on direct problem solving: exploring requirements and understanding the meanings of those requirements. However, by turning the techniques of knowledge elicitation onto epistemological questions of category, we can discover not the direct meaning behind requirements, but some of a person's semiotic models of the constructions behind those requirements.

My research looks to investigate a person's personal construction of data. I borrow from data flow diagrams with a similar intent to the RepGrid methodology, though the end product differs significantly. The idea of personal constructs, discussed by Kelly and Tan [12 and 13] and reformulated under many names: Terms that have been used to describe these things include "schemas" [14 and 15] "cognitive maps" [16 and 17], "technological frames" [18], and "mental models" [19]. I, like Tan, will use personal constructions as the operative term.

Kelly [12] describes a personal construction as a combination of philosophy and

psychology. A construct, being subjective, is a personal epistemological tool of categorization and differentiation: "A construct is a way in which some things are construed as being alike and yet different from Os." His thesis denotes constructs as framing devices where we can situate objects-as-signs in our way of knowing. He continues, "We have departed from conventional logic by assuming that the construct is just as pertinent to some of the things which are seen as different as it is to the things which are seen as alike." Here, the fact that an object is not categorized as something can be an important factor in a person's personal construction of reality. Constructs are bipolar, admitting knowledge of the sign/concept and its opposite rather than simple negation. The SDFN extends this bipolar methodology of construction construction by asking people to categorize elements as data, information, or knowledge. By articulating a tripolar construction, we not only can articulate the positive categorizations of data, but can more closely examine data as it transforms into specifically delineated categories.

Much would be lost if participants were asked to categorize "data or not data" as the "not data" construction comprises everything that is not data, and is therefore not particularly interesting as a means of indicating the ontological and epistemological affordances of data. By requiring positive categorization, relationships between data and other concepts can be elicited more easily than simple negation would warrant. However, I also recognize that a a given categorization may simply be irrelevant in respect to data (relevancy is a far more useful and pragmatic benchmark than negation). Kelly notes that personal constructions are bounded [12], and are not necessarily "convenient" methods of categorization. In that light, the interview methodology will allow participants to articulate other categories that do not belong to the trinary construct of data-information-knowledge.

The repGrid [13], is a similar concept elicitation method. Tan describes the IST uses

of the technique as: "a set of procedures for uncovering the personal constructs individuals use to structure and interpret events relating to the development, implementation, use, and management of IST in organizations." While it is more overtly focused on organizational modelling, and the interpretation of events, it is a study of cognitive processes in an organizational setting to more effectively articulate information system requirements. The repgrid relies on participants sorting a pre-established schema of entities or objects, defined as a common set of "nouns or verbs" to constructs, the framing understanding around those concepts. Tan describes repGrid concepts as: "Constructs represent the research participant's interpretations of the elements. Further understanding of these interpretations may be gained by eliciting contrasts resulting in bi-polar labels. Using the same example, research participants may come up with bi-polar constructs such as "high user involvement – low user involvement to differentiate the elements (i.e., IS projects)." The creation of framing dichotomies echos the construct framework of Kelly and then allows users to sort elements within those constructs with a variety of different methods.

However, the repgrid is not the best tool for understanding constructions of data: while it does articulate a dichotomy, it fails to expose the manipulations attached to data. Elicitation of affordances and transformations of data is crucial to understanding a person's construction of data in sufficient detail to provide useful tools designed for them. Furthermore, while the statistical reliability of the repGrid is appreciated, especially as it can be subject to content analysis through simple frequency counting, the lack of an explicit period of participants to articulate their self-schemata robs interviewers of the potential insights of an articulated schema.

A representation grid draws on the personal construct framework for its own purposes of organizational knowledge modelling. In many ways, a "RepGrid" is a means

of evaluating a social construction of reality, as discussed by Berger and Luckman [20]. The social construction of reality echos the idea of personal constructions (though never explicitly calls out the term) by evoking the different realities of objects, "Different objects present themselves to consciousness as constituents of different spheres of reality. I recognize the fellowmen I must deal with in the course of everyday life as pertaining to a reality quite different from the disembodied figures that appear in my dreams. The two sets of objects introduce quite different tensions into my consciousness and I am attentive to them in quite different ways." This evocation of personal constructions framing the affordances of interaction was one of the other inspirations behind this project. While Berger & Luckman articulate the primacy of our shared reality, this investigation explores one area where that shared understanding may break down.

Shared understandings of reality as encoded as self-schemata and expressed as understandings of terms. While this practice should just as easily be expressed as a linguistic pursuit, the aim of this investigation is to uncover elements of that primal construction of reality, not in differences in linguistic expression of that construction. I have found that the best way to explore an individual's construction of reality is to ask them to express that reality in database design. The act of rendering the real-in-mind into diagrams expressing that causes an awareness of the self-schemata to coalesce simply by bring it into the forefront of consciousness. Through introspection into cognitive activity, self-schemata are formed: "attempts to organize, summarize, or explain one's own behavior in a particular domain will result in the formation of cognitive structures about the self or what might be called self-schemata. Self-schemata are cognitive generalizations about the self, derived from past experience, that organize and guide the processing of self-related information contained in the individual's social experiences." [21] It is this very process which the creation of the data flow diagram occasions in regards to an

individual's data manipulation activities. Furthermore, it is this act of schemata creation and subsequent discussion that I aim to elicit with the SDFN.

The idea of schemata *qua* personal constructions of reality influencing human computer interfaces and system design is not novel. Though in the HCI field, the term "mental model" is used. Wilson and Rutherford were exploring this very topic in 1989. Specifically, while they identify a significant variation in the definitions of the term "mental model," they generalize the term to: "a representation formed by a user of a system and/or task, based on previous experience as well as current observation, which provides most (if not all) of their subsequent system understanding and consequently dictates the level of task performance." [22] The definitions they synthesize this from extend back into the seventies, and there is no fundamental disagreement that the practice of human-computer interaction is, in some way, the practice of presenting an interface to these mental models.

It is important to note that there are philosophical distinctions between the terms mental model, personal construction, and self-schemata. A personal construction is, in many ways, the philosophical reality of a term. The construction provides for understanding of when and how to use the term for all use cases as well as its personal and cultural semiotic identifications. A self-schemata is the articulated and explicit epistemological conceptions of the term: it is the developed understanding of an individual understanding how they categorize and use a term. A mental model, on the other hand, is the situated understanding in procedural memory. These mental model are, themselves, socially constructed through routines in organizations [23]. The mental model is the procedural manifestation of he personal construction in the recognized semiotic affordances of the concept of data.

Extending the mental model to expected manipulations of data, rather than ex-

pected interactions with a system is the providence of the DFD, though the DFD holds to an objective reality which synthesizes many mental models. The SDFN, therefore, is a way to inspect the subjective mental models of humans as they relate to the expected interactions and transformations that their world applies to the thing they call data. As the term is never formally taught, we must evolve our models by experience with the world. Rasmussen asserts that mental models evolve with world-experience: "A mental model of a physical environment is a causal model structured in terms of objects with familiar functional properties. The objects interact in events, i.e., by state changes that propagate through the system "Kelly argues that individuals use their own personal constructs to understand and interpret events that occur around them and that these constructs are tempered by the individual's experiences." [24]

As our experience with the world differs, so to must our models diverge to make individual predictions about the systems we encounter in our subjective, constructed, reality. Through articulated schema creation, we can expose a person's mental map in a sufficiently valid framework for database designers and philosophers to puzzle over.

5.1   Hierarchies of Data

This work is not the first to ponder the nature of data. There exist two significant and pre-established relationships of data to information and knowledge: Ackoff's and Tuomi's. My findings mostly tend to echo the realities of data described by Ackoff or Tuomi. While not every interview or survey articulates a hierarchical relationship between data, information, and knowledge, it is clear that Ackoff's work has entered the "common knowledge". A number of interviewees discussed a hierarchy of data first promulgated by Ackoff. While they never cite his influence, their descriptions of relationships between data, information, and knowledge match his quite precisely.

Interestingly, this hierarchy has a management basis, rather than one grounded in philosophy or practical understanding of how we appreciate and use data. There are many signs of cognitive dissonance between what is perceived as the traditional hierarchy and how data is used in practice. Ackoff has updated his hierarchy many times, including a book "Management F-Laws" [25] wherein he not only states that "to managers, a pound of wisdom is worth an ounce of understanding" but also belabors the useless metaphor with the extension that an ounce of wisdom is worth "65,536 ounces of data."

Tuomi presents a more philosophically rigorous conception of data in his cyclic hierarchy, with data feeding into information, which in turn feeds knowledge, which provides the ability to collect more data. In the analyses, I will refer to this as a cyclic hierarchy.

5.1.1 Ackoff's Traditional Hierarchy

The traditional hierarchy linking data, information, knowledge, and wisdom in a strict hierarchy of dominance and importance was created by Ackoff [26] in 1989. In a summary of his work, Bernstein notes [27]:

> Ackoff was a management consultant and former professor of management science at the Wharton School specializing in operations research and organizational theory. His article formulating what is now commonly called the Data-Information- Knowledge-Wisdom hierarchy (or DIKW for short) was first given in 1988 as a presidential address to the International Society for General Systems Research. This background may help explain his approach. Data in his terms are the product of observations, and are of no value until they are processed into a usable form to become information. Information is contained in answers to questions. Knowledge, the next layer, further refines information by making "possible the transformation of information into instructions. It makes control of a system possible" (Ackoff, 1989, 4), and that enables one to make it work efficiently. A managerial rather than scholarly perspective runs through Ackoff's entire hierarchy, so that "understanding" for him connotes an ability to assess and correct for errors, while

"wisdom" means an ability to see the long-term consequences of any act and evaluate them relative to the ideal of total control (omnicompetence). While a scholarly perspective on this hierarchy might prioritize the processes of inquiry and discovery, Ackoff does not account for them. But his concept of omnicompetence, which refers to "the ability to satisfy any and every desire" (Ackoff, 1989, 8), does encompass the satisfaction of user-defined needs.

In this ontology, data are subjective observations. Curiously, despite data being subjective observations, Ackoff does not suggest any need for filtering (a common theme in subjective/observation conceptions of data).

5.1.2   Tuomi's Cyclic Hierarchy

Tuomi's ontology is simple and counter-intuitive: knowledge is a framework of the world from which we build information. Information provides a local framework from which to extract data from the world. Thus, the apex of the hierarchy is data, which then filters downwards to modify knowledge and information. This approach represents an abductive approach towards the philosophy of knowledge and the research about data as the hypotheses being tested are information, generated from knowledge-of-world, rather than induced from data points [28]:

The generally accepted view sees data as simple facts that become information as data is combined into meaningful structures, which subsequently become knowledge as meaningful information is put into a context and when it can be used to make predictions. This view sees data as a prerequisite for information, and information as a prerequisite for knowledge. ... [Exploring] the conceptual hierarchy of data, information and knowledge, showing that data emerges only after we have information, and that information emerges only after we already have knowledge.

Tuomi's conception of a reverse hierarchy is useful to my research in two significant ways. Obviously, it allows me a prior idea to which to compare analyzed hierarchies of data. Although Tuomi's research is not explicitly about user's conceptions of data, his analysis of hierarchies is an acceptable complement to it. By presenting a novel

relationship hierarchy, Tuomi challenges the "everyone knows" mentality of much of knowledge management.

Tuomi's research into the fundamental questions of knowledge is one of the fundamentals of my research, for he demonstrates that it is possible to have a different understanding of data from an intuitive-philosophical standpoint. This demonstration of difference allows a questioning of the nature of data and acts as an external source of validation for my analysis.

This ontology of data is supported by a study of "intelligence" published in *Nature* by McNab and Klingberg [29]:

> Thus, high-capacity individuals (who can remember more information at once and who tend to do better on aptitude tests) might simply be better at keeping irrelevant information "out of mind," whereas low-capacity individuals may allow more irrelevant information to clutter up the mental in-box. The difference may just be a matter of having better spam filters.
>
> Some of our own recent work on differences in controlling access to working memory has provided evidence favoring this mental spam-filtering idea. In one experiment, measuring electrical signals emitted by the brain enabled us to show that high-capacity people were excellent at controlling what information was represented in working memory: they let in information about relevant objects but completely filtered out that about irrelevant objects. Low-capacity individuals, in contrast, had much weaker control over what information entered the mental in-box; they let in information about both relevant and irrelevant objects roughly equally. Surprisingly, these results mean that we found that low-capacity people were actually holding more total information in mind than high-capacity individuals were-but much of the information they held was irrelevant to the task.

The idea of consciousness as filter discussed by all of these researchers is not particularly novel, although the localization of filtering activities by fMRI to those physical regions of the brain suggests that this ontology has a closer connection to our biological minds than does Ackoff's.

5.2   Links with the Literature

Supporting the research of Tuomi and Ackoff are Zins, Galison, and Voloshinov. Dr. Chaim Zins, whose surveys in many ways validated Ackoff's "standard" hierarchy with his experiments, despite showing that there are many definitions of data, information, and knowledge. Galison contributes the idea of Trading Zones, where people with different constructions of concepts can come together and generate local areas of meaning to facilitate communication. A means for the different meaning of the same words is shown by the evaluative accents of Voloshinov and supported by the modern account of social steagnography by Boyd within the framework of Percian semiotics.

5.2.1   Zins' concepts

Zins performed a "collective knowledge mapping" of a number of researchers using a critical Delphi methodology over three rounds of research [30–33]. Although his approach looked at definitions directly, he found two different concepts of data and a similar ontological split over subjectivity, objectivity, and communication.

While I have utilized his works for justification of my literature, I did not closely examine his conclusions to avoid biasing my analyses. Zins, in 2003, identified areas of difference similar to those I identified [32]:

> Six distinctive concepts. Having established the distinction between the subjective and the universal domains, we are in a position to define the three key concepts data, information, and knowledge. In fact, we have six concepts to define, divided into two distinctive sets of three. One set relates to the subjective domain, and the O- to the universal domain. [Data-Information-Knowledge] in the subjective domain. In the subjective domain, data are the sensory stimuli, which we perceive through our senses. Information is the meaning of these sensory stimuli (i.e., the empirical perception). For example, the noises that I hear are data. The meaning of these noises (e.g., a running car engine) is information. Still, there is another alternative as to how to define these

two concepts-which seems even better. Data are sense stimuli, or their meaning (i.e., the empirical perception). Accordingly, in the example above, the loud noises, as well as the perception of a running car engine, are data. Information is empirical knowledge. Accordingly, in the example above, the knowledge that the engine is now on and the car is leaving is information, since it is empirically based. Information is a type of knowledge, rather than an intermediate stage between data and knowledge. Knowledge is a thought in the individual's mind, which is characterized by the individual's justifiable belief that it is true. It can be empirical and non-empirical, as in the case of logical and mathematical knowledge (e.g., "every triangle has three sides"), religious knowledge (e.g., "God exists"), philosophical knowledge (e.g., "Cogito ergo sum"), and the like. Note that knowledge is the content of a thought in the individual's mind, which is characterized by the individual's justifiable belief that it is true, while "knowing" is a state of mind which is characterized by the three conditions: (1) the individual believe[s] that it is true, (2) S/he can justify it, and (3) It is true, or it is appear to be true.

[Data-Information-Knowledge] in the universal domain. In the universal domain, data, information, and knowledge are human artifacts. They are represented by empirical signs (i.e., signs that one can sense through his/her senses). They can take on diversified forms such as engraved signs, painted forms, printed words, digital signals, light beams, sound waves, and the like. Universal data, universal information, and universal knowledge mirror their cognitive counterparts. Meaning, in the objective domain data are sets of signs that represent empirical stimuli or perceptions, information is a set of signs, which represent empirical knowledge, and knowledge is a set of signs that represent the meaning (or the content) of thoughts that the individual justifiably believes that they are true.

Signs Versus Meaning. Defining the Data-I-Knowledge phenomena as sets of signs needs to be refined. There is a fundamental distinction between documented (i.e., written, spoken, or physically expressed) propositions and meanings. "E = MC2","E = MC2", and "E = MC2" are not three different types of knowledge. These are three different sets of signs that represent the same meaning. In other words, they are three different utterances of the same knowledge. Knowledge, in the collective domain, is the meaning that is represented by written and spoken statements (i.e., sets of symbols). However, because we cannot perceive with our senses the meaning itself, which is an abstract entity, we can relate only to the sets of signs (i.e., written, spoken, or physically expressed propositions), which represent it. Apparently, it is more useful to relate to the data, information, and knowledge as sets of signs rather than as meaning and its building blocks.

My work profoundly agrees with the discoveries he made, though my research focuses far more on data and differentiates three different orders of data to his two.

5.2.2 Trading Zone

Also of interest is the way that the interviewees demonstrated the idea of a trading zone.

As Galison defines it [34],

> These considerations so exacerbated the problem [of physicists communicating] that it seemed as if any two cultures (groups with very different systems of symbols and procedures for their manipulation) would be condemned to pass one another without any possibility of significant interactions. Here we can learn from the anthropologists who regularly study unlike cultures that do interact, most notably by trade. Two groups can agree on rules of exchange even if they ascribe utterly different significance to the objects being exchanged; they may even disagree on the meaning of the exchange process itself. Nonetheless, the trading partners can hammer out a local coordination despite vast global differences. In an even more sophisticated way, cultures in interaction frequently establish contact languages, systems of discourse that can vary from the most function-specific jargons, through semi-specific pidgins, to full-fledged creoles rich enough to support activities as complex as poetry and metalinguistic reflection. The anthropological picture is relevant here. For in focusing on local coordination, rather than on global meaning, one can understand the way engineers, experimenters, and theorist interact. At last, I come to the connection between place, exchange, and knowledge production. Instead of looking at laboratories simply as the places at which experimental information and strategies are generated, my concern is with the site – partly symbolic and partly spatial – at which the local coordination between beliefs and action takes place. It is a domain I call the trading zone.

The requirement of locally true definitions applies across the original trading zones between cultures and, more interestingly, to the various cultures of physics. In the interviews, I noticed some evidence for trading zones in the interview material. Specifically, when the various participants referred to the terms "raw data" and "derived data," they seemed to be using a local definition of data that did not correspond with their own philosophy, strictly speaking. Instead, they were referring to various sensor products that were, indeed, "raw data" to every member of the team.

The extension of the Galisonian trading zone concept is not new to this research. In fact, business researchers have used the idea of trading zones and some sophisticated

ideas of boundary demarcation for quite some time, going so far as to use graphs and knowledge maps (as opposed to my SDFN) to identify different groups. Wilson and Herndl use this methodology when they describe their understanding of knowledge maps and trading zones [35]:

> The knowledge maps we created and shared with project participants encouraged cooperation and mutual understanding rather than the slash-and-burn rhetoric of demarcation events. When technical experts discuss the parts and subfunctions they have made, they get to describe their local practice, explain their knowledge, and open up their community-specific lexicon within the ecological relations of the boundary object. As they trace the lines connecting the boxes on the knowledge map, participants articulate communities of practice: each distinct but also connected through the boundary object. Because it is plastic and robust, the knowledge map balances the demands of identification and division in Burke's terms. As boundary objects, the knowledge maps help to create a rhetorical space that is best understood through Galison's notion of the trading zone.

This methodological description of their work is focused in the anthropological study of finding sub-cultures, rather than language differences. Despite this, my analysis produced many of the same results as theirs (methodologically speaking, if not with respect to content) because both my research and theirs tried to understand different philosophies/cultures with the metaphor of trading zones.

Just as they extend the concept of trading zone to consultants on the Washington beltway creating local definitions for a program already approved by the Pentagon, I extend the idea to how different conceptions of data interact. Their use of Graphviz-generated graphs as talking points to determine cultural ramifications matches my experience of using said graphs to generate philosophical insight:

> In the case we have been exploring, the knowledge map is crucial to the emergence of something like Galison's (1997) trading zone. Participants develop Galison's "possibility of communication and joint action" (p. 803) through the map as it emerges. The map continually structures how the team understands and explains the project.

The differences between their study and mine are quite pronounced, although just

as they observed different groups creating temporary trading zones through the use of knowledge maps, I observed something similar with the pidgin concepts of "derived data" and "raw data." While waiting for interviews, I saw researchers passing around sheets of paper with pictures of phenomena on them and referring to them explicitly as raw data. This practice almost certainly serves to inform the local definitions and create the trading zone necessary for successful research practice.

5.2.3 Evaluative accents

An evaluative accent is the set of interpretive filters a recipient applies to incoming communication, thereby changing its meaning based on the biases applied by the recipient, shared understanding, and cultural mores. It was originally used to explore the effects of Marxist propaganda, but it can also be an interesting way to explore how trading zones operate effectively.

> V. N. Voloshinov suggests an idea of an evaluative accent [36]:
>
>> Any word used in actual speech possesses not only theme and meaning in the referential, or content, sense of these words, but also value judgment: i.e., all referential contents produced in living speech are said or written in conjunction with a specific evaluative accent. There is no such thing as word without evaluative accent.
>>
>> What is the nature of this accent, and how does it relate to the referential side of meaning? The most obvious, but at the same time, the most superficial aspect of social value judgment incorporated in the word is that which is conveyed with the help of expressive intonation. In most cases, intonation is determined by the immediate situation and often by its most ephemeral circumstances. To be sure, intonation of a more substantial kind is also possible. ...
>>
>> All six [uses of a single word in a removed quote] by the artisans are different, despite the fact that they all consisted of one and the same word. That word, in this instance was essentially only a vehicle for intonation. The conversation was conducted in intonations expressing the value judgments of the speakers.
>
> Beyond the verbal intonation is context and use. In a more modern sense, people

of different generations use what amounts to steganographic[*] encryption in their status messages, relying on context and the source of quoted material to produce different meaning for different people.

An instance of this social steganography appears in the following example to pass different meaning through a Facebook post to the subject's friends and mO:

> When Carmen broke up with her boyfriend, she "wasn't in the happiest state." The breakup happened while she was on a school trip and her mother was already nervous. Initially, Carmen was going to mark the breakup with lyrics from a song that she had been listening to, but then she realized that the lyrics were quite depressing and worried that if her mom read them, she'd "have a heart attack and think that something is wrong." She decided not to post the lyrics. Instead, she posted lyrics from Monty Python's "Always Look on the Bright Side of Life." This strategy was effective. Her mother wrote her a note saying that she seemed happy which made her laugh. But her closest friends knew that this song appears in the movie when the characters are about to be killed. They reached out to her immediately to see how she was really feeling.

The use of "Always Look on the Bright Side of Life," as Boyd discusses, is an example of a successful steganographic encoding of a message. Her friends could decrypt the hidden message because they shared a private context of culture with Carmen, a shared evaluative accent [37].

Although that process is called social steganography, its unintentional practice causes gulfs in evaluative accent. Failed obscure jokes are an example of an incorrectly parsed communication. The obscure joke, in this case, relies on a shared commonality to be correctly anticipated by the recipient, and this mode of receptive listening is informed by the evaluative accent. In more common use, the language in a business memo may be so full of "business-speak" that someone who is not used to the company may

---

[*] Steganography is the act of hiding messages within other messages, where, only if you know the pattern or encoding scheme, can you identify the hidden message.
Social steganography is the use of shared context to provide polymorphic (meaning-changing) meanings to one's social communications, depending on context and other available social cues.

misunderstand the provided references. This misunderstanding is especially deadly if it makes sense within the reader's incorrectly applied accent. As the statement can be parsed by the listener, only the mismatch of reality models in later conversations can hint at the source of the problem: the misinterpreted statement.

Voloshinov believes that the evaluative accent partially belongs with the speaker, but also that there exist "side bands" of communication, such as intonation and body language, that are specifically interpreted by the recipient of any communication, in the recipient's context.

Literal intonation has very little to do with specific constructions of data. Yet the term is used in everyday, technical, engineering, and scientific speech. The full weight of the evaluative accent, as seen in the interviews, falls into the use of context and role. While originally it was seen as way to frame ideologies; almost a post-hermeneutic way of explaining some failures of Marxism [38]. The idea of an evaluative accent can be combined profitably with the philosophy of the trading zone.

The idea of an evaluative accent corresponds well with the idea of a trading zone. The construction of misunderstanding the same language could hardly be the result of a simple linguistic misunderstanding. When a data-as-subjective-observation person says "data" to a data-as-objective-hard-numbers person, both of them are using a "functionally correct" definition, a definition shared by many people. They are encountering a trading zone. As Galison states, "In the trading zone, where two webs meet, there are knots, local and dense sets of quasi-rigid connections that can be identified with partially autonomous clusters of actions and beliefs." [34] My identification of different conceptions of data certainly corresponds with these diverse beliefs. And those beliefs inform the evaluative accents that people use when they use and receive the term "data."

While different uses of data may be boundary objects for more profound cul-

tures, [39–42] this minimal investigation can scarcely provide an anthropological look into the various research cultures in existence. I can present research into the intentional creation of a local language. The use of "raw" and "derived" as a semiotic prefix presents a linguistic indicator to switch evaluative accents to recipients who are *aware* of that indicator. The practice of forming local trading zones, by repeatedly presenting symbols to the other party in an environment where people are aware that bridging must occur [34], is the non-ideological practice of causing sufficient cognitive dissonance in the recipients for them to "bud off" a new evaluative accent for interpreting the incoming sign-set.

The evaluative accent of the local definition of a word can be understood in the context of Peircian semiotics. As Aktin notes [43]:

> In Peirce's theory the sign relation is a triadic relation that is a special species of the genus: the representing relation. Whenever the representing relation has an instance, we find one thing (the 'object') being represented by (or: in) another thing (the 'representamen') and being represented to (or: in) a third thing (the 'interpretant').

The interpretant serves as the developed representation of meaning. Thus, we can understand local trading zones to be the product of evaluative accents present in the interpretant [44].

5.3   Justification

IST both suffers and benefits from its multidisciplinarity. One of the key tools taught in IST-training programs is merely the ability to understand the jargon of the other sub-disciplines of IST. As a database professional, I can still speak the terminologies of networking, web design, and enterprise systems. The process of communicating in these various jargons, however, necessitates different mental models of the world[*]. In addition,

each of those individual levels will have its own uses for the term "data". The field is not amenable to a single probe, and even if it were, each sub-discipline understands reality in its own way, as it must to solve problems according to the constructed protocols of that profession.

One thing underlying all of IST, however, is the use of the term "data." Every aspect of IST uses data, but their understanding of what constitutes that data is significantly different. Moreover, in this computerized age, everyone interacts daily with data to some degree. The difficulty is in the question: what is data?

There is a need for to understand how people understand data because conflicting definitions of "data" inform communications. Peoples' inherent conceptions of data inform how they interact with the constructed data of the world[*]. Some people consider data to be objective facts, Os consider it to be subjective observations, and still Os consider it to be electronically stored signs[†].

When people discuss data, information, and knowledge, their understanding of data informs their understanding of information and knowledge - be it synonymous, Ackoff's hierarchy from 1989, or any of the other hierarchies suggested by the literature[‡]. When these different understandings collide, the best case is that the people involved recognize that they have different understandings and create a local trading zone with words that have functionally identical meanings to both people. In the worst case, both people use the term in the way to which they are accustomed, and errors go uncaught

---

[*] Exploring computer problems and their solutions is an exercise in quickly changing levels of abstraction. Small, vital, and technical details fight tooth and claw against the broad vision of the designer [45].

[*] From a pragmatic point of view, many linguistic elements are socially constructed and our understanding of them is shaped by our linguistic interactions with other people. Data, being something categorized by humans, is a great example of a linguistic construction [20].

[†] For more details, see the results section, page 274.

[‡] For more details, see page 27.

until large mistakes are made.

Data is defined by its use. It is a socially constructed term[*] rather than a reflection of some property of the universe. Therefore, data is subjective relative to the person using the term. I have identified a need to probe other peoples' understandings of data. It is easy to mistake professional training as a single, true, definition of data. The problem with intuitive definitions is that their elegance may not be used or tested in reality. For research to be useful to practitioners, it must deal with the philosophical problems that they face, not add another definition onto the large heap. This research aims only to provide a tool and a reason for practitioners to use that tool.

5.4   Aims

I want to help improve communications, and I believe that a means for understanding different constructions of data could be one way to do so that has not been thoroughly explored. It may offer a theory towards explaining some errors in intergroup and intragroup communication. Furthermore, it may offer some direction towards exploring the philosophical basis of error by offering another take on system maps transmitted through communications[†].

I will, while exploring our ability to define and communicate data to people around us, lay a foundation for the exploration of the reality created by our use of data in computer systems. Our systems use data in multiple levels, from the hardware and simulated hardware through software and into fantastic constructions and games that

---

*   While the term Data, as language is socially constructed, there are a large number of people who feel that the content of Data, as measurements of reality, cannot be so constructed [46].
†   A system map is simply a person's internal mental model of how a thing operates and of how to get it to transition between different states. These maps may be communicated through instruction or alluded to [47].

embody and then produce their own data from any philosophical meaning. This study will not explore the various sub-constructions of data present on the Internet, in games, or in virtual worlds. Nevertheless, I hope that the methodology I create and validate can be applied to all sorts of computerized data: from the traditional bits down a wire into a simulation of a physics experiment inside Second Life [48]. As humans use and create all of these tools, our constructions of reality inform them. To serve that end, this research creates and tests a methodology that can probe peoples' understanding of data.

In database design, the hardest task is trying to understand the client's reality. Modeling a current organization's memory structures, its files and paperwork, and the relations among them in the minds of practitioners is an extremely difficult task. To facilitate understanding, this methodology is a tool for designers. The tool may allow them to understand what their clients think data actually *is*.

By understanding the type of data being modeled, database designers make two significant gains. First, their data models can correspond with how their clients think about reality, and thereby create intuitive relationships and map the computerized model to their client's mental model more capably. Second, and in some ways more critically, they can then explain the database design *to* their clients in their clients' language, potentially shortening design times by reducing miscommunications.

In the same way, the proposed methodology should help extend normal modeling practice: simply making designers more aware of the different types of data constructions may make more responsive designs possible. The demonstration of different conceptions of data is important to designers because it offers another meta-aspect of reality to be captured and incorporated.

I also want to create a method that can help extend HCI design practice. This methodology should be applicable to all sorts of design, as it is a tool for rendering

clients' realities and not a specific kind of technical reality. The discovery of practical meaning of terms, ideas, and affordances [49] of data is another tool with which HCI designers can understand how to render data presented in an interface. A tool that can make elements of private jargon explicit, and that is focused on that task (rather than treating it as a happy byproduct) can significantly contribute to the HCI design cycle.

This research investigates individual constructions of data, because there is no clear consensus on the exact nature of data, much less on the exact nature of data in technical design. However, as there is no recognized domain of the philosophy of data, this research, as a more practical matter, must lay the simplest foundations for that multidisciplinary field.

My basic discoveries, both methodological and philosophical, should have pragmatic results. I hope to create a methodology that improves communication and database design. I explore how we socially construct and use the term "data". From this investigation, I can offer potential insights into how we create trading zones between different cultures of data use. While true understanding of the nature of data may be outside the scope of this present research, the construction of a foundation is not. Any methodology created must be robust enough to provide useful observations and a compelling story.

6 Methodology

The field of requirements generation is heavily overpopulated with methodologies. These methodologies on the whole, generally presume that the participants are attaching similar meaning to the terms they use, especially when the terms are seemingly uncomplicated ones like "data", "information", and "knowledge." The Social Data Flow Network is an elicitation methodology that can be used prior to normal requirements generation. This methodology helps to map the shared and unshared components of a group's social construction of reality as it relates to data flows.

The SDFN has its methodological roots in the data flow diagram (DFD). Information technologists currently use a DFD as a tool for probing current data flows within an organization. I have designed the SDFN as a compliment, allowing a practitioner to uncover an individual's subjective constructions of data, one unburdened by the methodological constraints of the DFD. The SDFN can be used as an artifact for sparking discussion around practical definitions without the investigator having to enter the interview and ask participants about their personal constructions of data directly. Once data has been characterized by the participant, other requirements generation methods can then be employed to extract a formal understanding of what is needed, paying special attention to where different individuals understand the components of the same process differently.

By allowing people to probe their own constructions of data, the SDFN helps them to express their own understanding in their own language without worrying about being judged incorrect. By creating a sense of cognitive dissonance[*] between the participants'

---

[*] While Festinger's original work is important here, I believe that this model represents the satisfaction of

application of categories and their theoretical definitions, the methodology discussed in this section will serve as a way to illuminate how people understand the nature of data. It seems quite feasible to extend this methodology to other research endeavors.

6.1   The Social Data Flow Network

This chapter documents the methodology I used in conducting the interviews with the company. This section is organized first into definitions, an exploration of the SDFN as a concept, and then a practical discussion of running an interview centered on the SDFN.

6.1.1   Terms

This section will introduce the major terms of the SDFN and how those terms are used. The introduction of a new methodology, especially one borrowing from many different fields, is fraught with definitional dangers.

An entity is a noun: a role that can manipulate data. A flow is a noun, representing the *flow* of communication or symbols between entities. An entity dictionary is a way of brainstorming entities to make the participant feel more at ease.

6.1.1.1   Entity

An entity is something that plays a role receiving, manipulating, or transmitting data. In the SDFN, this act of input or output is represented by a noun described in a few words, which then have an oval bubble drawn around them. This bubble is a node in graph theory, with all of the corresponding attributes. The nature of the role is not restricted

---

constraints imposed by the categorization of terms as per Shultz [50 and 51].

to a person or physical entity. It is anything that can be made-a-thing-of such that it makes an independent manipulation of data.

Roles are anything that can be conceptualized as an independent manipulator of data. What differentiates an interesting role from one worth skipping over is whether the role somehow transforms data passing through it. There is no restriction on the number of roles that can belong to one person or thing. Just as one person can do multiple jobs, one can also have multiple roles. One hypothesis for future testing is that the role determines the perceived affordances of data. Every role has its own unique activities and therefore uses data in its own way. As the framing of the role changes, the definition of data may change along with it.

Participants should never describe themselves as a singular entity due to ambiguity. Their description of an entity as "self" is ambiguous to other people reading the chart, who do not understand the tacit assumptions of role and interaction from the same perspective as the participant. Instead, participants should articulate the potentially many roles that they play in an organization as separate entities. While participants self-articulating roles adds a certain artificality to the interview, the self-identification of roles also allows participants to adopt some of the framing of those roles. It is thought that the increased precision gained from artificial role distinctions is worth the contrived nature of the process.

Every role should be unique. However, there is no requirement for a one-to-one mapping from person to role. Because people and things are adaptable and can serve many roles in an organization, artificially forcing the participant to select one and only one definition of self would be contrary to the intent of this exploration. This requirement allows and encourages people to represent passing information to themselves in the guise of the different roles they play.

The avoidance of ambiguity is crucial. It makes the SDFN easier to interpret by other people and it forces the individual creating the diagram to define the nature of the entity explicitly. It is far too easy to use the self as a catchall to avoid the cognitive dissonance of thinking about thinking. It is important to document discrete and unambiguous roles, even though they may map to the same person, because it is the role that understands data, not the person. Furthermore, these different roles-as-self can pass data to one another. I used the following example in the interview: An entity as lecture designer (myself) would pass requirements to the database lab developer (myself) who would pass data to the lecturer (myself). Each of those roles has different requirements for the nature of data. Crucial insights would be lost if they were all collapsed into one entity with self-pointing flows of data.

Entities however, should not be ready-to-hand[*]. Devices that take on independent roles are fundamentally different from those that function as parts of another entity. The keyboard used for typing these letters into this document should not be considered an entity in the SDFN sense because except when engaging in this self-reflexive behavior, the entity "author of dissertation" does not explicitly pass data to the keyboard-rather, the "author of dissertation" passes data to the computer for processing. The keyboard is part of a large entity and does not manipulate data in my own construction of data. Instead, as an input device, the keyboard is an extension of the computer and represents an interface for the electronic recording of symbols.

At the same time, entity creation rules should not be strictly enforced, as each person may have his or her own conceptions of what an entity could be. A role can be a

---

[*] Ready-to-hand roughly means tools that form an unconscious extension of the self. However, I will avoid a discussion of Heideggerian Daesin and other terminology. To learn more: Dreyfus's discussion of Heidegger is not too painful (p.230 for ready-at-hand) and Marshall is using the idea in interface research [52].

person, machine, place, or group. An entity is any noun that the interviewee regards as accepting or receiving data meaningfully. Participants must define their own entities, as their own conceptualization of roles is one of the strongest sources of insight into their understanding of their construction of data.

6.1.1.2 Flow

A diagram consisting solely of entities, known as an entity dictionary, is not particularly useful. To represent relations among these nouns, however, we need flows. A *flow* indicates a transfer of something between one entity and another. We are concerned with the nature of the transfer instead of the act of the transfer, a verb describing how the transfer is accomplished is not particularly useful during categorization.

This expressed relationship, usually, will be a self-categorized flow of data, information, or knowledge. These flows are edges, represented by arcing lines between one or more entities, although most flows link one entity to another, singular entity. There is little objectivity in these indications of relationship. A flow represents a documented relationship, instantiated from the recipient's understanding of reality, not necessarily a true thing in the shared reality of all participants. As the SDFN is intended as a tool for exploring constructions of data there is little need to find a design that corresponds to the real world and the stakeholders' needs. On the contrary, the subjective expressions of reality can be compared against each other to identify where areas of miscommunication arise.

Practically speaking, flows must be represented as arcing lines between one or more entities. The arc allows readers to differentiate the labels of each flow, with a clear distinction between the over and the under component. Recursive flows, which link an entity to itself, are discouraged, as they tend to represent ambiguous and broad entities.

Participants selecting recursive flows should be encouraged instead to delineate the starting and ending roles as entities more clearly.

Each flow has a label and a category. The label describes the content of the flow. The category relates the flows to other flows and ideas in the diagram. Each flow, above the arc, should be labeled with the *contents* of the flow. The label is a one- to three-word description of the "stuff" being transmitted. This description must be unambiguous and unique to the contents of the flow. If two entities are transmitting the same content, care should be taken to ensure that the exact same thing is being transmitted. Minor content variations should be indicated by adding adjectives or other modifiers to the name: "Results" becomes "Summary of Results" and "Formal Results." Each label indicates a result being transmitted, but the different nature of the things changes the understanding of the thing. Reducing ambiguity is the job of the interviewer and is one of the hardest parts of conducting a SDFN session.

For practice in clarifying the nature of results and in the type of thinking needed to conduct an SDFN session, I recommend the game called "Zendo" by Looney Labs[*].

---

[*] The rules of Zendo can be found here: http://www.koryheath.com/games/zendo/
The essence of the game is that players, through the use of transparent colored pyramids, must use inductive logic to find a "secret rule."
An example of a secret rule is "A [set of pieces] [is true] if it has at least one green piece." And through rating constructions of their own true and false the leader of the game describes a universe of discourse with the secret rule as the governing element.
The critical element of the game, for purposes of this research, is that the leader of the game must, by the rules, refine ambiguity from any guesses the players may make. "Clarify the Guess: If the Master does not fully understand your guess, or if it is ambiguous in some way, the Master will ask clarifying questions until the uncertainty has been resolved. Your guess is not considered to be official until both you and the Master agree that it is official. At any time before that, you may retract your guess and take back your stone, or you may change your guess. If any koan on the table contradicts your guess, the Master should point this out, and you may take back your stone or change your guess. It is the Master's responsibility to make certain that a guess is unambiguous and is not contradicted by an existing koan; all Students are encouraged to participate in this process." The process of clarifying guesses to eliminate ambiguity is exactly identical to clarifying entities and the labels of flows. Besides being a fun game, it is crucial practice to get a feel for the level of precision required in the SDFN.

Specifically, if one can run multiple sessions of the game successfully, the same skills in clarifying statements and assessing the nature of things will be used in this methodology.

The core of the SDFN is the process of categorization. The SDFN encourages participants to discriminate and categorize flows. By relating different flows through the use of category, it is then possible to induce the definition of the category through its flows. The category should be written under the arc. In computerized renderings, the over/under distinction is less important, so long as the label and category of the flow are clear.

When creating the flow, the participant should first be prompted to label and then to categorize the flow according to a pre-formulated short list of categories. This list of categories should contain the most common expected categories of participants. By prompting the participant with a list, the interviewer focuses the categories on the topic of the participant's choice. However, participants should always be able to add their own categories to this list. For example, in the interviews, I always prompted participants with "Data, Information, Knowledge, or other."

There should always be the option for other. But the other category should never remain as other; the participant should name it. Some participants used categories such as Emotion, Wisdom, or Request. In no case was a flow allowed to remain other. These new categories were created on the fly and used as part of that participant's diagram from then on.

At the same time, people should not classify their own domains without any initial guidance. All but the most self-reflective will be paralyzed by the many choices available and not entirely clear on the distinctions the interviewer wants them to draw. Thus, my question took the form of "Data, Information, Knowledge, or O" rather than "How would you categorize this flow: data or not-data?" Denoting sample categories creates a negotiable universe of discourse for the categories.
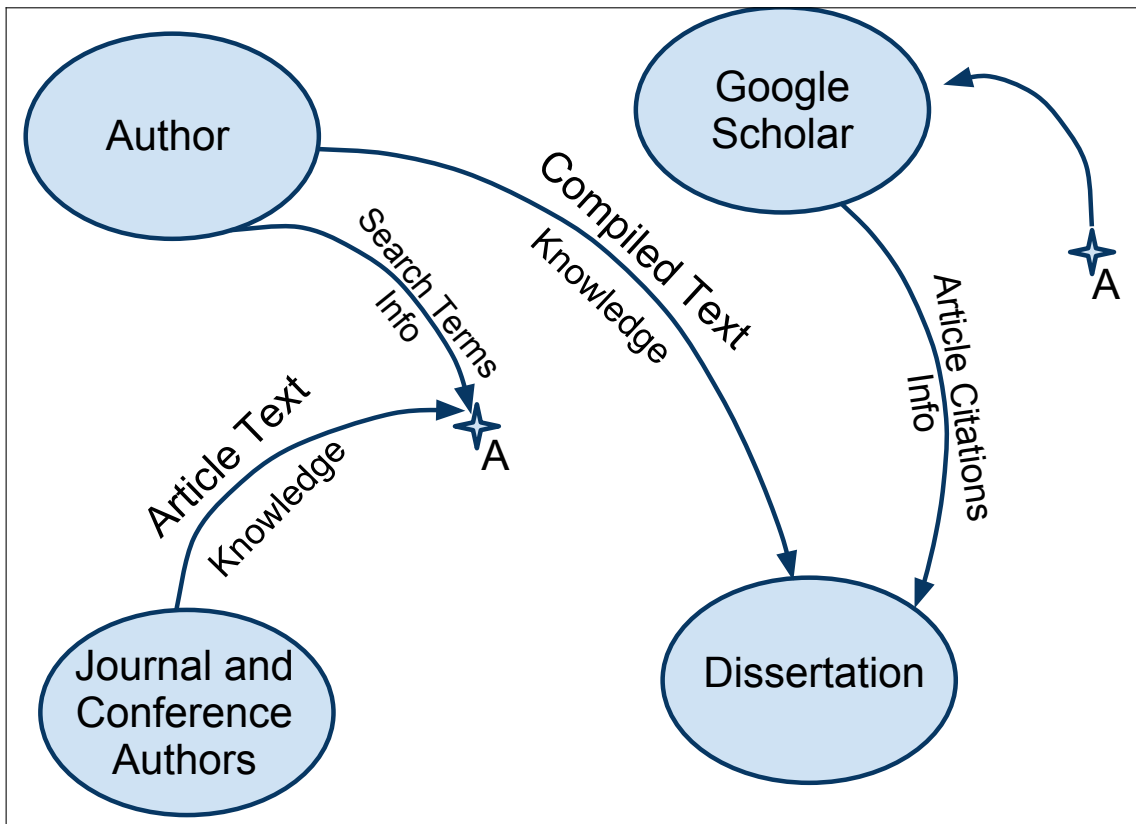
Figure 6.1:   A trivial SDFN used to illustrate the idea of "flows" and worm-holes. Crossed lines become unbelievably messy, and so the "wormholes" are a far better method for routing lines across other lines.

If a diagram becomes too crowded, it is quite acceptable to make "wormholes" on the paper design during the interview. A wormhole is some symbol (usually an *) and accompanying identifier[*] placed more than once on the paper. Each symbol sharing an identifier should be considered connected, which may allow for easier routing. In extreme cases, a wormhole may have one arrow leading from it to represent the inbound connections of all of the flows connected to the other wormhole. The only real restriction is that the creation of wormholes should be unambiguous and clear both at the time of creation and afterwards. Sometimes, in the case of major changes, it is better to redraw

---

[*]   A character, number, or unique symbol all serve well.

the design than to use too many wormholes. This action is sometimes quite desirable, due to the edits that the participant may introduce in the entities, flows, or topology on the second draft.

6.1.1.3 Entity Dictionary

Entities and Flows are the core parts of any SDFN diagram. However, not all participants may have the ability to easily understand the nature of entities. For that reason, and as a precursor to group-based SDFN creation, I engaged some participants with creation of an entity dictionary, a simple list of entities that may be involved in the SDFN.

An entity dictionary is a simple, non-authoritative, brainstorming device in case the participant is unsure about where to start. Instead of starting the SDFN with two entities and a flow connecting them, I will encourage the participant to imagine all the different entities with which they engage on a daily basis, to name them, and to describe their roles. The immediate feedback, both positive and negative, on each described entity teaches participants to think in terms of roles. Once they have filled a page, most will have internalized the meaning of entity.

Through the creation of this entity dictionary, a number of interesting themes will appear, based on participant enthusiasm or repetition. I was especially careful to pay attention to offhand comments about entities or the participant's work during the creation of the dictionary, as these comments will most likely indicate interesting topics for the interview. The dictionary should be started by encouraging each participant to name an entity that represents them in some role, and then the scope should be gradually broadened to things and people they work with.

To those familiar with the DFD methodology, the idea of the entity dictionary is

almost completely opposite to that of the "data dictionary" of the DFD. While the Data Dictionary is a device for the accurate specification of data in the data flows, compiled during and after the creation of the diagram, the entity dictionary is a piece of scaffolding designed to help participants think the right way about entities.

Unlike a data flow diagram, the entity dictionary is not authoritative[*]. In a DFD, all flows must be decomposed[†] to their atomic definitions[‡], which correspond with database or programming structures. This requirement exists because the DFD has its roots as a programming design, and therefore must be able to explicitly define the data structures of a program. Because the SDFN is probing a non-computerized theoretical area, the requirement of precision is unnecessary and counter-productive, as it distracts the participant from their task. The object of the SDFN is to probe functional definitions, not to have all participants arrive at the same constructed definition of the UoD.

6.1.2 Creating the SDFN

This section will describe the process of creating a SDFN in full. This describes the methodology used in the study presented in the next chapter. In brief, the SDFN begins through the explanation of terms, a summary of the ideas expressed above. If participants do not understand the nature of entities, an entity dictionary should be created. When participants understand entity and flow, a topic is chosen and the diagram is created.

---

[*] Authoritative: a canonical listing and extremely precise description of the structure and components of variables.

[†] Decomposed: simplified by breaking the components of a flow (or transformation) apart into separate components. An example of a decomposition may be an "Address" flow, that is subsequently decomposed into 4 flows "street address + city + state + postal code" In the same way, a transformation can be decomposed. "Mail a letter" can be decomposed into "Look up address -> Find Zip Code -> Assess Postage -> Attach letter"

[‡] For example, a "string" is defined as a "series of characters from a to z and A to Z as well as numbers, spaces, and punctuation. This level of excruciating detail is necessary for accurate implementation in a computer.

After the creation of the SDFN, it is used as an artifact into a short open-ended interview for self-reflection.

The terms and nature of the SDFN should be gently explained to the participant. If they seem unsure, engage in the brainstorming tactic of creating an entity dictionary before proceeding. From conversation during the introduction, the participant and interviewer should agree upon the topic, referred to as the UoD[*]. The topic should be drawn from the participant's common work experience, to give them sufficient memories to draw upon. The pre-defined topic sets the limits of exploration and acts as a UoD. When those limits have been reached, the SDFN is completed. During the interview, avoid using terms like "universe of discourse" or even "ready-to-hand" because the jargon only distracted from the topic.

Instead, in the small chat at the start of the interview, ask them to speak about their job. Then ask open questions about "interesting" aspects of their job, ones that involve data in some way. This casual conversation is crucial for reassuring the participant and steering the direction of the discussion. If the small talk was not enough, move to the creation of a heavily scaffolded entity dictionary, to see what entities they are most interested in, and thereby define the UoD.

Participants begin by describing or selecting two entities within the UoD. One entity should be a role associated with the participant for ease of imagination and the other can be anything with which the participant interacts. Quite a lot of prompting will generally be necessary during this first interaction. Prompting should take the form of open-ended questions, guiding the participants to first establish their own roles as entities. Once they have described themselves, they should identify the role they interact

---

[*] The Universe of Discourse is the bounded realm under investigation.

with as another entity, and then be guided into describing a flow between those entities. Through the use of guiding open-ended questions and the interviewer serving as scribe, each participant should create the entities, flows, label the flow, and then categorize the flow. The interviewer should never label or categorize anything.

This process of identifying flow and entity should continue in an iterative loop until the interviewee starts tiring. Generally, the topic will be sufficiently broad for a 30- to 40-minute diagramming session. If you end earlier, repeat with a different topic in a new diagram. After identifying the first pair of entities, however, the order changes. The subject should be encouraged to identify a flow first, and then to add entities as necessary. Participants should only create one new entity at a time, and then try to relate other entities to that one.

The graph, for purposes of clarity and ease of expression, should remain at least weakly connected[*]. Isolated subgraphs should be moved to their own papers and explored as completely different graphs. Separating the graphs can create distance between the topics. This distance allows one topic to be completed and then another role to be assumed when talking about the other topic.

The interviewer has a number of tasks during this process. He or she should provide enough scaffolding[†] so that the interviewees feel comfortable in suggesting their own

---

[*] Roughly speaking, a weakly connected graph means that every entity must somehow be attached to all the entities present in the SDFN. While "subgraphs" (groups of entities not connected to the rest of the graph at all) are possible, they tend to increase confusion and should be dealt with separately. "A directed graph is called weakly connected if replacing all of its directed edges with undirected edges produces a connected (undirected) graph. It is connected if it contains a directed path from u to v or a directed path from v to u for every pair of vertices u, v. It is strongly connected or strong if it contains a directed path from u to v and a directed path from v to u for every pair of vertices u, v. The strong components are the maximal strongly connected subgraphs." Wikipedia – http://en.wikipedia.org/wiki/Connectivity_(graph_theory)

[†] Scaffolding: structured guidance to the participant to reduce choice paralysis and help direct them to the correct actions in the circumstance. Different people need different amounts of scaffolding, and it can be progressively removed as the participant learns what they need to do. The metaphor is well discussed by stone in relating to children's learning, but can also be applied to interface design [53].

flows and entities. This will require progressively less scaffolding as the first few entities will provide both positive and negative feedback. It is vital to gently clarify incorrect entities and flow descriptions the moment they are suggested. The interviewer must insure each flow added is unique, understandable, and directed. Although correcting flows after the fact is encouraged as the participant refines his or her terms and understanding of the diagram, ambiguity must be caught immediately, before it can sabotage the SDFN.

> The process of clarification can be seen in this transcript:
>> Interviewer: What other flows are there?
>> Participant: Well, it just sends back results.
>> Interviewer: Same results or are these results different from these results?
>> Participant: They are different. But not in nature. Just in ... obviously, I'm not going to take every result I take from the code and send it on. Because that would be ridiculous.

When a participant uses an entity word in a different way, it is important to catch the usage and ask questions about it. Clarifications also serve as negative results, as "what do you mean by that?" changes their mental term for an entity as the term is refined. In contrast, simple and low-key responses like "Cool, so would you classify that as data, information, knowledge, or O?" are positive feedback, indicating a mild approval and acceptance of the concept. In the beginning, it is better to be more detailed about the nature of entities so that, by the end of the interview, the labels on flows and entities are just flowing naturally.

This iterative approach is also useful as it saves significant and boring theoretical explanations at the start of the interview, which may bore the participant, make them hostile (as some do not like being explained to), or be redundant because they're not listening anyways.

The objective of the SDFN is a page or two of bubbles connected by arcing flows[*]. This paper graph can be trivially digitized in Graphviz. Graphviz is a graph layout program that accepts a text description of the desired diagram and then renders it graphically. The application of Graphviz to the problem saves significant post-processing time in labeling and diagramming flows. Although this research was rendered in Graphviz on Linux workstations, any program that can render graphs can be used for post-processing.

Post-processing involves roughly three steps. First, in one file, describe the list of entities and the relationships among them[†]. Entities should be defined first with distinct labels. The distinct labels are very useful because they provide a way to ensure the quality of the subsequent graph. Graphviz is quite permissive with entities. Typos in entity names in either the entity or relationship section will be happily accepted as valid input by the program. Identifying unusual entities that are not expected on the final output is a great way of checking for node validity.

Edge validity can be checked by counting the total number of edges of each entity[‡]. There should be a 1:1 relationship between the paper level of connectedness and the diagram. By counting the number of edges around each node and comparing that total to the original graph, one can trace errors in the diagramming to specific entities and then fix those errors.

---

[*] See figure 6.1.
[†] See appendix B page (389) for code.
[‡] Starting from the top of an entity, make a tiny mark at the edge chosen, then circle clockwise around it, counting each edge. The count should be the same for the entity on paper and the entity rendered in the computer-based visualization. This practice is more effective than counting every edge in the diagram at one time because when the count is off, it is easier to figure out what edge is missing, but faster than comparing edge by edge.

6.1.3 Running an Interview

This section will discuss the necessary items and methodology for running the interview. There are only two physical requirements: good paper and two good recorders. A backup recorder is essential because these interviews are impossible to duplicate: as people resolve their internal cognitive dissonance throughout the interview, their answers change. It is therefore impossible to re-run the interview, though running follow-up interviews tends to be quite fruitful. It is important to prepare for all the ways in which an interview might fail. A repeated interview covering the same ground should instead focus on a discussion of categorization choices on the interviewee's already completed SDFN. The SDFN should have clarified their internal thinking as to what their personal construction of data was so all that remains is to re-record their ideas.

During my interviews, I used a mini-recorder and my laptop. The laptop, despite being large and distracting, served as an excellent recording device because it recorded directly in the audio post-processing program Audacity. Audacity is highly recommended both as an interview-recording program and as a sound post-processing program. It is important to process the recordings before transcription due to inevitable background, A/C, and RF noise. Phones should be turned off during the interview as they generate inordinate amounts of RF noise that can severely corrupt the recording.

Ease of access is a function of recording availability and limits the utility of many mini-recorders. Extracting recordings from some recorders involves considerable effort and requires proprietary software and cords. It is important to test the full extraction process from all of the candidate recording devices before engaging in an interview. If it is not easily feasible to extract common file formats from the device, select another device. Optimally, the device will produce an mp3 audio file, as that is the *de facto* compressed

audio standard. Voice, being easily compressible, is a perfect fit for mp3, and many hours of recording can be stored with ease. An earlier uncompressed format (wav) is also suitable, being compatible with any modern computer. The wav file sizes are, however, much larger. Before the interview, make sure there is sufficient space on the devices for twice the estimated interview length.

Paper selection is significantly easier. A large pad of paper is sufficient, though higher quality pads are desirable as they will tear less easily, and absorb the ink from the pens. Fast drying pens are preferred, though any tip will work. Each sheet on the pad should be labeled as it is used with the number of the interview, the page count, and the date. In case the pages are arranged out of order, this information is sufficient to reconstruct the drawing order and interview.

6.1.4 Timing Considerations

The interviewer should allocate around 15 minutes on both sides of the interview for equipment preparation. During the interview, another 10 to 15 minutes should be spent on breaking the ice and making the interviewee comfortable. Creating the SDFN will take half an hour to an hour, depending on how complex a diagram the participant desires. Although it is theoretically possible to compile the answers for a SDFN diagram very quickly, try to encourage the participant until either a page is filled or he or she is clearly horribly lost.

A subsequent discussion, once the SDFN has been completed, is completely optional. Some of the people interviewed want to reflect, whereas Os, uncomfortable with the procedure, do not. Due to this huge variation, there is no standard duration for this discussion, because it can go as long as the participant would like it to go. It is unusual, however, for it to go more than half an hour. If the participant is still interested after

half an hour, attempt to schedule a second, follow-up interview. The SDFN creation tends to be quite draining, and new insights may appear after a few days off for internal self-reflection.

Preparation is fairly trivial with enough pre-interview logistics work. It is important to have liquids and treats for both parties available, as there will be a significant amount of oral discussion. In the meeting area, try to position the discussion around a corner of a table. Having large separation between the interviewer and participant is contraindicated on both theoretical and practical levels. Theoretically, it is a bad idea to introduce any sense of distance or remoteness, as it will just increase the difficulties of icebreaking. Practically, the sheet will change hands many times, and a short distance will allow both parties to read edits and additions as they happen. Normally, the interviewer will serve as scribe to render the participant's descriptions in a common and consistent format. The participant should nevertheless see what is being scribed in real time, to offer feedback and corrections of his or her own thoughts.

The final element of preparation is to ensure the operation of both recorders. The recorders should be positioned out of the direct line between interviewer and participant. If possible, they should be positioned to pick up the participant clearly and isolated from the table to reduce the thumps and scratches transmitted by the table. Recording devices will pick up hand movements and emphatic gestures that hit the table depressingly well.

Have some sort of subtle timing device to ensure the interview is proceeding according to schedule. Make sure that it is possible to look at the timer without disrupting the concentration of the participant. A watch in this regard is a poor choice, as looking at a watch is a social cue for many people. Cell phones pose a similar problem (and in any case, they and other radio devices should be off during the interview to prevent transmission interference.) Try to record at least 30 seconds of silence before the interview

begins, and turn off the option on the recorders to not record white noise, because those measures will help with post-processing operations.

6.2   Conducting the Interview

The interview necessary to explore a participant's individual construction of data has three phases. Initially, the interviewer should collect demographics through the introduction. The introduction is primarily a means to diffuse anxiety and to gain the critical basis for comparison between parties. The second phase is that of constructing the SDFN. The SDFN exposes the practical understandings of data of the participant through repeated categorization. The final component of the interview, the denouement, tends to be a discussion of the participant's understanding of data uncovered by the SDFN. As the process can cause a construction to change as it is articulated, this self-reflection period is an excellent opportunity for the participant to air their thoughts and describe their new or revealed understanding.

6.2.1   Introduction

The introduction serves multiple purposes. Primarily, it diffuses anxiety, explains the background of the subject, and creates a scaffold for the intuitive prompting of the SDFN. In these interviews, people display many different sources of anxiety.

The most common is a sort of performance anxiety, wherein they do not believe their opinions are sufficiently privileged to describe their "understanding of data." Another common difficulty is job anxiety. Participants may feel that they are revealing secrets of their job to an outsider who, either as a spy for management or for some other reason, would steal the secrets to the participant's detriment. It is vital, in this

stage, to reassure the participants of the intent of the interview and to make them feel in control-as they in fact are.

The other goal of the introduction is to provide the interviewer with an understanding of the background of the participant. This background understanding will provide for demographics and will hint at the topic of the SDFN diagram. By investigating their work and educational experience, it is possible to gather data regarding any possible links between work, education, and their understanding of data. Understanding educational background is also important because it shapes the nature of the jargon used, and is an explicit way of changing vocabulary.

As the participant discusses their work experience, especially in relation to their understanding of data, incidents that are important to them will arise. By drawing out these incidents for any significance of data flows, one can choose a topic for the SDFN that is both engaging to the participant and a fruitful for examination during the SDFN. If repurposing this methodology for other tasks, at this point the task-specific goals should be emphasized, because by choosing a topic for discussion, the participant is implicitly assuming a role and engaging in a particular mindset.

After the participant engages in the discussion, it is important to explain the nature of the SDFN. Lightly explain flows and entities, the purpose of the diagram, and the nature of categorization. This explanation should be far less philosophical than even the descriptions presented above. A flow, to participants, is "any flow of data, information, or knowledge between one entity or another"; an entity is "a person, place, or thing that can interact with the flows." This is a significant point of divergence for participants. Some people will understand the nature of entities quite clearly, as shown by their body language, and Os will not. If it looks as though the participant does not understand, correct that problem by building an entity dictionary. The discussion of categorization

should explain that, "The content of the flow will go above the flow. Content is roughly what is flowing between the two entities. Then I'll ask you to categorize the nature of the flow, whether it's data, information, knowledge, or other."

If the participant is confused about entities, help them to create an entity dictionary. Ask them to describe typical entities from their workday and to describe themselves in various roles. Then ask them to describe other roles and things with which they work. The building of the entity dictionary provides the maximum scaffolding for teaching them about the nature of entities.

6.2.2   SDFN Building

After the participant was comfortable with the topic, and an entity dictionary was built (if appropriate), the SDFN began. I avoided asking for definitions of data to avoid contaminating their categories with half-remembered definitions from their educational days.

The SDFN is designed to encourage participants to intuitively define their understanding of data. Classification is a way of probing operational (rather than theoretical) understanding. Repeatedly confronting people with their "gut reactions" creates a cognitive dissonance[*] between the theory and practice that the participant will articulate during the process.

It was important to engage the participant as a subject-matter expert. The SDFN should explore a safe topic within the subject's expertise. A project, a process, or everyday interactions are excellent topical areas, as long as the participants have a strong familiarity with the domain. The choice of topic is important because it empowers the

---

[*] See Schultz for a theory of cognitive dissonance [51].

participants. Their experience in the domain reduces their uncertainty and fears of being wrong. Explaining what you do every day and are good at to someone who is interested and willing to listen also tends to be pleasant for most people, because of the validation[*] inherent in the discussion.

6.2.2.1 Methodology of the SDFN

When constructing the SDFN, I acted as primary scribe. While the participant should have access to a pen so that he or she can scribble corrections, the interviewer will do most of the drawing. Because the activity of the SDFN is to iteratively construct flows of "data, information, and knowledge" between entities described by the participant as the subject-matter expert, this section will discuss the structure I provided to participants.

Describing a flow always began with entity declaration. The participant declared which two entities the flow is between and then declares the flow itself. My prompting for categorization changed throughout the interview. Initially, the questions were quite explicit. "Who starts the flow? What do they do?" In this high-scaffolding variant, I explicitly identified the source of the flow and then guided participants to identify the destination and then the nature of the flow. By reducing the focus of the question to the smallest possible parts, I helped the participant not to feel confused by trying to think about too many unfamiliar things at once. Early in the interview, breaking the questions into tiny sub-questions allows for prompt feedback. As the participant learned through positive and negative reinforcement, their awareness of expectations reduced the need for tiny sub-questions.

I mentally examined an entity before committing it to the diagram. When validat-

_____

[*] Validation is a confirmatory statement that increases a person's self-worth [54].
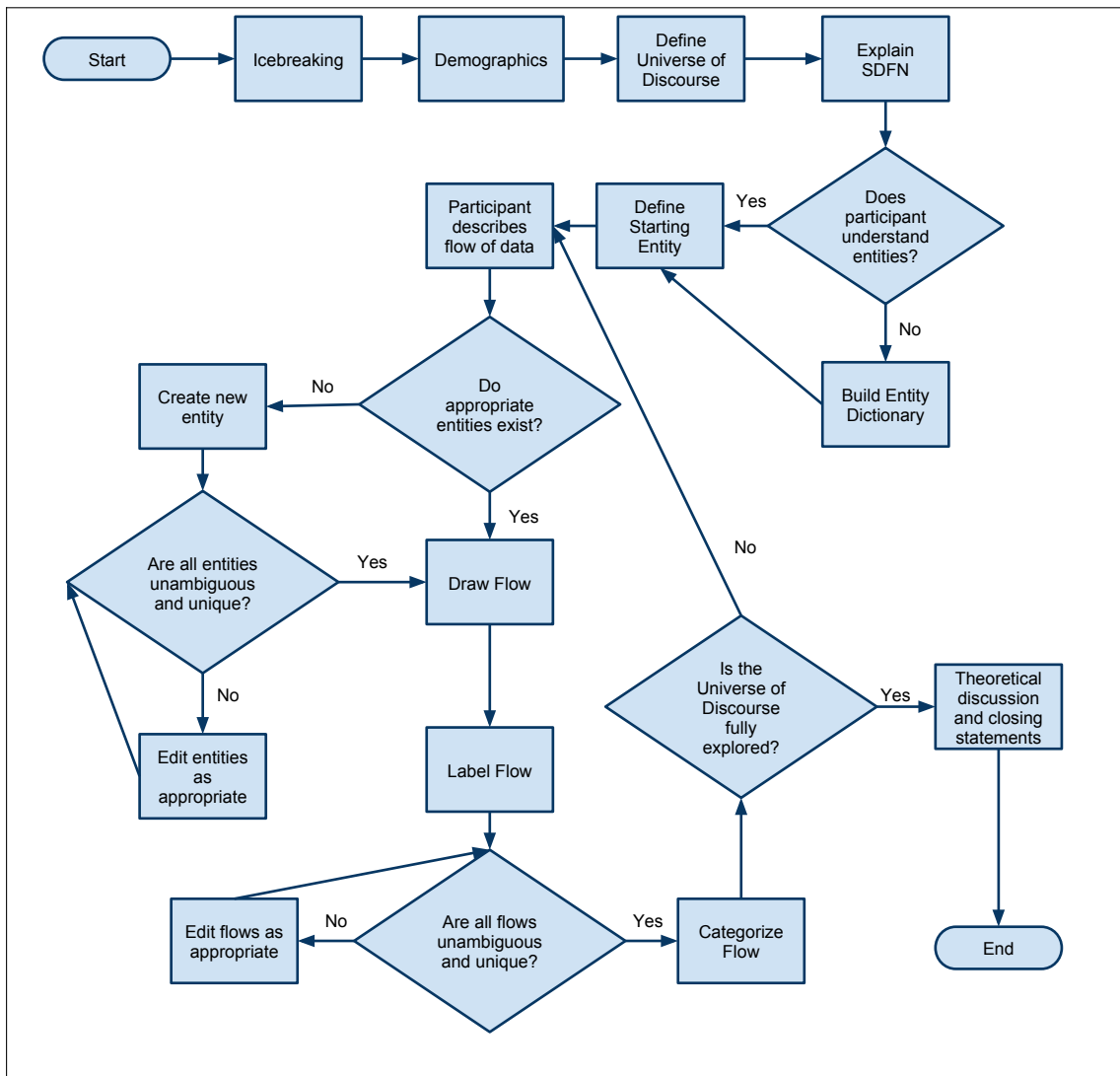
Figure 6.2: This is a flowchart exploring the complete SDFN interview process. Different portions of this chapter will refer to different elements.

ing the source entity, my decision tree was simple: Does the entity exist? If it does not exist, is the role that the participant described short and unambiguous? Entities must have short names, as names take up valuable space. If the entity bubble is more than about an inch in diameter, it takes up too much space and is likely to require redrawing the diagram.

I encouraged the participant to generalize to the point that the entity can rea-

sonably cover all objects in its class. "Person" is too vague; it does not give the reader any sense of what role the person occupies. "Brian" is both too vague and too specific: it identifies a specific person, but does not suggest his role (thereby requiring more explanation) and does not allow for similar types of person. Good entities would be "Dissertation Author" or "Casual Reader" or "Examiner." They can be generalized to one and only one role; they are simple (few words), and they allow for anyone who fits that role to be classified without unnecessary recourse to edge cases. By using examples from my personal experience, I was able to consistently give the same scenario example across multiple interviews, without looking as if I was reading from a script.

With the source constructed, the participant needs to define the destination using the same methodology. Later, the scaffolding can mostly be withdrawn and the entire process summarized with an "And then?" as the participant understands what is being asked. The transition should be gradual, rather than abrupt, and is predicated on the error rate of the participant. If the error rate increases, I increased the scaffolding to compensate.

Once the entities are identified and written, the participant should describe a flow. Using an arcing path with a clear "above and below," connect the two entities. These arcing lines, representing flows, are oriented to the entities, not the page. I rotated the page if it allowed for cleaner arcs and more space above and below for description.

With the role description in mind, the participant was asked to note what the flow contains with a theme and variations on: "This is a flow of…" I made sure, when asking these questions, that they were as open-ended as possible. Flows also must be unambiguous. Ambiguity may be introduced via the introduction of 0, similarly named flows. When a flow is described, check all the extant flows for identical and similar names. In the case of an identical name, inquire whether the same contents are flowing. If the

names are similar, make sure there are enough adjectives around the flow to distinguish the two. I edited the prior flow if it makes more sense to do so.

When the flow has been described, and the noun (with or without adjectives) has been written above the arc, then prompt the participant to categorize the arc as data, information, or knowledge: "And is this flow of x data, information, knowledge, or O?" Unusually, this element of scaffolding is never completely removed, although it may be shortened as appropriate if the participant has already categorized things of its nature on the diagram. While other is a category, and the participants should be allowed the luxury of defining new categories, the greatest utility is derived if most participants classify most flows as data, information, or knowledge. As such, insure that only few new categories are made.

6.2.3 Theoretical Discussion

Having spent at least a productive half an hour on the SDFN building, participants were then encouraged to conclude the session with a theoretical discussion on their own revealed constructions of data. Not all participants will want to have this discussion, and beyond a vague prompt of "And would you like to talk about your thoughts on the philosophy of data?", it is not worth the effort to force reevaluations here. This discussion generally explores the ontological and epistemological questions and novel categorizations that arose during the SDFN creation.

Participants, if not uncomfortable from the unusual thinking demands of the SDFN, generally engaged in a self-reflective discussion. It was vital to ask open questions that build a scaffold for the participant's self-discovery.

Of interest in the discussion are how the participant transitioned from one of the categories to another. The SDFN allows for a very solid investigation of interesting

questions of categorization. In this case, I was exploring how the participant categorized things as data, information, or knowledge, their boundaries, and their transitions.

The main opportunity of the discussion is to study how the participant thought about the interaction between categories. In many interviews, the participants would construct a hierarchy in the theoretical discussion and discuss how data became information. The self-generation of a theoretical ontology of data from the practical categorization experience is the main purpose of this methodology.

other items to query include any departures from the normal flows. The participant, when creating the SDFN, will prefer certain categories to Os, depending on topic and personal preference. When given the opportunity, I tried to ask about the unusual categories: instances in which the participant either made up a category, combined or concatenated categories, or used a rare classification. For example, when certain participants strongly favored data, only a few flows were classified as knowledge, mainly because they fell outside the scope of the discussion. Subsequent conversation then focused on those flows to try to get a balanced understanding of their nature.

When discussing combined or concatenated flows, it is important to understand the type of metaphor that the participant is using [55]. While this is mostly important in terms of analysis, my participants normally used low key but figurative metaphors. During this process, I tried to ensure that they elaborated on each metaphor and clearly distinguished the type they were using. A container metaphor ("data holds information") is different from that of a concatenative metaphor ("data alongside information") and both are different from a combinatory metaphor ("data and information").

A brief discussion of metaphor is in order. Participants gave me clues to their personal construction of data through their use of metaphor, especially their use of verbal affordances. Anytime the participants said, "filter data," it was a strong clue that

they were very interested in subjective masses of data that needed to be winnowed down. Discussions of precision and accuracy, or any kind of implicit meta-data were pointers to objective elements.

> An example of container metaphor in action:
>> Interviewer: ... To you, what is Information?
>>
>> Participant: It's something that's not physical, basically. That means it could be a communication, it could be a conversation or a story. Something verbal.
>>
>> Interviewer: So Information is any non-...
>>
>> Participant: Data to me is physical, basically. It's an entity that [fuzzy word] an entity of some sort.
>>
>> Interviewer: You would say that letter is Data?
>>
>> Participant: I would say that letter is Data. But what is on that Data is Information. Just to be a bit confusing.
>>
>> Interviewer: Tell me more.
>>
>> Participant: Coming from the field we're in, in [removed]. You've got different areas. It would be, as I've said, Data or record. That's just the physical entity.

I notice the "what's on that Data is Information", because "on" is a "container" word.

>> Interviewer: Data is a container...
>>
>> Participant: Yes, of the Information.
>>
>> Interviewer: And Information is content of what type? Is there something common to all Information?
>>
>> Participant: it's not easy. I'm struggling. It's a struggle. Information, what is Information? It's just... No, I'm drawing blank.
>>
>> Interviewer: We have right now Data
>>
>> Participant: And then of course you get Knowledge.
>>
>> Interviewer: We have Data, Data is a container for Information.
>>
>> Participant: I'm quite happy with that.
>>
>> Interviewer: You say physical at some point?
>>
>> Participant: Yes, physical in [unintelligible]. It doesn't actually have to be physical in a piece of paper, but it can be physical as in an e-mail message.

Whereas, in a different interview, we have the combinative "and":

Interviewer: Analysis, explanation. Analysis is?

Participant: Something like: model outputs or calculations. Explanation is what that actually means in context of your [work environment].

Interviewer: Class of analysis is Data, Information, Knowledge?

Participant: It's probably more on the Information. Well, it's Data and Information. And the explanation is Knowledge, it had better be.

Interviewer: When you say that, what do you mean?

Participant: You hope that when someone gives you their explanation, you know more than before they told you. Not always true. They can tell you stuff, and you can go "Well, I understand even less than when I started." Because if it's completely contradictory to your understanding, you are now really confused. And in the cross cultural... you're often doing these meetings with interpreters in the room. Am I not asking it right? You don't necessarily get an interpreter that speaks Technical [other language]. Often they're here for some other meeting, and they just bring someone from one of the marketers who will be bilingual. Now, they all are bilingual, but particularly senior people will choose not to speak English because it's embarrassing when they speak badly. That said, we can't speak [other language], so who are we to criticize?

The keyword "and" of "Analysis is data and information" meant that they are combinable and can sit in one flow. The concatenative idea is a lot more difficult. Although in the following example I allude to a container, the interaction between data and information is not the strict container of the first example, but rather has more "alongside" affordances[*]:

Participant: As part of the Information flows to those, I may include a little bit of raw Data, but not very much.

Interviewer: Does this raw Data ...

Participant: Usually photos or a graph.

Interviewer: So you would say that photos are raw data.

Participant: Yes.

Interviewer: Would you say that the raw data is contained in the Information? i.e. you send them interpretation of measurements. As part of that interpretation you have to send them some of the measurements that are really interesting. Would you say that the Data is inside the

---

[*] Metaphor provides the affordances of the thing being related in the metaphor. Therefore, a container metaphor affords "putting into." It is by analysis of which set of affordances is hinted at most strongly by the participant that important clues are gleaned towards the participant's conception of data.

Information flow, and we can just label this as Information? Or would you say that it's Information + Data?

Participant: I'd keep it inside the Information flow, because if it was just raw Data. They could very easily reach what I think is the wrong idea – misinterpret.

Interviewer: Therefore, you're not going: "Here's the Information, here's the Data." You're going "Here's the Information, here's some Data inside the Information to back it up."

Participant: Yeah, that's right. But with just the raw numbers and no context, that's Data.

This discussion, if fruitful, will lead into definitions. Having looked at the ontological transitions in the discussion above, the participant may now be prepared to examine the ontological definitions of the various categories. Here was one of the more treacherous spots of the presentation, because it would have been extremely easy to put words into the participant's mouth through suggestions or overly specific leading questions.

Instead, I tried to allow the participants to use their own inductive process on the categories and transitions they have defined. Normally, the basis of the definitions will occur in the discussion of transitions, but it may not happen in every case. If possible, guide the participants to identify and discuss their own thoughts of how they categorized something as data.

Although the relationship questions are normally deeper, leading into this discussion through transitions means that the various affordances and other philosophical handles of data, information, and knowledge will be discussed first. Data, lacking form, has no "natural" or non-constructed affordances. The reification of data through the SDFN caused the participant to suggest their own affordances, and thereby strongly hint at their conception of data. The other component is to ask them to discuss their understanding of how they know something is data or of how they categorize it. The only structure possible here is that provided by the SDFN itself. The participants were also prompted to explain their categorization methodologies.

6.3   Survey

After the completion of the interview process, a decision to create a survey was made. The purpose of the survey would be to test two things: would it be possible to replicate the success of the interview technique in a more automated form and would the different constructions of data be evident in a more varied audience? Unfortunately, the survey suffers from both methodological and coding flaws, and is therefore presented for intellectual curiosity only. It is clearly a necessary direction for future research. One of the most critical problems is the framing of the survey, explitly asking for a differentiation between data, information, and knowledge: "This survey is exploring what you think about Data. To do that, the survey will present a list of short "scenarios". We will ask you to categorize the scenario as involving Data, Information, Knowledge, or something else, depending on your own understanding of the terms." Unfortunately, this posits an artificial distinction between data, information, and knowledge that the participants may not originally have perceived.

During the process of collecting data, an unexpected opportunity arose: a mailing list of retired intelligence officers and agents was interested in my research. To take advantage of this opportunity, I created a survey. Optimistically, the first survey was a direct copy of the interview process, starting with a complex demographic interrogation, asking the participants to create flows and entities, and then asking them to self-direct their own investigation into their understanding of data.

It was a complete failure.

In the first survey pilot test, 18 people attempted to complete the survey. Only the person who had participated in my interviews had any idea of what the survey was talking about, and even that attempt produced no useful data. Most of the participants

failed to complete the survey after taking the demographics section as an exemplar of the whole, and stalled horribly at the "now describe an entity" section.

I believe the people who attempted the survey ran into two significant problems. The first, and more critical, was the symptom known as "tl;dr" or "too long; did not read." An associate, practiced in survey creation, suggested that no one taking a survey would read more than three sentences of instructions. As these surveys presented multiple paragraphs detailing and defining terms, it was clear that the obstacle of tl;dr was in full effect.

More subtly, though, the very abstract and theoretical nature of the questions was a problem in creating scaffolding. In the interview, because I was able to provide assistance and incremental steps according to *my assessment* of the participant's comprehension, I do not believe that any participant found the process exceedingly difficult. Rather, because the survey was self-guided, its impersonality was its primary point of failure. In the interest of making a survey that people could finish quickly, I had created one that was not able to adapt the scaffolding processes that made the interviews successful. Therefore, the only people to complete it were those who *already* knew about the concepts being discussed: one of my interview participants, and an academic who specialized in teaching the DFD methodology.

From this failure, I learned that a new methodology would be required. The primary lesson was that the direct translation of interview techniques failed. My intuition was that the success of the interview was based on the feedback given by the interviewer, not the structure of the interview *per se*. In an online survey, people expect mostly to click answers, rather than to type essays in a web form. Very few long-answer questions are appropriate to such a format, however, and a survey comprised entirely of them is wholly inappropriate for anything but a final exam. The informal nature of a survey makes

the kind of focused concentration required of long answers quite difficult, especially considering the lack of any reward besides the completion itself. It was also a mistake to establish expectations in the demographic area of the survey and then violate them on the next page through a longer theoretical component.

When considering what to include in my second attempt, I could not simply consider all the myriad ways that the first survey failed. It was also important to understand the few ways in which the first survey succeeded. The two principal successes of the first survey were in the demographic section and in the tool itself. The demographic section successfully captured interesting demographic information at a high granularity. The tool, Limesurvey, performed far beyond expectations. It is well written, database-agnostic, secure, free, and open source. The mechanism for importing and exporting surveys is streamlined and very functional.

A slavish copy of the methodology of the interview was clearly unsuccessful, and so any theme and variation on that would almost certainly share the same fate. I had to reconsider what question it was that I was trying to answer. In the first survey, the question developed into, "With a self-constructed SDFN, can you articulate your own philosophies of data, information, and knowledge?" The respondents presented a very straightforward answer: "No." The essence of the SDFN is in the process of categorization. Although the interview length lent itself to a thorough exploration of the self-declared roles and their own data transfers, the essence of the SDFN was in enticing categorization of many different, and distinct, flows.

I realized that it was possible to remove person-specific flows and allow people to classify a general set of scenarios. I wanted to explore a specific question: "How does a specific role categorize data, information, and knowledge." The question of role was tricky, despite the success of the demographics section; the participants' answers

did not suggest which role headspace they were considering their answers from. A hypothesis, while creating this, was that people would have different answers to the categories depending on the role in which they were thinking at the time-an explanation substantiated by the remarkably different interview answers one participant gave when interviewed twice about remarkably different topics. I needed to assess the person's role, rather than just his or her generalized demographics while keeping the results completely anonymous. As the scope of the prior project was in many ways its fatal flaw, minimalism was the rule of the day in the second attempt.

The survey opened with "This survey has requested that you answer it from the perspective of one of the jobs that you do. Please describe the duties of that job (in general).". Earlier, I asked participants to: "We believe that people can have different philosophies, depending on what job they're doing. For this survey we ask that you think about the scenarios from the perspective of one of your jobs." The phrasing of the first sentence was unfortunate, invalidating the survey's "scientific reliability." It is my belief that the question solicits all necessary demographic information without extending beyond the participant's comfort zone of anonymity.

The survey questions after this point all had the same format. They would begin with: "I am trying to understand what you think of as Data, and why. The questions below ask you to categorize the scenario, and then explain the categorization. Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario."

Then the scenario is presented. Here are all the scenarios in order. The scenarios were chosen such that the highlighted phrase would help to differentiate between the three constructions of data found during the interviews. This strategy was not particularly effective.

- Alice receives a letter from Bob.

- Alice receives a letter from Bob containing instructions on how to build a machine.

- Alice receives a letter from Bob containing a short story he has written.

- Alice determines the locations for parts of a Rube Goldberg style machine to cook her breakfast.

- Alice receives a letter from Bob. The letter is a time chart of what shows he has watched on TV for the last week.

- Bob receives an e-mail from Alice, it is a record of the daily temperatures outside her apartment for the last week.

- Bob receives a flash drive from Alice. It contains mp3 music files.

- Bob attends a symphony with Alice and enjoys the live music.

- Bob ignores the traffic noise outside the symphony.

- As Bob is mugged walking home, the mugger demands his wallet and watch.

- Charlotte finds a microfilm in a hollow coin, it contains a list of numbers and times about something unknown.

- Charlotte finds a microfilm in a hollow coin, but cannot decipher the code.

- Charlotte finds the secret key to the code, and realizes it's a letter for technical support to the spy's handlers.

- Charlotte finds a microSD card in a hollow coin, it contains a planning program for something unknown.

- Charlotte creates a statistical profile of a spy, to predict their actions.

- Dave lectures to a classroom about database design.

- Dave grades quizzes from a relational algebra course.

- Dave discusses the reasons behind one of Eve's incorrect answers.

- Dave writes a survey asking people to describe their impressions of a user interface.

- Dave saves an empty word document in preparation for his later work on a conference paper.

- Eve writes poetry describing the winter wind.

- Eve interviews students for the campus TV station and gets short quotes for her topic.

- Eve looks at the weather report and decides to bring an umbrella.

- Eve receives a letter from an ex-boyfriend, telling her to take her stuff back.

- Frank selects which instrument readings to include in his experiment.

- Frank designs an experiment

Each of these surveys asked the participant to categorize the highlighted phrase as data, information, knowledge, or other. The other then provided a text box for elaboration. Participants were then given a large text area to explain their choice if they wished. This survey structure allowed for similar kinds of self reflection as found in the SDFN, but did not adequately phrase the questions.

6.3.1 Tools and Techniques of the Survey

With minimalism in mind, the first survey was completely reinterpreted. The final section of the first survey, which asked people to write an essay on their own conception of data, was removed, the demographics section was reduced to one question, and the flow diagramming was reduced to simple categorization.

While people were quite willing to answer the demographic questions, as stated earlier, I feel that the initial questions distracted from and reduced peoples' attention span for the subsequent survey. The only demographic question that really matters is about the participant's mindset.

The new survey asked participants to "vividly imagine" a role, thereby artificially putting them into that role's mindset. Asking participants to engage in a specific mindset is effectively asking them to play a role: to pretend to think in ways that are foreign to their current state of mind. They retain their authentic deep expertise in the domain that

they have chosen. By engaging in role-playing, participants assume the understanding of data of that role as the scenarios are filtered through the mental maps imposed by the role.

By asking participants to vividly imagine and then *describe* that role, the survey made it possible for them to reveal as little or as much as they wanted about their thinking methodology without disclosing potentially identifiable information. The survey gave the following instructions for the role:

> This survey is exploring what you think about Data. To do that, the survey will present a list of short scenarios. We will ask you to categorize the scenario as involving Data, Information, Knowledge, or something else, depending on your own understanding of the terms.
>
> We believe that people can have different philosophies, depending on what job they're doing. For this survey we ask that you think about the scenarios from the perspective of one of your jobs.

While this was a long answer question, the fact that it came first and was asking them simply to describe what they imagined seemed to allow for it to be effective.

After vividly imagining and describing a role, the survey moved into the pure measuring phase. This phase involved presentation of a one-sentence scenario with a term highlighted in bold[*]. The first scenario was, "Alice receives a letter from Bob." The survey requested: "Please read the following one sentence scenario. Categorize the highlighted word or phrase in the context of the scenario." Participants were asked to classify via a drop-down box whether a given scenario was "Data, Information, Knowledge, or O" and, if they chose O, a text-box appeared so that they could enter their own classification. Happily, this option was often utilized, suggesting that most participants did not choose a false category of data, information, or knowledge.

When participants selected a choice from the survey, they were then encouraged

---

[*] The full survey, in printable version, can be found in Appendix A.

Figure 6.3: Scenario 1 from the survey. The dropdown box asks participants to choose Data, Information, Knowledge, or other. If they choose 0, a text box appears for more details.

to explain themselves: "Please explain in one or two sentences why you categorized the scenario that way." This phrasing offered participants the choice to engage in as much self-reflection as they wanted about the only critical thing that mattered: the act of categorization. While it would have been desirable to have a more comprehensive survey, I feel that the second survey incorporated the lessons learned from the first and was able to produce some surprising results, regardless of its small scope.

The audience for this survey was chosen very informally from three roughly distinct groups. The first group "chosen" was from my social networks via Twitter and Facebook. This group was initially contacted to pilot test the survey. The power of social networking tools in this kind of research cannot be overstated.

As the first few results came in, the survey looked sufficiently effective at capturing the participant's understanding of data to launch it without modification. The second group consisted of the respondents to the flyer[*] sent to the INTELST Forum, a mailing list coordinated by the U.S. Pentagon and Army. The people who responded to the flyer were then e-mailed a link to the survey. Very little can be said about this group, save that they are all active or retired intelligence professionals from both the military and civilian side of things. The responses from this group were fascinating, and clearly reflected a personal construction of data different from that of the Os.

The third group was recruited through a presentation at the company, summarizing the findings of my interview research and inviting participants to take the survey. Unlike the more focused set of my "initial trial" group, a large subset of researchers and staff from the company was invited to my talk. This sampling allowed me to invite many distinct people to take the survey, and I feel that quite a few different jobs from the

---

[*] See Appendix C.

company were represented in the final results.

7   Interviews

This chapter presents redacted but raw transcripts from the interviews I conducted at the company. The SDFN is a novel methodology. As an untested methodology, and as the experiments and "analysis" were both conducted by me, the transcripts are a useful check for the reader. I appreciate the great length of these transcripts and offer excerpts in my reflections in the next chapter for readers not interested in excruciating detail.

The excruciating detail is necessary, however, for readers interested in potentially implementing my methodology. Not only do the raw transcripts provide a second perspective on my methodology, but they demonstrate my elicitation techniques through the act of recording rather than my synthesized instructions as found in the previous chapter.

Fundamentally, the success or failure of this elicitation method must be decided by the reader, not simply through my analysis of the results I obtained. This chapter, therefore, occurs before my reflections in an attempt to provide the reader an unbiased look at the material which informs my subsequent analysis.

The redaction process sought to eliminate all identifiable nouns, some verbs, and most positions from the text. All of these items have been replaced by curly braces, a descriptor like noun, or verb, and a number. The number is consistent for every instance of that term, allowing the reader to note patterns of terms without necessarily needing the initial terms themselves. Each diagram has undergone the same redaction process as the interviews, though their quality may suffer significantly from the reduced page size.

Each interview is ordered by the rough groupings of the reflections chapter, with interviews taking the same order in each chapter. This ordering is intended to support

readers flipping between chapters to check assertions or their hunches against my reflections. This order, however, is not the chronological order of the interviews, as that is being intentionally obscured to help protect anonymity.

7.1 Interview 1

This interview is presented in detail with annotations as a reflective case study. The technical pilot interview was the first interview conducted, and it served two purposes: to vet the equipment, and to provide a test of the methodology. After exposure to my methodology, my advocate could then use their experiences to persuade their co-workers to participate in my interview. Due to the more public nature of this interview, as well as the work with uncertain equipment, my advocate and I chose to work with a harmless topic: our mutual participation in an online game.

7.1.1 Collected Drawings

Please see the next page. The drawings are presented as full-page size for optimum readability.

7.1.2 Annotated Transcript

As this was the technical process, figure 7.1 illustrates the creation of an entity dictionary. As we can see, the entities are just sketchy bubbles with entity names in them, which may or may not appear in later cases.

The following is my explanation of the process to the participant:

Interviewer: Well, we'll be getting back to basically this question after we build the data flow diagrams. This is something to let simmer in your subconscious.
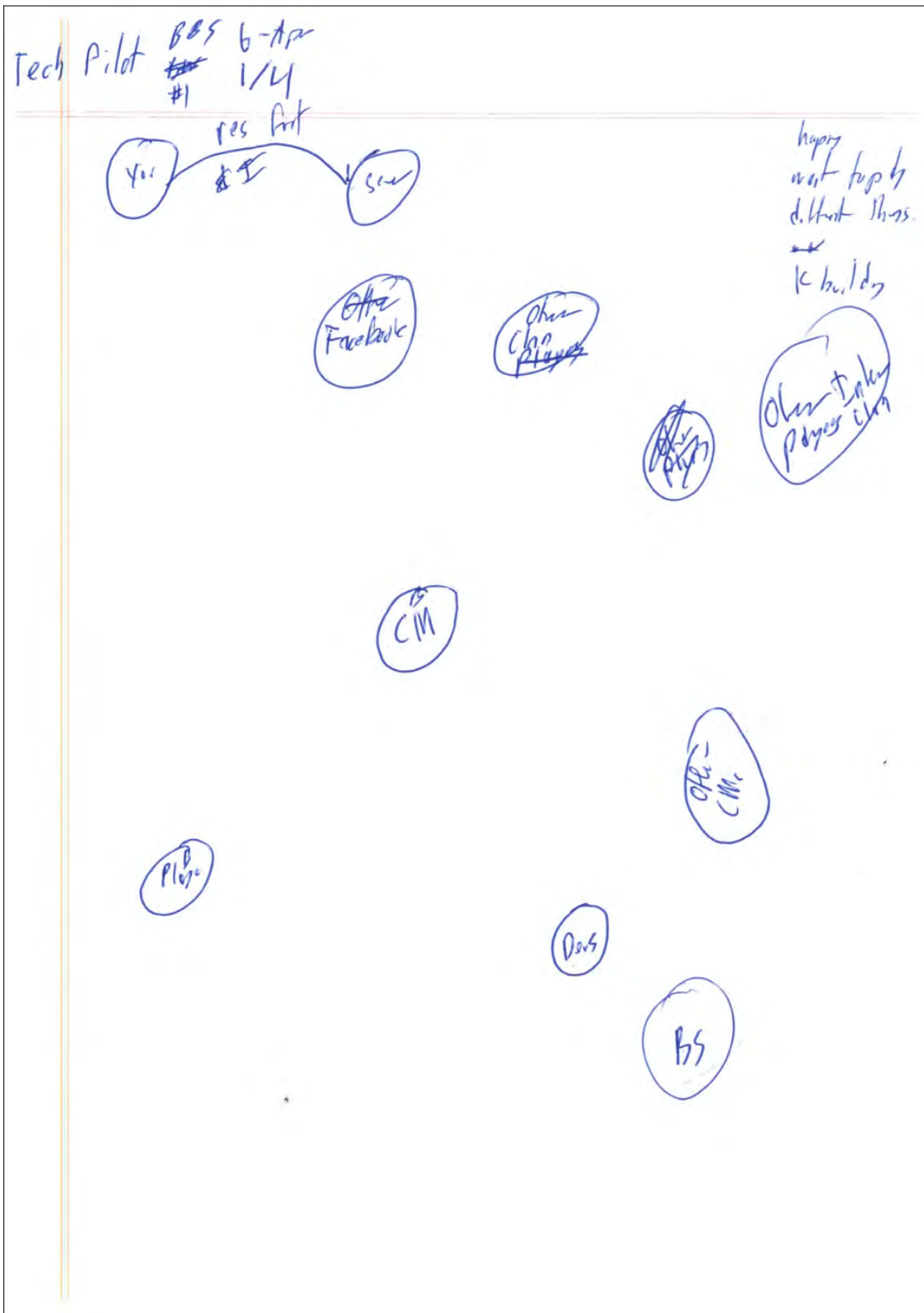
Figure 7.1: A sample entity dictionary. The participant was brainstorming possible entities for us to explore.
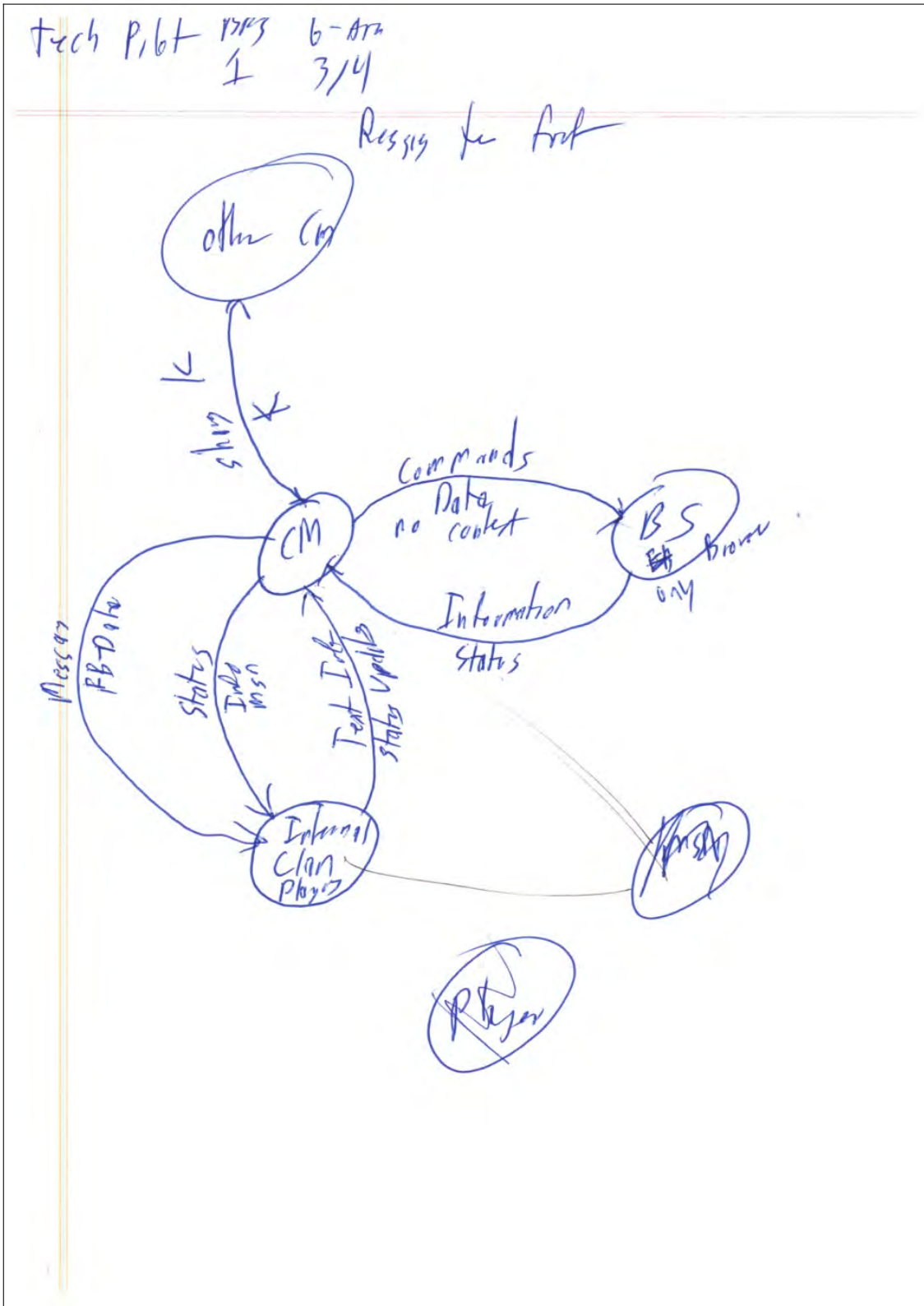
Figure 7.2:   The first "SDFN diagram." Note how each flow is categorized be-
low the flow and labeled above the flow, showing the necessity for curved
flows.
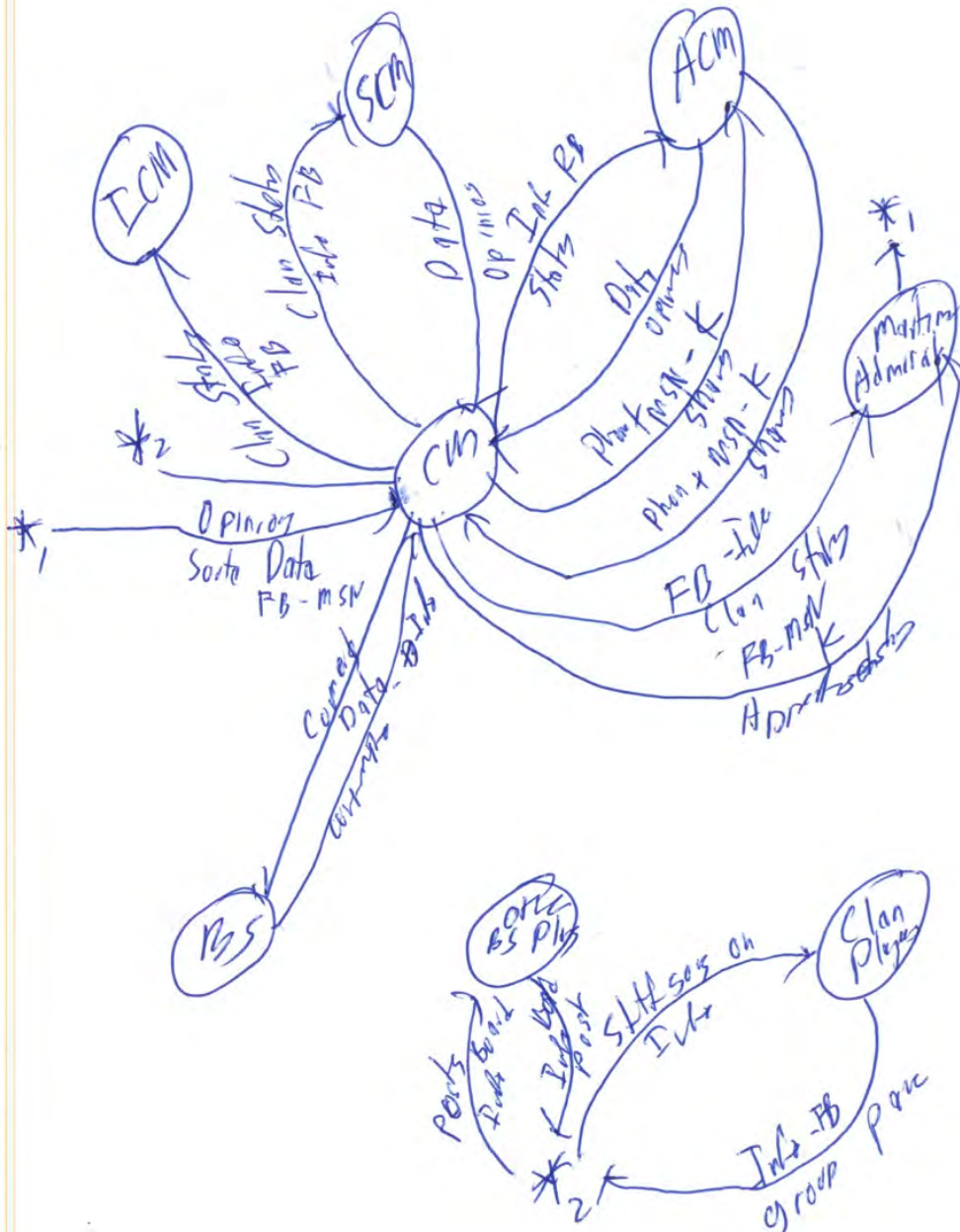
Figure 7.3:   Another SDFN diagram. Note the presence of wormholes.

Participant: Yeah it would... it's something that we should discuss more. Yeah, there's context of knowledge, and there's context of data. Maybe in a superficial analysis they don't meld. But if you think about it more deeply, you go, 'Oh, hang on.' It's not just making arbitrary distinction that this is knowledge and this is data. Think about, "Why do I consider that knowledge? Why do I consider that data?" That really is knowledge because its stuff I know not just there You're doing the Ph.D.

Interviewer: Well, yes, but I'm doing the Ph.D. based on what you tell me. So, when we're building this data flow diagram. What I'm going to be doing is two symbols. Well, we're going to be making circle symbol and a circle symbol will have some sort of designator that is important for you in it. It represents a person, or an entity, that communicates, transmits, verbs, data. Whatever you consider data to be.

Participant: Does that include knowledge?

Interviewer: Does it include knowledge?

Participant: Point Taken.

Interviewer: And then are going to be arcing lines sometimes with an arrow on it to another entity. The arcing lines we'll label with whatever you would classify the data/knowledge/information, whatever term you want to do. It's, for example, you said that there was a 'fail res'. As an example I'm not sure you would use this, one transmission of data would be you to the server, res the fort. We would label this line, 'you, server, rez fort.' And if you want to, you could say this is data, or that this isn't data, this is information. Whatever you think is important to define about that transmission. We'll label that line as.

Participant: Oh, so are they different classifications?

Interviewer: There are whatever classifications you want. What I'm going to be doing here, is looking at how you classify these things, and what small groups you identify and how you think that other people classify these things. And then compare it to how other people classify the things, and look at how the perceptions shift from person to person. There's no detail too trivial for this discussion. Because I'll be using it to look at what other people will be doing.

In this, the theory of the SDFN is not deeply explored. The important thing to do before the start of the process is to gently explain the process and diagramming techniques that will be employed in the rest of the interview. As the participants tend to be unsure at this stage, it is also important to avoid giving them definitions of data, information, or knowledge as they may be looking for specific cues to suggest which construction of data to use.

In this SDFN, I did not ask the participant if they wanted to do an entity dictionary, I just started with the process of articulating the entities. It is acceptable if the participant is chatty during this process. Not only will this help to define the universe of discourse, and establish them in their own minds as subject-matter experts, but also the process of being chatty is allowing them to slip more deeply into the role they are discussing.

> Interviewer: This is literally trying to render how you perceive the game. ... We'll start with a circle. How would you like to label this circle? What do you think is a representation of you? We can write you here, we can write the computer here.

> Participant: We can talk about me as a Clan Master. I'd put that one there. That's to start with. I'd like to have another one, for my different roles. Cause I'm also a player of the game. Which is different from their role as a Clan Master for sure. Because they're often in conflict.

> Interviewer: What other entities can we identify in general?

> Participant: Across the whole game?

> Interviewer: Well, that will transmit data, whatever that is, or information or knowledge. When I say data, feel free to assume I'm also talking about information or knowledge if that helps.

> Participant: That's good. Cause often, if you get too transfixed on your own view of data, jeez, there's lots. Like the developers. You've got individuals within there. Now, I guess this is when the granularity comes out, because you can talk about other clan masters as a circle. But they're individuals amongst themselves. So I can identify. Within that, you have the concept of a clan master. There's [Active Clan Master 1] but then they have other concepts such as –

> Interviewer: [Active Clan Master 1] is a person, yes?

> Participant: Person, yeah. You've got active CMs, which [Active Clan Master 1] can be part of this group. And you've got inactives.

> Interviewer: Inactive CMs?

> Participant: Yeah. So there's individuals, but within there, you could break that. And then say: "Well, look [Active Clan Master 1]'s here, or something like that. And [Clan Master 2]'s here. We'll often talk about semi-active as well. There's me as a thing and there's also Os within, and break it down like that.

> Interviewer: So what we'll want to do here is create some sort of representation so that you can talk about classes of people. Or individually people if you feel that they're important to be talked about as an individual. So, from this, you can say, "Well, this is a communication from me, to the active Clan Masters." "And I do this sort of communication." "Or, this is a communication from me to [Active Clan Master 1]." Whatever

represents what you're doing.

Participant: Look, cause sometimes – Well, the information we can talk about, what, information, data? is transmitted. That's why I sort of did them separately. Cause I want to transmit me and [Active Clan Master 1] will talk about something differently than we'll talk about with Os here. That can be a slightly different form of communication with these and with that. You deal with the individual differently even in a group.

Interviewer: And that's what I'm trying to tease out here. Thank you. We've got you as a player, you as a clan master, the Devs. Right now, we're just going to make bubbles and we'll take these bubbles to a third page when we're drawing lines. This is brainstorming.

Participant: We'll keep it sort of then at a higher level.

Interviewer: Whatever you want to do, if you want me to draw it, I'll draw it. If you want to draw it, go for it.

Participant: Nah, you might as well do it. So there's other CMs then, as part of the clan. And, I'll use that term to classify all the people that can rez the fort, per se.

Interviewer: Okay, so you define Clan Master as someone who's able to rez the fort?

Participant: Well, not really. But I'll group them together. There's what we call Martini admirals in [Clan] which aren't Clan Masters per se, but they have almost the same power as a clan Master. They can't be kicked. But I'll group them together. Because they're really, as you know in [Interviewer's Clan]. There's either those at the top and there's the rest of the players. So I'll call that other clan masters. And then there's other players. That's probably an easier, higher level distinction up there. And the next natural bubble is "other clans."

The other aspect of the SDFN that I was teaching the participant about here was the appropriate scope of an entity as well as the desired granularity of their universe of discourse, as it is bounded by both scope and detail: only so many actions are of interest, and some actions are too trivial to diagram.

The mistakes of figure ?? illustrate the participant demonstrating a misunderstanding of the SDFN, drawing their own hierarchy of authority within their organization. The creation of these side artifacts as part of the entity diagram is acceptable, especially as a way of pinpointing desired levels of granularity in entities. Participants should not think of the SDFN as a hierarchy.

The start of the SDFN can be quite subtle. In figure 7.2, the rezzing the fort SDFN

began as a "walk me through the process:"

> Participant: So the next would be – Maybe if I described the process ...
>
> Interviewer: Walk me through the process.
>
> Participant: Well, this is the case. I say "It's time clan war." This is where I bring in the extra bubbles.
>
> Interviewer: Let's trace this and see where we get off these bubbles from the process.
>
> Participant: We're at clan war. "I want to rez the fort." A command to rez the fort, it sends me back. Now, I guess to introduce the other bubble here as other clan members.
>
> Interviewer: So, we're going to say internal clan members? Can we say other than members because that conflicts with clan master?
>
> Participant: Clan Players?
>
> Interviewer: Players. ICP.
>
> Participant: I'll leave you with the acronyms. I tell them something now as well.
>
> Interviewer: Now, when you tell them something.
>
> Participant: There's a lot. That's a really detailed line between us and
>
> Interviewer: Do we want to multiple lines?
>
> Participant: Yeah. That's cool What we're dealing with at the moment. I'm going through the process of rezzing the fort, which is a common thing we want to do.
>
> Interviewer: So shall we label this rezzing the fort?
>
> Participant: So I'm telling them, I'm sending them information. That they can now war. That they can start.
>
> Interviewer: And so this is a what? Is this a status, is this a command, is this something else?
>
> Participant: It's information because it does require context. But it's like a status it's saying: "you can now war. We can start fighting."
>
> Interviewer: So it's a status. What other flows do you have to the internal clan players?
>
> Participant: Apart from that? We obviously maintain that they're sending stuff back to me.
>
> Interviewer: What are they doing there?
>
> Participant: The line going back. They're sending me, also status updates. Whether they're ready to play, whether they're there. How much AP they have and things like that.

Interviewer: And they're sending you these as?

Participant: Textual information.

Interviewer: So it's text information?

Participant: Received via MSN.

Interviewer: Do we want to have MSN here or is MSN not at this level?

Participant: No, MSN is at that level. I'd say it is. I see MSN as – it's true, these would in essence, I don't see them as MSN. MSN is a like a tool or a spanner. As an intermediary because I send it here (MSN bubble) and then to there (Player bubble.) And that is true because I don't talk directly to them, per se.

I start by exploring the entities that we described in the entity diagram along with a process that has come about out of small talk. The process of diagramming a single process is about the right complexity for an SDFN. As we can see, the advent of a second SDFN in Case-4 meant that the process of "rezzing the fort" was a little too simple. It costs nothing to make a second diagram if there is sufficient time remaining.

The other important element is the requirement of asking questions. The point of the interview is to tease out the understanding of the participant, and to do that, they have to keep talking. Open questions, confirmations, and other prompts keep them talking without guiding down them any specific direction.

Interviewer: So what we'll want to do here is create some sort of representation so that you can talk about classes of people. Or individually people if you feel that they're important to be talked about as an individual. So, from this, you can say, "Well, this is a communication from me, to the active Clan Masters." "And I do this sort of communication." "Or, this is a communication from me to [Active Clan Master 1]." Whatever represents what you're doing.

Participant: Look, cause sometimes – Well, the information we can talk about, what, information, data? is transmitted. That's why I sort of did them separately. Cause I want to transmit me and [Active Clan Master 1] will talk about something differently than we'll talk about with Os here. That can be a slightly different form of communication with these and with that. You deal with the individual differently even in a group.

Interviewer: And that's what I'm trying to tease out here. Thank you. We've got you as a player, you as a clan master, the Devs. Right now, we're just going to make bubbles and we'll take these bubbles to a third page when we're drawing lines. This is brainstorming.

Participant: We'll keep it sort of then at a higher level.

Interviewer: Whatever you want to do, if you want me to draw it, I'll draw it. If you want to draw it, go for it.

Participant: Nah, you might as well do it. So there's other CMs then, as part of the clan. And, I'll use that term to classify all the people that can rez the fort, per se.

Interviewer: Okay, so you define CM as someone who's able to rez the fort?

Participant: Well, not really. But I'll group them together. There's what we call MA in [Clan] which aren't Clan Masters per se, but they have almost the same power as a clan Master. They can't be kicked. But I'll group them together. Because they're really, as you know in [Interviewer's Clan]. There's either those at the top and there's the rest of the players. So I'll call that other clan masters. And then there's other players. That's probably an easier, higher level distinction up there. And the next natural bubble is "other clans."

Interviewer: So I've got other clan players, and then other clans?

Participant: Um. No, I would put them together, so just other clans.

Interviewer: What about other clan masters? Or is that other clans?

Participant: I would put them as just other clans at the moment.

Interviewer: What about non-clan players? Are they relevant to this?

Participant: No, they're not relevant.

Interviewer: Is facebook relevant to this?

Participant: Yes it is actually.

Interviewer: And I will absolutely want to why you think that.

Participant: Yeah, that's interesting, maybe for our future wrap-up discussion.

Interviewer: So we've got FB, we've got other clans, we've got other players.

Participant: So that's other clan members. They're like, they're other clans. Like [Enemy clan name.] And these are other clan players in our clan.

Interviewer: Oh, other internal clan players. And we've got other CMs, Now we have the Devs. Do we want to say battlestations?

Participant: Yeah, I would just assume we're talking about [game].

Interviewer: No, do we want to have [the game] as an entity?

Participant: As the game? We could do that...

Interviewer: We don't have to use any of these. We've got you as a clan master, and you as a player. Is there anything else you think we will want to render beforehand?

Participant: In terms of my professional life as a clan master? No, let's go with that for now.

Interviewer: It's not stuck in stone. Sorry, cast in steel. Let's start by diagramming just the basic data flows. Where should we start? Just the most trivial.

Participant: Well, I guess that the most trivial is between myself and the game.

Interviewer: And now this is you as

Participant: Me as a clan master to the game.

Interviewer: CM to [game]

Participant: That's why I was happy to have [the reference to game as an entity] in there.

Interviewer: So there's data flowing from you to the game?

Participant: Yep.

Interviewer: Or... stuff? What would you label this as?

Participant: I'd label that as sort of data. I just send it stuff. It doesn't need any context. It's stuff like "I. Am. Going. To. Rez. The. Fort."

Interviewer: When you say 'stuff' What do you mean?

Participant: It's almost like discrete packets.

Interviewer: So, packets of?

Participant: Data?

Interviewer: Commands?

Participant: Like a command. Saying: "I want to do something. I want rez the fort now."

Interviewer: So is it OK to label this as commands, or does it encompass something other than commands?

Participant: It does, sometimes...

Interviewer: So commands and..

Participant: No, that's true. It's pretty much just commands. like I would send it I could do many things but they're all related to telling the game to do something. I don't call it knowledge, because it's like it's just, maybe they're just talking about context thing. It's something fairly simple, discrete. It's not open for arbitration or anything like that. It doesn't require arbitration, just "do it."

Interviewer: Let's build from here. What other trivial communications or interactions do we want to label here?

Participant: Well, it would be [game] back to me.

Interviewer: You as Clan Master?

Participant: Yep. That would be - its funny - I'd describe that coming back as information. It's telling me that its rezzed the fort or its kicked a player. it's just done something. "I've promoted someone!"

Interviewer: So you'd say this is information?

Participant: Yeah.

Interviewer: What information are you getting?

Participant: Well it's information about the status of something within the game

Interviewer: So status information? So you'd be getting communications of status which are information?

Participant: Yes. That's right. Yeah. "The fort is now up." I can tell the HP of the fort has gone up or that someone's changed a rank.

Interviewer: What other interactions are there in just a trivial level?

Participant: With other bubbles? Or just with [game]?

Interviewer: Let's elaborate from each of these bubbles and just grow it.

Participant: So the next would be – Maybe if I described the process ...

Interviewer: Walk me through the process.

Participant: Well, this is the case. I say "It's time clan war." This is where I bring in the extra bubbles.

Interviewer: Let's trace this and see where we get off these bubbles from the process.

Participant: We're at clan war. "I want to rez the fort." A command to rez the fort, it sends me back. Now, I guess to introduce the other bubble here as other clan members.

Interviewer: So, we're going to say internal clan members? Can we say other than members because that conflicts with clan master?

Participant: Clan Players?

Interviewer: Players. ICP.

Participant: I'll leave you with the acronyms. I tell them something now as well.

Interviewer: Now, when you tell them something.

Participant: There's a lot. That's a really detailed line between us and

Interviewer: Do we want to multiple lines?

Participant: Yeah. That's cool What we're dealing with at the moment. I'm going through the process of rezzing the fort, which is a common thing we want to do.

Interviewer: So shall we label this rezzing the fort?

Participant: So I'm telling them, I'm sending them information. That they can now war. That they can start.

Interviewer: And so this is a what? Is this a status, is this a command, is this something else?

Participant: It's information because it does require context. But it's like a status it's saying: "you can now war. We can start fighting."

Interviewer: So it's a status. What other flows do you have to the internal clan players?

Participant: Apart from that? We obviously maintain that they're sending stuff back to me.

Interviewer: What are they doing there?

Participant: The line going back. They're sending me, also status updates. Whether they're ready to play, whether they're there. How much AP they have and things like that.

Interviewer: And they're sending you these as?

Participant: textual information.

Interviewer: So its text information?

Participant: Received via MSN.

Interviewer: Do we want to have MSN here or is MSN not at this level?

Participant: No, MSN is at that level. I'd say it is. I see MSN as – it's true, these would in essence, I don't see them as MSN. MSN is a like a tool or a spanner. As an intermediary because I send it here (MSN bubble) and then to there (Player bubble.) And that is true because I don't talk directly to them, per se.

Interviewer: MSN for information communications?

Participant: It can be annoying if MSN is down, because we lose the ability to communicate.

Interviewer: Now, does battlestations use MSN as an information conduit?

Participant: No

Interviewer: So, how can we differentiate these?

Interviewer: What other flows are there?

Participant: Add myself now, as the player bubble. I don't know if we want to start a new sheet here whether we want to say this is going to be sort of rezzing the fort Now, I guess if we say, "we've rezzed the fort, we send the messages, we get updates, we send [game] commands, it sends back that we're ready to go or that we're up."

Interviewer: So, is this a complete rezzing the fort sequence?

Participant: Pretty much. Not quite

Interviewer: Not quite. What are we missing?

Participant: That we've come to the other CM bubble here, I guess. And that would be. It's a tough one, it's really a unified sort of... it's a knowledge maybe?

Interviewer: Is it one directional?

Participant: It's definitely two-directional, it's much more than the other things were Where we're sharing really information on "well, we want to rez, can we rez, what's going on? Who's up?" "No one's showing up. Shit I don't want to do this." It's that sort of... group 8-10 people on the list that communicate. So it's not even like... I'm certainly not sending commands. I'm not really sending data. I'm sort of sending messages. But anything or nothing could come back.

Interviewer: And you'd classify this as knowledge?

Participant: Yeah. But we really are sharing stuff.

Interviewer: Back and forth sharing?

Participant: "Oh, I can't make it just then, I'm busy this weekend." "I don't want to go now. Can we leave it for another time?" It really is a knowledge thing. Because, collectively we're gaining a better context of what's going to happen to then permit these other things.

Interviewer: What would you say this should be labeled as?

Participant: I'd certainly describe it as knowledge. Because you can really gain something out if it.

Interviewer: We need to talk about that more. But I actually meant the knowledge flow. What flow of knowledge is it? How can we refer to this flow of knowledge? Or is it a flow of knowledge of knowledge?

Participant: I don't quite get you.

Interviewer: We have commands are data. We've got a flow of data that are commands. Here we have a bi-directional flow of knowledge that is organization? But it's not the organization that you're... what category of knowledge is this? What kind of activity is it?

Participant: Sharing? Sharing knowledge? It really is sharing. That's what I said.

Interviewer: There's a process of sharing that goes on before and during the rezzing of the fort process?

Participant: Yeah. It's primarily before. It's certainly part of the thing. That is not the complete picture. It doesn't just happen on its own. Even if it doesn't happen all the time, there's still a sharing of knowledge about what's going on and who's doing what.

7.1.3 Personal Reflection

The personal reflection for Interview 1 is presented here as part of the expanded discussion around this interview. All other reflections for other interviews can be found in the next chapter.

Participant used the term information to describe a communication of meaning. In this interview, the term was used to describe communications of "status." In the sense of the diagram, status is ambiguous. Because people describe their status to the people in charge, it is a communication between people and not a technical communication. Yet it is also the game reporting the status of the data representation of the player to the player. Supporting this, the term "information" describes communications that are explicitly privileged above data.

Participant used the term knowledge to refer to expertise. Knowledge can be shared and is asserted to be a communicable view of reality. The players in charge of the group of players explicitly engage in knowledge sharing, and impart that view of reality to their apprentices.

Participant, in the SDFN diagram, seemed to use data to refer to contextless communications. These communications can originate from a computer or a person, but they fall into two significant categories: activity causing and unprivileged.

In the activity-causing context, participant described communications *to* people and computers. While communications from computers are information, the commands to the computer are data to the game, and explicitly contextless. Data can also be transmitted to a person and is a simple alert designed to cause activity. Curiously, in the same category, messages from the players in charge to the rest of the group are also data. They seem to be contextless and simple instructions. There seems to be a different ontological structure between commands to the game and commands to the players, despite both being described as "data."

The unprivileged communications context seems to be attached to communications where the senders cannot know what they are talking about. In a sense, such a communication must be viewed with skepticism: it is a minor or inferior form of informa-

tion without reliability. In this context, it describes communications from apprentices to the people in charge. Participant believed they do not know enough to offer information. As such, their communications are only data, explicitly described as "opinions" that do not have any basis for action.

The SDFN diagram suggests that the participant has two different understandings of data simultaneously and suggests no way to reconcile the two.

7.2 Interview 2

Participant: So we can do a typical Information request for example. It starts off with somebody asking me a question.

Interviewer: Which entities to start with?

Participant: Entities as in an individual?

Interviewer: Role entities?

Participant: Let's do with a position 86 because they're close.

Interviewer: To you as?

Participant: In my role as an position 87.

Interviewer: There is something... what is this flow?

Participant: That flow could be either a person 1:1 chat or maybe phone or email or something like that.

Interviewer: What is it a communication of?

Participant: Information I suppose you could say.

Interviewer: How can we label it? The position 86 is communicating with you Information. What is the context of this Information? What are they communicating with you?

Participant: They're asking a question.

Interviewer: So the position 86 communicates with you a question. You would say this question is Information? This is the noun 97 we'll be doing.

Participant: Just keep prompting me.

Interviewer: Then what happens?

Participant: It depends upon the question. If we do a typical question. Say they're looking for research work in a particular area. It would mean that I search a particular database.
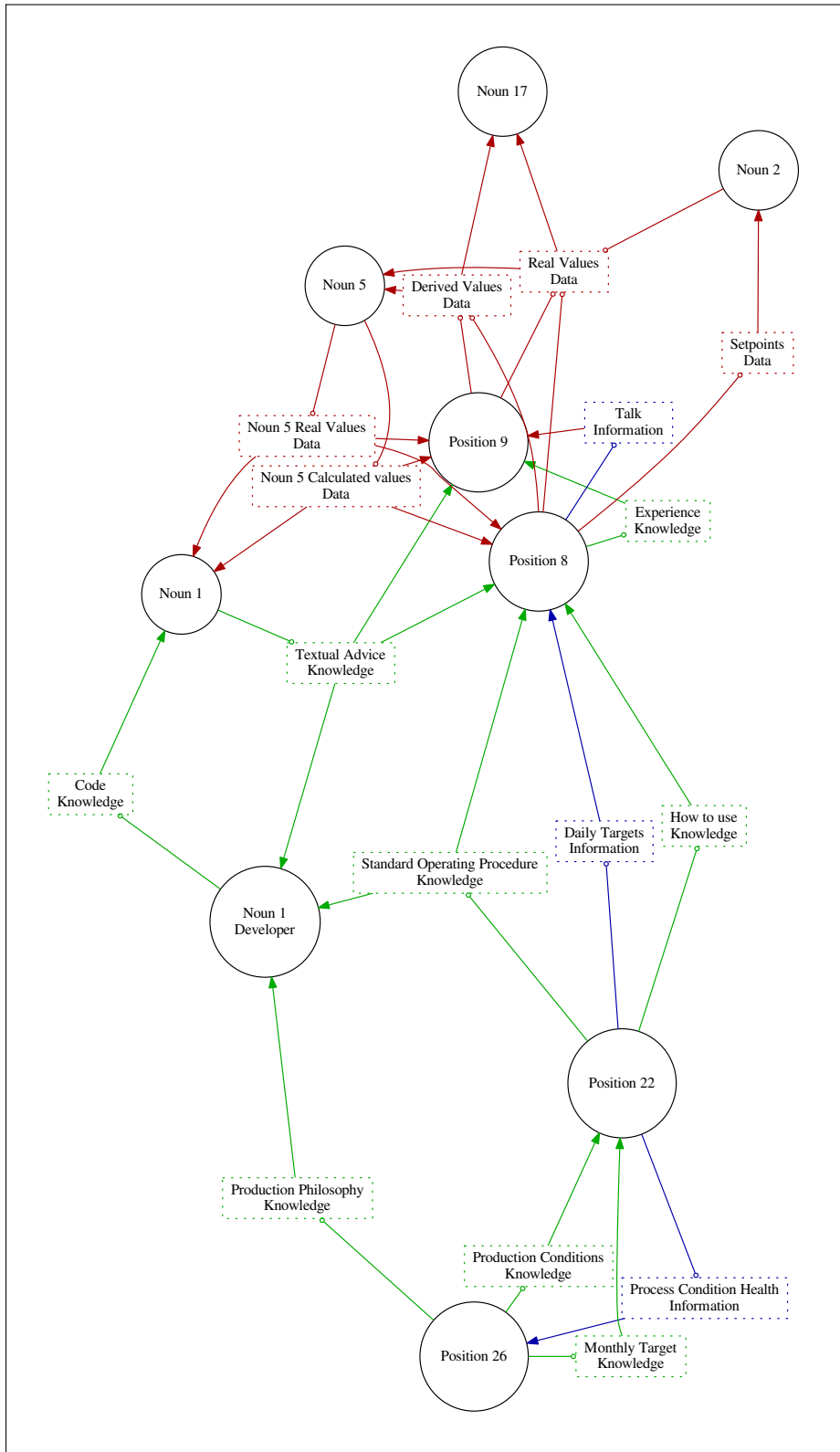
Figure 7.4:   The SDFN Diagram for Interview 2

Interviewer: What entities are involved in you figuring out what database to search?

Participant: That's from experience I suppose.

Interviewer: Is there a flow from one entity to another or is this innate experience?

Participant: It's experience.

Interviewer: Is this experience, Data, Information, Knowledge, or O?

Participant: Whether is it Data or Knowledge. I suppose it's Knowledge really. Data to me is not – I say it's Knowledge.

Interviewer: You use your Knowledge as experience to find a database. Is it safe to generalize to "database" here or shall we be more specific?

Participant: No, database.

Interviewer: What flows are there from you to the database that your expertise has chosen?

Participant: The Knowledge is that I know which one to search.

Interviewer: You have internal Knowledge of which one to search. Is there some sort of confirmation action that you do to confirm that Knowledge or is that tacit?

Participant: It's tacit.

Interviewer: you now know which database to search. What is the first thing you do when you have found that DB?

Participant: You answer that question.

Interviewer: So you send something to that database?

Participant: As I said, it's looking for perhaps a research topic. So we're going to that topic request.

Interviewer: From the question you would generate a topic request. This topic request – this flow, can be categorized as Data, Information, Knowledge or O?

Participant: Data.

Interviewer: We start with you as position 87 sending a topic request through data flow to the DB. Then what happens?

Participant: You either get a result or you don't.

Interviewer: Let's assume the best case.

Participant: You get a result. You get Data which will then flow back.

Interviewer: So we have a flow backwards which is Data and this is a result?

Participant: Which would then be communicated back to the position 86.

Interviewer: Do you perform a transformation on this result?

Participant: By transformation, if it's actually a physical entity, you might actually provide that to the person. Not just Information, if it exists like a copy or something.

Interviewer: So one case would be that you and the position 86 get this result. This is in one case. Are there other cases?

Participant: That's assuming we're successful. If it's not successful then we may need to go elsewhere. That would be the first choice of checking that DB.

Interviewer: This iterates around and around until you...

Participant: It might not be a database, it might be asking individuals for example. So you go to a different source.

Interviewer: Let's diagram that. Which entity are you talking to as a class.

Participant: It would be another position 86.

Interviewer: How can we label that other position 86 different from this one?

Participant: position 86 B. So I've gone down this loop, got nothing. Then I would contact that person.

Interviewer: Before we go into that, is the getting of nothing a data flow?

Participant: Yeah it's a flow.

Interviewer: What it's a flow of? Let's start by indicating a flow. This is basically a null result.

Participant: It'd be a flow back to position 86 A saying that didn't work.

Interviewer: So we've got this null result. Is this null result flow from the database to you Data, Information, Knowledge or O?

Participant: ??? <From diagram it looks like Data>

Interviewer: Then you perform – you send something back to the position 86?

Participant: It would be Information back to them.

Interviewer: And what is this Information?

Participant: Information saying "I now will have to actually look further." That's when we start to look at people like position 86 B.

Interviewer: What covering name can we give to that "look further" communication? Is it like a status update?

Participant: Yes.

Interviewer: Just to step back for a moment... let me know when there's enough. We give a status update back, they may or may not refine their question. Eventually you may need to talk to a position 86. What is it that you communicate with a position 86?

Participant: It is that original question, that original Information.

Interviewer: Do you perform any manipulation of this question?

Participant: Possibly.

Interviewer: What manipulation do you perform?

Participant: It may be just narrowing down. Which could be Knowledge I know of various... they may have the wrong year of the research project.

Interviewer: Perhaps a refined question?

Participant: Yes

Interviewer: Is this refined question Data, Information, Knowledge, O?

Participant: Information isn't it?

Interviewer: You send, as position 87, to position 86 B, a refined question? What do they do?

Participant: They may know it in their brain already or they could refer me to someone else, position 86 C, or they may actually have their own systems, actually.

Interviewer: This, in a way, is atemporal. And, because of that, there's no causation. Let's model the set of normal responses that the other position 86s can give to you. What flows back are there from the position 86?

Participant: I suppose one is that they cannot help at all: "Dead end."

Interviewer: Is this Dead End Data, Information, or Knowledge?

Participant: Information. The other one is that they may refer you. That would be Information.

Interviewer: referral as Information.

Participant: Another one is that they check their internal systems, they may have DBs, something like that, and provide the answer.

Interviewer: So they are communicating with an entity.

Participant: Let's say they have their own DB.

Interviewer: They send a flow to a DB. What are they sending to the DB?

Participant: It would be that original question –the refined question.

Interviewer: The refined question as Data, Information, Knowledge?

Participant: It's Information isn't it? No, it's Data, sorry. It's Data.

Interviewer: Do they perform a transformation on this?

Participant: No, I don't think so.

Interviewer: And the DB returns what?

Participant: Something positive, basically.

Interviewer: Could we say it's the result / null result thing?

Participant: Yes. They get the same, a repeat.

Interviewer: They have a DB result, what then happens?

Participant: They refer it back to me.

Interviewer: As part of the referral or as a different Information flow?

Participant: It would be the same.

Interviewer: so they get a hit and they refer the hit back to you.

Participant: Yes that's right

Interviewer: and you perform some sort of status update and give them a response. What other entities do they consult? In a generalized case?

Participant: The research system? Their own brain? They may talk to somebody else. And it just repeats.

Interviewer: If they have neither referral nor a dead end, is there something else they can send back to you?

Participant: No, that would be the end of it, wouldn't it?

Interviewer: In their expertise they have some answer in their expertise is a referral?

Participant: Yes.

Interviewer: Say they dead end you, then what happens?

Participant: Go back to position 86 A to find out just how important it is. That's all to do with the priorities.

Interviewer: How can we model that flow? Is it part of the Status update or is it different?

Participant: I suppose it's status, it would just be a different priority in status.

Interviewer: You give them a status update. Do they then give you anything?

Participant: They either say keep going or they might have a refined question. This could go round in a lot of circles.

Interviewer: I'm getting that sense. They send back to you …

Participant: Information to say either stop or more Information to help assist.

Interviewer: Can we combine those in one covering statement.

Participant: It's just a re-request or continuation – priority.

Interviewer: Because they're clarifying what is a priority.

Participant: Yeah, that's right.

Interviewer: You send them a status update back one way or another. They prioritize…

Participant: Yeah.

Interviewer: Are there any other entities involved in this noun 97?

Participant: It again depends on the question. If it's definitely internal. You stick with internal, you may get a dead end. Or if it's something that perhaps somebody outside the organization might be able to assist with then you do the sort of steps with other people.

Interviewer: You then do position 86 external.

Participant: Exactly, or DB, external. That sort of thing.

Interviewer: is there an entity which is neither a database nor a position 86? Or is there a different type of person that's meaningfully different?

Participant: It could be another position 87.

Interviewer: Let's model that flow. You as position 87 A, talk to position 87 B. And what do you communicate with them?

Participant: The refinement – it would probably be the priority one by this stage.

Interviewer: Refined question with priority. Information?

Participant: Yes.

Interviewer: And then?

Participant: And then they would do something quite similar to what I just done. They would consult a system or an individual.

Interviewer: Same flows?

Participant: Same flows.

Interviewer: If they return, from their sources, dead end or null result.

Participant: Which is possible.

Interviewer: What happens?

Participant: Back to position 86 A via myself.

Interviewer: What do you get and what does the position 86 get?

Participant: Not a lot. But Information though. Or as in 0.

Interviewer: What label do we want to put on not a lot? We can absolutely call it not a lot.

Participant: No result, basically.

Interviewer: And you communicate that back to position 56 A through status update?

Participant: Communication, status update.

Interviewer: If they don't get no result ...

Participant: Sometimes they're quite happy. Sometimes they're just testing it out. That's what position 86s... they don't actually want you to find anything. Particularly if it's a patent or something like that. They go through all that

Interviewer: due diligence.

Participant: But other times they're very unhappy, so it depends.

Interviewer: Besides no result, what would you get from position 87 B?

Participant: Probably a bill, an invoice. Just a status update basically saying they've searched this, this, and this, and that it hasn't been successful.

Interviewer: Do we want to label this invoice or status update or both?

Participant: both.

Interviewer: Invoice is what?

Participant: It would be data, wouldn't it? Data, Information, what was the other one? Knowledge? No. It would be Data.

Interviewer: And they search status update you.

Participant: Which is Information.

Interviewer: What happens if they do find a result

Participant: they do? Then they provide the Information as in Data. Or they may do a status saying we have to pay them for this part.

Interviewer: So you get search status and invoice one way or another.

Participant: Precisely.

Interviewer: You said you get the Data as Information?

Participant: The status would be Information, the actual article or whatever it is is Data. The result... If we had a result that would be Data.

Interviewer: The search result would be Data as opposed to no result which is Information. We have external position 87s iterate over that stuff according to the priority. Other entities? What do you think position 86 A is doing with this request?

Participant: They're probably – they could be one of the people tasked in their team to work this out. There could be a team entity over there.

Interviewer: What interactions do you see position 86 A having with their team?

Participant: There would be some sort of Information communication between themselves and the team.

Interviewer: Then, to team, Information.

Participant: It would be the status update.

Interviewer: Do they perform a manipulation of the status update

Participant: They would just convey that Information to the team.

Interviewer: Does the team send any flows back to position 86 A?

Participant: They may to – it depends on this loop, doesn't it. It may have happened to get that priority for example. To say yes it was important to keep going. Then we went through this loop down here.

Interviewer: So is the team prompting A to send priority? Does A manipulate the team's prompt?

Participant: They stick it in the e-mail for a month or so and wait.

Interviewer: That actually is a manipulation. It's checking informally "is this important?"

Participant: Don't know. They may or may not.

Interviewer: If they don't, there's nothing here. If they do?

Participant: I wouldn't know unless it was in the trail from an e-mail, for example. They've sent it out and they sent the original team's response back, then I can decide what they've done. But otherwise, I don't know if they've manipulated or not.

Interviewer: What is your guess here?

Participant: Guess is that they don't manipulate.

Interviewer: so there is a flow from team to you of priority. But there is also a flow from A to you for sending this. But the impetus for the question is always from A.

Participant: it looks like a bit of a mess. That's what some days are like. You go around in circles.

Interviewer: Something slightly different. We have data flows to and from the DB system and the invoice. Are there any Data flows in that zone?

Participant: There would be if it turned out to be... we haven't finished off the – we've gotten the result.

Interviewer: Do you perform a manipulation of the search?

Participant: Maybe in a tiny way. Putting it in a format so that it's more acceptable to us here.

Interviewer: Which is absolutely manipulation. What shall we label this manipulation as?

Participant: Just formatting isn't it?

Interviewer: Formatted search result, which would happen with any of these entities.

Participant: Yes.

Interviewer: This formatted search result is Data, Information, or Knowledge?

Participant: It's Data.

Interviewer: Is there a flow of Knowledge at any point in this noun 97. Can there be a flow of Knowledge?

Participant: It could be from position 86 B? or it could be from the position 87. So we're assuming here that they're searching a system. They may, as I said, just be searching their tacit Knowledge.

Interviewer: Then they would communicate?

Participant: That's right. Write it, or pop it in an e-mail, that sort of thing.

Interviewer: If position 86 B is communicating back to you. What shall we label this Knowledge flow as?

Participant: It's just their Knowledge isn't it? Their experience.

Interviewer: So their experience?

Participant: Yeah.

Interviewer: What about Info Broker B?

Participant: Yes, I suppose it could be Knowledge there. It depends on – I'm thinking of specific Information questions. They're more like an intermediary that I'm thinking of position 87 B in that they may have gotten Knowledge from somewhere else, but they themselves, that entity, no. It's not the Knowledge, it's just the Data.

Interviewer: Can we assert that Info Broker b would just have position 86 Bx.

Participant: The sort of external systems similar to these internal. Yeah.

Interviewer: But they themselves wouldn't. They would just pass –

Participant: Intermediary.

Interviewer: You are then the recipient of their experience qua Knowledge. What do you then return?

Participant: Right, OK. Then I either just relay it. Or if it's too technical, I actually get them to talk to each other.

Interviewer: That sounds like two different flows. Flow the first is that you're relaying it. What is the flow here between you and research A?

Participant: It's just an exchange of Data.

Interviewer: It's just formatted search results?

Participant: That would be it.

Interviewer: You're formatting Knowledge as a search result.

Participant: That's right.

Interviewer: And that's Data?

Participant: Yes.

Interviewer: Or? You are referring B to A.

Participant: Yes.

Interviewer: Would this fall under formatted search results as well? With just different formatting? Or is it substantively different?

Participant: It's just Information. Isn't it? Not that much, no.

Interviewer: It's just Information?

Participant: Yes.

Interviewer: What Information is going?

Participant: Just Information saying yes, I do know where that ... contact.

Interviewer: This is obviously a superposition of many generalized cases. Are there any other common avenues of exploration in this?

Participant: No, that's good enough for that particular topic.

...

Interviewer: Let's move to a more theoretical discussion. You label yourself, in one of your roles, as an position 87. To you, what is Information?

Participant: It's something that's not physical, basically. That means it could be a communication, it could be a conversation or a story. Something verbal.

Interviewer: So Information is any non-...

Participant: Data to me is physical, basically. It's an entity that filters?? an entity of some sort.

Interviewer: You would say that letter is Data?

Participant: I would say that letter is Data. But what is on that Data is Information. Just to be a bit confusing.

Interviewer: Tell me more.

Participant: Coming from the field we're in, in Libraries and Record group 130. You've got different areas. It would be, as I've said, Data or record. That's just the physical entity.

Interviewer: Data is a container...

Participant: Yes, of the Information.

Interviewer: And Information is content of what type? Is there something common to all Information?

Participant: it's not easy. I'm struggling.

Interviewer: Frankly, this is why I'm doing my PhD

Participant: It's a struggle. Information, what is Information? It's just... No, I'm drawing blank.

Interviewer: We have right now Data

Participant: And then of course you get Knowledge.

Interviewer: We have Data, Data is a container for Information.

Participant: I'm quite happy with that.

Interviewer: You say physical at some point?

Participant: Yes, physical in t??. It doesn't actually have to be physical in a piece of paper, but it can be physical as in an e-mail message.

Interviewer: Physical as real?

Participant: Yeah, physical as in real.

Interviewer: Except that doesn't really help.

Participant: What about in terms of semiotics.

Interviewer: Is there some sort of semiotic connotation or denotation of data? So this is Data containing the Information of the entities we described.

Participant: Yes.

Interviewer: Container is actually the correct word.

Participant: I think so.

Interviewer: Because it gets fuzzy of what the nature of the container is.

Participant: Yeah.

Interviewer: We have Information.

Participant: Information is … Was talking about Knowledge. Knowledge to me is more than just Information. If we go backwards perhaps. Knowledge, you actually have to associate it with an experience.

Interviewer: So Knowledge is experiential?

Participant: Yes. Whereas Information is just there. It's provided nobody actually read it or analyzed it or anything like that. Whereas Knowledge has been more analyzed and used. And I'm getting very vague.

Interviewer: Knowledge is Information which has been analyzed?

Participant: Yes, yes that's right. It's the next step up, I view it as.

Interviewer: Analyzed and used Information. Can there be a – Knowledge is experiential. This experiential Knowledge allows us to do what?

Participant: That's how a doctor knows how to operate, basically. It's because they've had Information, and they've analyzed it. So that's the tool for them to operate. It's like somebody else coming in with just having some Information. But they don't have the actual analysis or expertise. Which seems to be the Knowledge that the surgeon has.

Interviewer: So a surgeon has expertise about the world, and his expertise about the world is Knowledge.

Participant: Expertise about his Job, world, whatever.

Interviewer: Can this expertise about his or her job be verbed into Information?

Participant: I suppose it could be?

Interviewer: which verb?

Participant: I don't know?

Interviewer: But there is a movement back.

Participant: Yes.

Interviewer: Is there a movement from Information to data?

Participant: No, not in my view. I view data as something completely different.

Interviewer: We've got up and down arrows [between Information and Knowledge] but we don't have an arrow there [to Data]. We have Data as a bucket.

Participant: Yes, as a bucket, in my mind, whatever.

Interviewer: of Information or Knowledge?

Participant: and/or Knowledge

Interviewer: so there's a difference in nature or kind between Data and these other two. But not between these two [Information and Knowledge]

Participant: No, these two get fuzzy.

Interviewer: Information can be turned into Knowledge?

Participant: Information can be turned into Knowledge.

Interviewer: Through analysis and use?

Participant: Yes.

Interviewer: Knowledge can be turned into Information through what?

Participant: I suppose it depends on who uses it basically. It's the user of that. It's somebody with ... not that. But what it was. Its like it's been that level of ... It's like dumbed down in a way. Somebody reads this and they've got the Knowledge, but somebody else reads it and it doesn't make sense to them, it's just Information. Because they don't have...

Interviewer: Where is the sense making component here?

Participant: I don't know.

Interviewer: The first question is: is there an external sense-making component?

Participant: I don't know what you mean.

Interviewer: If I'm reading that, and I'm a lawyer,

Participant: that's right, OK.

Interviewer: Versus that and I'm a painter with no Knowledge of the law.

Participant: That's where I would say that the painter would just be reading that, and it would just be Information to them. But the lawyer would then become Knowledge. Because they would read into it their experience or whatever.

Interviewer: The lawyer would apply his or her Knowledge? Or incorporate that into his or her Knowledge?

Participant: Would incorporate that into.

Interviewer: So that's part of the analysis

Participant: Yeah.

Interviewer: say you're teaching someone as an apprentice. Is there a flow of any of this sort in that kind of teaching?

Participant: Teaching you would start off with the Information, and they would through the exercises and the training, it would be … their Knowledge.

Interviewer: You would cause them to develop

Participant: Knowledge. Their Knowledge, yes.

Interviewer: You're not communicating Knowledge to them.

Participant: No they have to do something to create the Knowledge.

Interviewer: would it be safe to say that Information is some sort of vehicle for Knowledge?

Participant: Yes.

Interviewer: What kind of vehicle is it?

Participant: It's the primary vehicle, really.

Interviewer: So Information is the way that Knowledge is communicated?

Participant: Yeah.

Interviewer: Can Knowledge be communicated? Say you're talking to position 87 B, outside of this context, as equals. Can you ever give them a flow of Knowledge that's Knowledge a opposed to a flow of Knowledge encoded in Information?

Participant: I'm not sure.

Participant: I suppose – you do assume that they have a level of Knowledge, so yes, it would be an interchange. You'll start to use, perhaps, jargon or something, assuming they know what you're talking about.

Interviewer: Let's talk about jargon. What is jargon?

Participant: Jargon, I suppose, would be a special language between specialists.

Interviewer: A verbal shorthand?

Participant: And each discipline would have its own.

Interviewer: Jargon is what?

Participant: Just a cryptic way of exchanging Information.

Interviewer: Jargon is a cryptic or short way of–

Participant: A shorthand

Interviewer: a shorthand for Information.

Participant: between specialists.

Interviewer: that's not a Knowledge exchange.

Participant: No, it's Information exchange.

Interviewer: Can there be a Knowledge exchange without using Information?

Participant: I'm still, I don't know. There probably is but I can't think of examples.

Interviewer: Say I'm writing my dissertation. Clearly, because I'm typing into a computer, Data is the container, yes?

Participant: Yes.

Interviewer: Am I typing my Knowledge into a computer, or am I typing my Information?

Participant: You're typing your Knowledge. Because you've been through the noun 97 of filing the Information.

Interviewer: So I've collected Information through whatever, and I've got all the Knowledge that I've got. The noun 97 of putting Knowledge into a Data container, that does not touch Information? Or does it?

Participant: I don't know. As you say, you could just be putting it in there, it's not Knowledge, it's just word for word from somebody else. You're not relating it or transforming it or anything. So you're not adding to it. As I said, if I come back to the feeling that Knowledge has been added to.

Interviewer: As you know I'm going to be transcribing this recording. The noun 97 of me typing this recorded conversation falls where in this?

Participant: It's Information. It just depends upon if you make any changes. If it's a straight word for word, that's just Information because you haven't altered it in any way.

Interviewer: So our communications are Information.

Participant: Yes.

Interviewer: It's recorded on the computer as Data.

Participant: Yes.

Interviewer: Me typing it would just be changing the Data container?

Participant: yeah. That's right

Interviewer: Because it's just Information, because there's no analysis or use.

Participant: By the time it gets to your thesis, it turns into Knowledge.

Interviewer: Which I'm encoding into Data. But other people pick up a conference paper I generate from my thesis and theoretically read it. That conference paper, to them, is Knowledge or Information?

Participant: It could be either. It's more Information. It's initially Information until they choose to do something with it. Until it adds to their armory of whatever their skill set is.

Interviewer: Really, Information can produce Knowledge, but Knowledge can produce Information because someone is communicating their Knowledge to you as Information and it's not until you analyze it that it becomes their Knowledge.

Participant: Yeah. That's what it seems to be.

Interviewer: Could it be that the Knowledge to Information bit is abstraction?

Participant: Oh yeah.

Interviewer: If I'm communicating through a conference paper to you, my Knowledge, but you're getting it as Information, I'm abstracting our my use and analysis of it, i.e. my internal tacit Knowledge. Can we think of a better word than abstraction?

Participant: No.

Interviewer: I want to make sure I'm not putting words into your mouth.

Participant: That's OK. There's no word there at the moment.

Interviewer: Final question. We have Data, Information, and Knowledge. First, is there anything to either side of Information or Knowledge?

Participant: Not that I can think of at the moment.

Interviewer: Knowledge is the worldview and Information is abstractions that can eventually be?

Participant: Yes.

Interviewer: If I'm making a prediction. Let's do a really trivial one. I predict when I release this pen, it will drop. What is that?

Participant: That's probably Knowledge, because ...

Interviewer: If you say to me, go research X, what is that?

Participant: That's just Information, just telling you to a task of some sort.

7.3   Interview 3
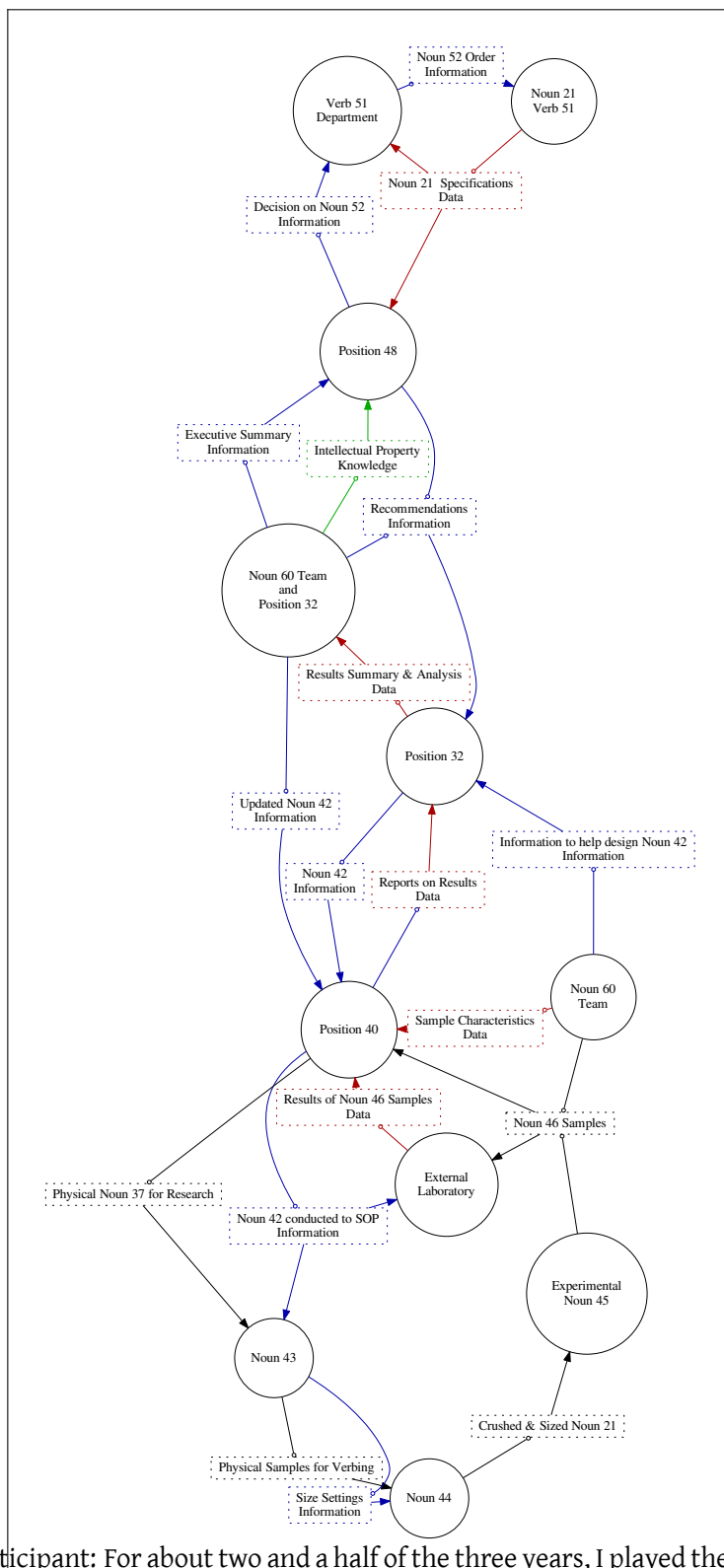
Participant: For about two and a half of the three years, I played the role

Figure 7.5:   The SDFN Diagram for Interview 3

of position 100 of the project. And even that name was an interesting name, because it was felt that the invent, although a very clever person, was not a good position 41 of people. So they needed somebody who could come in and noun 41 the project. And, of course, keep the position 101 happy at the same time. So there's a little bit of interesting stuff there. It was fine, I did never fight with the position 101, we both fought Os together.

Interviewer: So we've got the position 100, we've got...

Participant: So there was a group 104 that the position 100 was the leader of. We were a fulltime group who mainly sat together.

Interviewer: Let's look at this interaction. Did you send anything to the group 104?

Participant: I sat among the group 104. We're talking about a time before there was personal computers, before there was e-mail. And so, there were of course, telephones, etc.. So we sat together. Typically, we would talk for a period each day about what we were trying to do and who was doing what.

Interviewer: Let's model that talk. That thing is, even though you're sitting together, there are flows of stuff. When you're sitting together in his meeting each day and you're talking to each other. What was one of the categories of things you talked about, that you told them?

Participant: In the early days, what we were trying to do, there was the verb 102 part of it. So we were developing verb 102. Then there was another thing early on with equipment verb 52. It was quite important to get the right equipment because the aim that we were heading for, that we realized we were heading for, was some kind of automated noun 75. It was agreed without having been proved, that the position 101, he had the ability to control the noun 103 better than the normal position 90s. He said it I do it by using the noun 111, and I use the noun 111 this way. People were suspicious and thought that he was taking more cues than were actually in the noun 111 and people wanted to say: "Can a machine understand this noun 111? and can a machine take the same kind of noun 97 changes?"

Interviewer: Let's look at the very trivial equipment verb 52. You as position 100 are talking with these folks about equipment verb 52s. ...

Participant: What we within the group 104 needed to do was to work out what was the most important pieces of equipment to have for purposes of conducting a noun 75. We were discussing it and working it out together. And of course we've got another bubble of the people who were actually going to fund it. That was the group 105.

Interviewer: So the group 104 sent to the group 105 a proposal?

Participant: yes.

Interviewer: Would you say that this proposal is Data, Information, Knowledge or something else?

Participant: Probably Information.

Interviewer: so the group 104 sends to the group 105 a proposal that is Information.

Participant: It is probably Information noun 95ed to persuade the group. It's not Data, it's brought together somehow. You mightn't agree.

Interviewer: The thing is, I'm a passive receptor. I by definition agree. Does the group 105 send anything back to the group 104?

Participant: The group 105 will either agree with the proposal or not agree with the proposal or do nothing.

Interviewer: Are these separate flows, or is it a response flow?

Participant: It's a response flow. Typically there would be two ways of doing this, there would be a formal way where there's a meeting where the decision was taken or there would be a less formal way.

Interviewer: Are they different?

Participant: The result would be pretty much the same.

Interviewer: Is the result, Information, Data, Knowledge, O?

Participant: I dunno. We'll go with Information at this stage.

Interviewer: Don't feel compelled to choose Information just because of the other two.

Participant: It could be Knowledge, I suppose.

Interviewer: What's the best fit?

Participant: I'm struggling with these definitions in my own mind.

Interviewer: If we're struggling with the definitions we can do a theoretical conversation if you think that that would help you more or I can continue poking your intuition.

Participant: Let's go intuition. You might discover what I actually think along the way.

Interviewer: group 104 position 100, what are they sending you?

Participant: I think mostly Data.

Interviewer: what is one of the Data flows they send to you?

Participant: There's lots of stuff about just the daily occurrences, progress.

Interviewer: Are daily occurrences data?

Participant: I think so. The way I'm thinking of them, it's not anything that's been considered or noun 97ed, it's just Information. I shouldn't use the word Information, it's just what's happening.

Interviewer: you can absolutely use the word Information if you want.

Participant: It's just what's happening.

Interviewer: If you want to use Data and Information as synonymous, you can.

Participant: I'll try not to, I'll try to use some sort of hierarchy of those words.

Interviewer: If you don't consider it to be a hierarchy...

Participant: I do consider it to be a hierarchy, but it's not very clear in my mind.

Interviewer: so they're sending you daily occurrences in a data flow.

Participant: Lots of data, which I'm handling to try to understand what's going on. As a position 100, in this particular thing, position 100 was an euphemism for what I was being asked to do. And what I was really being asked to do was to be the leader of the group. Without offending the position 101.

Interviewer: Would you say that your role as group leader was different from your role as position 100, or was it just a polite label?

Participant: It was a polite label. But I tried to run a very democratic type of group. But I think when you work in a hierarchy, people don't have to have it all that democratic. But you have to have some reality that people can say what they need to say anytime. And they can question any decision.

Interviewer: position 100 back... So you're assembling this. Are there flows of stuff there, perhaps to yourself or yourself in other roles?

Participant: I think in this particular case there were lots of flows of things, because in the group 104, in the main group 104, there was a position 106, a position 107. I'll call him the position 107. There was the position 101 and there was the position 108. Now, we all had to work together and my job was to lubricate all of these things to make sure that everything was happening correctly. There was a natural suspicion of the position 101 with the position 107. The position 101 and I were pretty reasonable friends. The position 108 and I were pretty reasonable friends. And for a lot of the time the position 107 and I were pretty reasonable friends. We never got to be enemies, but we got a little bit separated. There was a bit of tension across this way, and it was my job to make sure that everything happened as it should happen. The position 101 couldn't do the noun 109. And the position 107, wanted to be acknowledged for being a position 107 and didn't necessarily want to get ... he wanted to be a bit more of the – not necessarily to delve into the depths of the detail. A bit more of an adviser. And the position 108 just wanted to get on with it. Which is exactly the right sort of person to have. And it was my job to work on each of those issues. As I mentioned to you, this issue of noun 110 in this noun 111 was a tremendously important part of this endeavor, so I'll talk about that a bit more. Things like knowing the hardware and so fort that the group wants, we dealt with it. We worked out what we wanted and the group 105 gave it to us, etc.. But with this – with respect to noun 110. ... The noun 111 had various shapes that you could observe by eye .... And the noun 110 thing became a classification thing. Now, in the early days of this, the best we could get to was about 65

...

Interviewer: Your conjectures... So we've got position 100?

Participant: So therefore I became the position 113 of that meeting, which everybody allowed me to be.

Interviewer: Is position 113 a better term for modeling that meeting than position 100?

Participant: Probably. But I was a participant as well as a position 113.

Interviewer: So there's you as position 113, and there's you as participant. What role were you playing as participant?

Participant: I was the position 114 ...

Interviewer: So we'll go with position 114?

Participant: Sure.

Interviewer: Because there's time when you could be sending stuff to the position 113.

Participant: Yes, I could be. The position 113 always listened to me.

Interviewer: What other roles did we have at that meeting?

Participant: The position 101, and the position 108, and the position 107. That's who they were. You're asking me what their roles were? The position 107 was mainly a position 115 of the current verb 112 technology. The position 101 was very passive. So it was a passive role, but I don't know if I could put another word to it. Write down passive until we get to another word. And the position 108 was basically a supporter of the idea that we needed to change.

Interviewer: Did any of them play multiple roles?

Participant: So I played a multiple role. Those two roles. And I think actually our roles were changing during the day. .... And then I became the position 114 of that method. At the end of that meeting, I was very much an position 114 of that method. It's more perhaps part of the position 113 role. I had to keep this meeting to be constructive. A few, I think can remember, a few pretty straight truths were told to each other at various times of the meeting, but it was in a constructive way. The position 101 was pretty passive. He was wanting, quietly, everything to move to a new state. ... And the position 108 who was supporting it again was wanting to make sure that we got. I think there was probably in the three people, myself the position 108 and the position 101, we were all probably coming into the meeting thinking that the position 107 needed to move ground a little bit. There was this element of coercion, subtly there.

Interviewer: Let's start by looking at some of the flows of the meeting. Maybe with this coercion noun 97. What flows were there?

Participant: I think we probably started in a facts and data way. I would, as position 113, would have asked the position 107/position 115 to just give us an update on where he was getting to.

Interviewer: That request for update, should we call it request for update or something else?

Participant: Call for Information, I think, maybe Knowledge. If we had Knowledge.

Interviewer: Do we want to call it for something? What was it a call for?

Participant: I think in casual use of the word, it was a call for Information.

Interviewer: Was this call for Information, Data, Information, Knowledge or O?

Participant: It was Information and Knowledge. that's what I was wanting.

Interviewer: What were you sending over this call? What is the nature of that call?

Participant: I don't know how I would put it. Could you step out of your role for a moment and help me with what you think it might be. Do you have this often?

Interviewer: We don't have to identify it as Data, Information, or Knowledge. On the other hand, what you're doing here, you're telling them that they need to give you Information and Knowledge.

Participant: Maybe, what they give back. I ask for Information and Knowledge, but I think what they will give back are three things. Data, Information, and Knowledge. Probably all mixed in together.

Interviewer: So they give back all three.

Participant: It goes backwards to everybody.

Interviewer: What is the label of their response? What are you asking them to give, Information and Knowledge about? Their model?

Participant: Yes, it was about the progress about their model.

Interviewer: Okay, so we can say Call for Progress Update, because what you want is for them to give you Information and Knowledge about...

Participant: Exactly. Progress...

Interviewer: Now, what they give back to you, was it the progress update, or was it something else?

Participant: It was the progress update, and it was a whole heap of other noun 111s there, including non-verbal noun 111s about how happy they were with the world and those other things.

Interviewer: Let's break that out. They've got the progress update itself. Just that, would you classify that as Data, Information, Knowledge, or O? Just...

Participant: That's Data in my opinion.

Interviewer: But, along with that, they send non-verbal cues. To the same daisy-chain?

Participant: You've modeled it as a daisy-chain.

Interviewer: In the model it'll go there and then explode. But I need some way of... It's not a daisy chain, it's a simultaneous communication. They send the progress update qua Data. Then they send the same non-verbal cues in the same explody.

Participant: So the verbal thing can contain Data as well as Knowledge and Information.

Interviewer: What aspect of the progress update contained Information?

Participant: The Information about the overall performance about the noun 110 program.

Interviewer: So we have the progress update as Data. We have overall performance discussion as Information. And the second one is overall performance?

Participant: Yes. There's some details, some breakdown of that. It's not just a single thing.

Interviewer: But it's an informative flow.

Participant: Yes, it is.

Interviewer: Do they also send a Knowledge flow?

Participant: If you want to take it up to another level, the level of: are we achieving our goals and things like that, I think it's getting to be a more inclusive thing. It might be heading from Knowledge towards...

Interviewer: So they are sending something about goal status maybe?

Participant: There's something more than Information there.

Interviewer: So the goal status is Knowledge?

Participant: Yeah, it's a higher thing, and I think the ... I'm not quite sure where you put those. It might be part of Information. It's getting to be part of a bigger picture thing. It's bringing together a number of pieces of Information.

Interviewer: This bringing together, would you say it's Information or Knowledge?

Participant: It's hard to say in the abstract.

Interviewer: Would you say that it's both?

Participant: It can be, yeah. I think part of it's Knowledge, part of it's Information.

Interviewer: What about nonverbal?

Participant: I think that in this situation where, after a while in this meeting, I think although we were trying to be a very happy family, each of us would know if we were on the team of 3 or the team of 1. It

would be some feeling of – I don't have much support here or I do have a bit of support here. Which is probably the reason why the position 101 was a bit passive. He was trying not to be on one side or the other. As position 113, I should have been that way too. As position 113, I was in that role, but in my position 114 role, I was pushing my own...

Interviewer: In that case, who do you think was picking up these non-verbal cues.

Participant: I think everybody picks up the nonverbal cues. We're very used to working with each other. At this stage, we've worked with each other for a long time.

Interviewer: These non-verbal cues, Data, Information, Knowledge, O?

Participant: Mostly Data, I think.

Interviewer: So the position 115, in his progress update is spewing out all kinds of stuff, and you called for a progress update. You say "position 115, give me a progress update." probably in nicer tones.

Participant: I think we all gave progress updates, he wasn't the only one.

Interviewer: Did all of you give the same coordinated... set of flows?

Participant: I think so. So the person I called the position 115, the position 107. The position 108 gave his update, and I think the position 101 probably talked about the things he was particularly interested in. What he was doing was opening up a new idea. He was sort of thinking this other idea doesn't work, let's try something out.

Interviewer: So they all set progress updates to the position 113?

Participant: To the meeting. We're all in the meeting.

Interviewer: So people are sending it to the meeting, not to anyone specific. So we have position 113 to position 115 going "Call for progress" That's something, but we don't know what that thing is yet. Do you have any intuitions?

Participant: I don't know. It's not – I'm thinking it's a lower level thing, really. It's somewhere down – it might just be Data.

Interviewer: The position 115 to the meeting sent? A progress update, as Data, nonverbal cues as Data

Participant: But he also conveyed Knowledge. And Information as well as the Data.

Interviewer: He conveyed his overall performance as Information?

Participant: yes.

Interviewer: Or do you mean as part of these?

Participant: It was conveyed, the update contained all of those aspects of Data, Information, and Knowledge. But I don't mind if you show them by separate flows.

Interviewer: So all of these flows are contained within the progress update.

Participant: Because it's not as well organized as computers talking to each other.

Interviewer: We'll kill the progress update as a flow because we can decompose it as these other flows. So we've got nonverbal cues as Data, We've got overall performance as Information. And all of these we understand as part of the progress. As Information, we have the goal state.

Participant: Now the goal state, now we're drawing it this different way, I think the meeting is actually concluding with this Knowledge, from those feeds. It is producing the Knowledge.

Interviewer: Who does the meeting send this Knowledge to?

Participant: At this stage, the Knowledge stayed within the meeting. But we probably wrote it down at the time. We would have written it on the blackboard in those days.

Interviewer: So the meeting produced Knowledge sent to the blackboard. The Knowledge produced we can label as?

Participant: What went on the blackboard would have contained these conclusions, which are Knowledge. The blackboard also got some other things like to-do lists.

Interviewer: Did the meeting produce these to-do lists?

Participant: Yes.

Interviewer: Is this to-do list, Data, Information, Knowledge, O?

Participant: It's more or less Data, I think.

Interviewer: So, the meeting, sent to the blackboard conclusions and a to-do list. The position 115 sent to the meeting as a whole nonverbal cues and overall performance of his algorithm. Did the position 115 send anything else to the meeting?

Participant: Later on in the meeting, the position 115 was a strong participant in the discussion about how we were going to improve the noun 110.

Interviewer: Improvement discussion?

Participant: Yup. I think it starts with Data. And then you sort of start to integrate things.

Interviewer: The meeting integrates things. What does the meeting do with these integrated things?

Participant: Then it's trying to make decisions. The decisions are finally part of the to-do list.

Interviewer: So improvement discussion as Data goes into the meeting, it goes grindgrindgrind and it outputs the to-do list onto the blackboard. Does the position 115 get anything from the meeting?

Participant: He will get items on the to-do list. He will get a lot of nonverbal communication, etc.. about what goes on.

Interviewer: Is he getting that from the meeting, the position 114, or the supporter?

Participant: These are very hard to separate. I think the way that we were trying to run it, it was from the meeting. The collective.

Interviewer: So the meeting was sending back to the position 115 nonverbal cues?

Participant: Yeah.

Interviewer: These nonverbal cues

Participant: Data

Interviewer: Would you say that the meeting changed the position 115's mind?

Participant: Yes, because at the end of the meeting he agreed that we would change the approach.

Interviewer: Did the meeting send anything to him to change his mind or did someone else send something to him?

Participant: I think that at this particular occasion, the reason I said I was the position 114, once or twice in a lifetime you get moment of clarity. And what we were talking about then was what was called the learning set. So of these X noun 116s or Y noun 116s, however many there were, what had been done in the past was that the expert, our position 101, would say that this set of squiggles is 1a and this set is 3c. And the squiggles were something that, ... we're not necessarily particularly perfect. But then, around that time, we had just got our automatic data logging stuff. And we could see what these squiggles looked like on a screen spread out. And so, as we started to talk through these things, I think that the position I put to everybody was: if the learning set isn't perfect, how can we expect the recognizer to get it right? The basic thing that was said was: let's use our new equipment ..., and we will get things that we believe are absolutely perfect examples for the full period of the window and they will be our new learning set. There was no loss of face in terms of the methodology, the learning set was improved because we now were able to see it for what it was really with our new equipment. I think that really was the difference between the position 107 and those of us who were working a little bit more in the field. Was that the position 107 was thinking of this as much more abstract way, and we were saying, hey, this is ... And so that was really the difference and we said that probably within a couple of days we had assembled a new learning set and we had gone from 65

Interviewer: You as position 114 had a moment of clarity that you transmitted to the meeting.

Participant: Yes.

Interviewer: This moment of clarity, Data, Information, Knowledge?

Participant: It came out as Data.

Interviewer: So you communicated this moment of clarity as Data.

Participant: It's one of those things that people who are a little more extroverted tend to speak before they think. So the stuff that comes out can be pretty rusty. And I think the meeting was able to build on that. I clarified my own position, etc..

Interviewer: Let's look at that. You as position 114 sent a moment of clarity into the meeting as Data. Who built on that?

Participant: I think I got a little bit of encouragement, in the first instance, I built on that.

Interviewer: Where does that flow come from?

Participant: I would imagine that what happened was, that when that idea was first put up, people said oh yeah, there might be something in that. Then we said, the position 108: how could we do this? The idea was accepted almost immediately as being reasonable and then we wanted to know whether it was feasible to do it. And when it was proved to be feasible, people agreed that we should have a go at that.

Interviewer: Then you would say that the position 108, who was the supporter, sent a flow into the meeting of "let's discuss feasibility." This flow, Data,Information,K?

Participant: It's probably just Data, I think.

Interviewer: Then, people went yeah, let's talk about feasibility. The passive position 101 sitting there doing...

Participant: At this point, although he had an alternative thing up his sleeve, he would have then supported this idea because it was a new thing to try and after we talked about the feasibility we didn't think it would be that hard to try it.

Interviewer: So the position 101's support of idea, Data, Information, Knowledge?

Participant: Probably just data.

Interviewer: What other interactions did people have in the meeting?

Participant: I think that this is where, later in the meeting, where the to-do list starts to come out. After the moment of clarity and getting it clearer and clearer, writing up on the board what things look like

Interviewer: your conclusions?

Participant: yeah, I think.... That's right. I think we concluded from the Data gathering stage that we were in trouble, we needed to do something. And then, that's the stage... what can we do? I had my moment of clarity, I threw that into the meeting, we talked about it a lot. I think that the position 107 wasn't sure about it at all. And he was actually pretty concerned that it wouldn't be right. Because in all the books that he read, you get the expert to classify, and we were saying the expert isn't good enough to classify even though he invented this.

Fortunately, the expert agreed himself that maybe that was true. We said "we can classify better if we can see it clearer."

Interviewer: Did you just persuade him with nonverbal cues? Or?

Participant: I think that there was a very logical argument. So that, we're talking about all the people in the room scientific techo types that you're familiar with. People respond to a logical argument. I don't want to estimate non-verbal, because most of us techos most often think nonverbal doesn't even exist. I think, at the time, I certainly would have said that logical argument prevailed.

Interviewer: Who put the logical argument to whom?

Participant: I think that what happened, it's interesting now that we've written it on here. Once I had this moment of clarity, one of only a few in my lifetime, I was advocating that, because I could see it really clearly, and it was only a few minutes afterwards that I saw it as an axiom, that it was so clear that we needed to try this. Then the position 108/supporter, he was jumping on board. And he was saying "Yeah, I think it's a good idea and I know how to help to do this. Here's this." The position 107 wasn't sure because I don't think he was really defending something as much as saying: "I'm not sure, I'm not sure." But I think probably the logical argument was: we actually have the tools to do this now, we're not going to be waiting months to try it, we're can try it tomorrow, and who knows? And that would have been the persuading things.

Interviewer: and you as position 114 persuaded the position 115?

Participant: the meeting persuaded. I think that as I mentioned early on, I actually had pretty good relationships with all of these people 1:1. But there was not such a good relationship here or here, and so maybe I was helping everything to come on board because of the personal relationship side of it.

Interviewer: So we've got logical argument: Data, Information, Knowledge? What is this logical argument thing?

Participant: I don't know how we'd classify theoretical... it is one of those things that you might receive Data, but the Data are all naturally clicking to gear in the receiver's mind.

Interviewer: Let's look at that. You say that you receive Data, and that the receiver's mind "clunk clunk clunk." So... that set of organized, structured data, that really comprises a logical argument. "Given this, this, that, ..."

Participant: As soon as you say organized structured Data, which is exactly what it is, it starts to become Information.

Interviewer: We would say that this logical argument as a collection of ordered, structured, Data is Information?

Participant: Yes, I think so. I think that's where that changed. Because, at my moment of clarity, it was probably better that I threw this into the

meeting as Data. And then we all started clunking, clunking, clunking with it, which meant that this wasn't seen as me – my thing, because we all worked on it. But, as I say, it didn't take me very long, where this in my own mind had gone from an idea to being an axiom. Which, fortunately, proved to be true.

Interviewer: so we've got this logical argument from the meeting as Information to the position 115. But you also mentioned that you had a personal relationship between position 114 and position 115.

Participant: and that was just in terms of, mainly an out of work hobby. So I had quite a good relationship with him.

Interviewer: This personal relationship. Would you say that the drawing on this background, basically going: "I'm going to persuade you, because you know that I'm not completely full of it because of our personal relationship." would you say that persuasive technique is drawing on or communicating Data, Information, or Knowledge?

Participant: It's certainly got an organization to Data, at least. Whether it goes further than that, I don't know how these things are classified. I'm a bit unsure about how to deal with that. It's beyond data.

Interviewer: Is it Information?

Participant: I think he would have used it as Information.

Interviewer: He perceived the flow as Information. Because, he saw it as "here's this guy, I trust this guy because of our shared hobby. He's calling upon that shared trust to persuade me in this meeting. And he would see that as Information.

Participant: I don't think that, for this meeting, that would have been a major factor. Okay? It might have been that straw that eventually... "Yes, we'll do it." I think that what he was – he did tend to sometimes defend himself. I think, in this meeting, he was actually defending what he – knowledge of the technique. And he didn't want the technique to be bastardized in any way.

Interviewer: The meeting produced conclusions, these conclusions were Knowledge. Did anyone else produce Knowledge?

Participant: I imagine that various people quoted Knowledge.

Interviewer: They communicated their Knowledge that wasn't theirs. Who would have done that?

Participant: I certainly think the position 107/position 115 would have done that. Because, to some extent, his reticence to try new things was on the basis of his Knowledge.

Interviewer: So he was communicating his Knowledge to the meeting... And the Knowledge was, status quo?

Participant: I think it was about the state of the art of these techniques and that sort of thing. I think similarly the position 101 would have been from time to time, injecting his Knowledge of noun 117 into the meeting. And the position 108 as well, Knowledge of programming.

He would have been helping the meeting with the Knowledge of programming, because we had to use programming as a tool. These are all kind of constraints to how it is. I might have even been communicating Knowledge of group 130. position 113. So, in a way, although one is communicating Knowledge, you can possibly be communicating it as Data for everybody to churn around with.

....

Interviewer: Does Data become Information or Knowledge?

Participant: Coming into this interview, I thought that the hierarchy was Data, Information, Knowledge. Maybe wisdom or something above that. Or something crazy like that. Way out there. During the meeting, I think I... with your ... nondirectional coaching, you helped me to see where things were probably Data, where those Data were used by Os to form Information. And then, pieces of Information, there's something higher above Information.

Interviewer: So Information becomes Knowledge?

Participant: I think it's can become Knowledge.

Interviewer: Do any of these mediate any of the Os?

Participant: Can you run that by in a different method?

Interviewer: Before I do that. What happens with Knowledge? Does Knowledge jump back, somehow, to either of these? Can Knowledge become Information, can Knowledge become Data?

Participant: I think we were talking about it a minute ago. I actually think that you can go both ways. You can send Knowledge back as Data and you can probably send I back as Data, too. I don't know if that complies with the theories. This is just off the top of my head.

Interviewer: So Data can become Information. How does Data become Information?

Participant: I think Data becomes Information by a person or machine collecting the Data and drawing something out of it.

Interviewer: Can we put a less general generalization to something?

Participant: Somehow, integrating it. Firstly, if you've got say: numbers as being the data. You can tell is it going up, is it going down, is it jumping around? Is it... Those are Information.

Interviewer: Can other things be Data?

Participant: I think we discussed all sorts of things as being Data there. nonverbal stuff. Verbal stuff, not just numbers.

Interviewer: So, when we have verbal stuff, what transformations happen to verbal stuff qua data to something else?

Participant: So, when I take in someone's Data that they're sending, they're being assembled in various ways in my mind. And I'm starting to believe certain things from those assembled group of words.

Interviewer: From that assembled group of words that is Data, you generate Information or Knowledge?

Participant: Yes. I think you can do both.

Interviewer: So Data can precipitate Information or Knowledge.

Participant: I don't know to get Knowledge it has to go through the Information stage or not. I'm not real sure. I think that the analogy I've got in my own mind here is something along the lines of: "I've got a Newton's first law of motion or something in my mind, and I see the Data coming in, assembling it to some extent, and suddenly it fits that model of that, so the Data is almost become Knowledge without going through the Information.

Interviewer: You have a model in your head. Let's use Newton's laws of motion. You would say that those are... Information, Data, Knowledge?

Participant: I think they're Knowledge.

Interviewer: They're Knowledge because you can use them to...

Participant: I can use them in lots of different ways. I can certainly – I can predict the future from Data.

Interviewer: So you can predict the future from Data... So you've got Knowledge, i.e. the equations. And you've got Data which are...

Participant: something happens, something happens, I'll put it in the equation and I'll say, so that's the time using... and I can say "hey, that's going to happen."

Interviewer: Saying "Hey, that's going to happen." which one of those is that?

Participant: I think It might be Information. Because the Knowledge got you the Information in that case. So we're back here, we can go back there too.

Interviewer: So you would say that Knowledge can become Information with the modulation of Data.

Participant: Yes.

Interviewer: Because Knowledge + Data = Information.

Participant: In that case. I think I'd be a flunk in philosophy.

Interviewer: Nonsense. How else can Knowledge become Information? Say you're teaching me about how a Noun 2 operates. What's going on there? Besides me flunking...

Participant: If I was teaching you about how a Noun 2 works, firstly I'd tell you something really simple about what you put in the top and what you put in the bottom and what comes out. And then I'd probably tell you one of the most important things to understand is ... And therefore, how the real-world is constrained so that you can only get certain outcomes. So there are models that I would have in my mind,

and you would say: "OK, I want to make lots more Noun 23, so I'll put I'll put a lot more ... in the Noun 2."

Interviewer: So you can make rust?

Participant: Remember there's a lot of noun 46 in there. ... That's not very intuitive is it? So you were exactly right. And I would say, you put more ... in, you can't, because the top of the Noun 2 will go cold. And I only know that it goes cold because I understand the heat and mass balance of the Noun 2.

Interviewer: So telling me that simple stuff: Stuff goes in the top, stuff goes in the bottom, you would say that you're communicating to me...

Participant: That's more Information. But the model of how the Noun 2 works is Knowledge.

Interviewer: So you start out with Information. But you build upon that foundation with models. But I don't have the ability to understand the models without the Information.

Participant: No, I don't think so. Because you don't even have a language.

Interviewer: The creation of a jargon is the noun 97 of what? You're communicating your what to me?

Participant: In common parlance, you'd say Knowledge. I'm imparting my Knowledge. But a general description is hardly Knowledge.

Interviewer: What are you giving me in that jargon?

Participant: I think it's more or less somewhere between Data and Information. It's just a description.

Interviewer: Because you're giving me Data and Information. And the Data you're giving me is this word means that, or is that Information?

Participant: If I say all the words, that's definitely Data. When I give you a meaning so that you can connect it with something you already know, that could be getting to this stage of Information. But these are all self definitions we've been changing in the last hour.

Interviewer: Let's wrap on the oh so much fun concept of the normative assertion.

Participant: The who?

Interviewer: Saying "Should. You should do that. You should learn how a Noun 2 works." What's that?

Participant: I think I would – it depends on the circumstance. If I just met you and somehow we thought of ourselves as peers, rather than you being an old guru or something or other. So, if you said to me, you should do that, I'd take that as Data. Now, if you were the old guru, and you said to me: "You should do this." I would take it as being more than Data, I'd have to... there's it's not just one of the things that I'm assembling with other things. It's something that I'm almost being compelled to...

Interviewer: would you say the old guru is communicating to the newbie Knowledge?

Participant: By saying you should? In a way, I suppose. You could say that person in giving that piece of advice, has assembled a whole heap of things over a long period of time and that would be Knowledge to them.

Interviewer: So they're drawing on their Knowledge, but you're getting....

Participant: I'm getting Data, initially. And then I'm sort of saying, is that person an old guru that I should take notice of or should I just put this in with a whole heap of other pieces of Data and work out whether I'll do it or not.

Interviewer: And if it's someone you should take notice of it's promoted out of Data into?

Participant: I'm not sure where it's promoted. It's promoted somewhere, because it doesn't have to be digested the same way.

Interviewer: Could you say that is promoted into Information or Knowledge, depending? Can you make a case for either?

Participant: I'm not sure, when somebody says you should, I can't quite get it to be Knowledge in my mind.

Interviewer: It can be promoted into Information, you recognize that they've done operations on a whole heap of Knowledge and Data, and you're getting something that's privileged over other you-shoulds. Whereas as a road sign, it's just Data, right?

Participant: Well, I hardly have ever seen a road sign in my life. But there are some that are mandatory.

Interviewer: That railroad sign there.

Participant: That's mandatory. So I would obey it.

Interviewer: But even in noticing and obeying it, you would consider that it's communicating with its presence Data?

Participant: Yes. I think it's Data. I just wonder when we're talking about this how much research has been done on this. ...

7.4   Interview 4

Participant: A trivial example is we hold a meeting with the noun 54. And we are giving them data on noun 55, they're doing some interpretation and providing it back to us.

Interviewer: Let's model this. Who's us?

Participant: Us would be position 56s, position 57, and position 58ing.
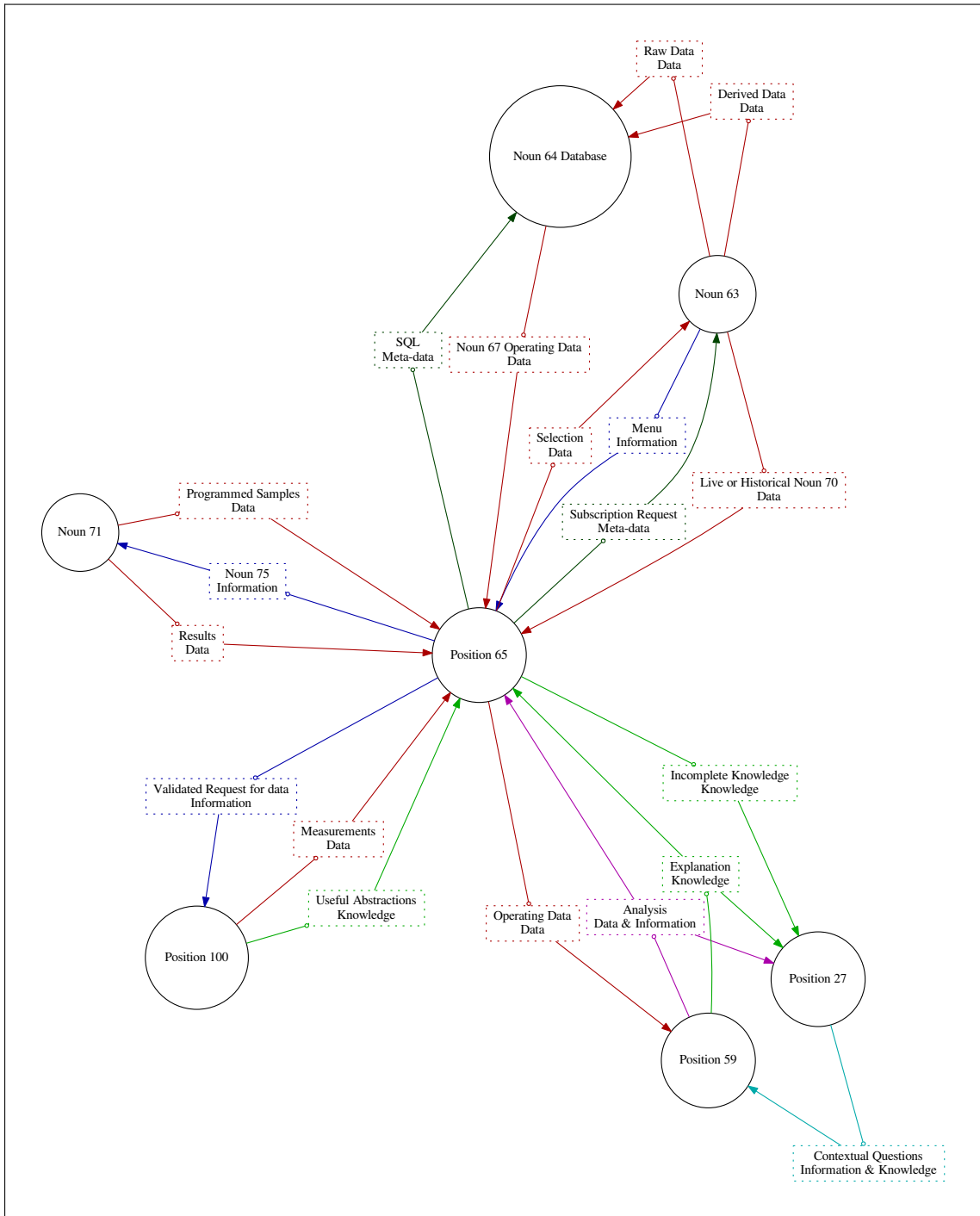
Figure 7.6: The SDFN Diagram for Interview 4

Interviewer: What do I put in the bubble?

Participant: [company].

Interviewer: other bubble?

Participant: position 59

Interviewer: You, send them, ...

Participant: Operating Data. Measurements from the noun 60. usually together with contextual questions. We've got a problem, this is what's happening, why?

Interviewer: Is it a separate flow?

Participant: It depends on what you want to model it as. The two go together...

Interviewer: Entirely up to you

Participant: Leave it together

Interviewer: This flow of operating data is Data, Information, or Knowledge?

Participant: Data.

Interviewer: You as [company] send operating data to the position 59. The position 59 then?

Participant: The position 59 – we need a separate one which is contextual question...

Interviewer: These contextual questions are: Data, Information, Knowledge, O?

Participant: Well, it's going to be both Information and Knowledge, because the way you ask the question, as far as: "I have this Data, I think this is happening (Knowledge), But, my explanation doesn't explain this, this, and this. What's happening? We actually put it like: "[company] thinks this, what do you think"

Interviewer: These contextual questions hold Information and Knowledge. Are I and Knowledge. Is it possible to distinguish a flow of just Information from a flow of just Knowledge?

Participant: They're usually two clauses in the same sentence.

Interviewer: That's fine. Information + Knowledge. The Information in the contextual questions is what?

Participant: The Information? Would be things like "we have found these noun 61s." The Data part was their analysis. But we found the noun 61s, that's Information. The Knowledge is, having done the analysis, they are made up of this and the likely chemistry is this, but that doesn't explain why they formed, tell me why they formed. So that one sentence would have the Information and the Knowledge.

Interviewer: What does the position 59 then do?

Participant: Ignores us. ... If they genuinely don't know, they probably tell you. This comes up in a few different forums. If they know something and it's happened in this project – and if they know something

is potentially a problem for you but you haven't encountered it yet, they won't tell you. When you have the problem if you ask the question the wrong way, they won't tell you. So it can be quite frustrating to do a whole pile of work and go "We found this, this, and this, and our explanation is this" "Yes, we saw this last year on this other noun 60 that's not yours." "Did you think to tell us?" "You didn't ask." ... But the form is those questions: "[company] has a this. We see this, this, and this. What is your experience?" ... What should happen is that they do some analysis on that data, informed hopefully by our Knowledge, and they return to us their analysis and explanation.

Interviewer: Do we want to have separate flows for these?

Participant: Might as well.

Interviewer: Analysis, explanation. Analysis is?

Participant: Something like: model outputs or calculations. Explanation is what that actually means in context of your noun 60.

Interviewer: Class of analysis is Data, Information, Knowledge?

Participant: It's probably more on the Information. Well, it's Data and Information. And the explanation is Knowledge, it had better be.

Interviewer: When you say that, what do you mean?

Participant: You hope that when someone gives you their explanation, you know more than before they told you. Not always true. They can tell you stuff, and you can go "Well I understand even less than when I started." Because if it's completely contradictory to your understanding, you are now really confused. ...

Interviewer: Let's explore that for a second. When their explanation conflicts with your model...

Participant: Here, model might be our model of understanding about the system... In the first instance, you try to go "Well, hang on, well, this this and this. But hang on" What more usually happens and you try to get a lot done in the meeting, so we go away and we talk amongst ourselves and we follow up in an e-mail for clarification. ...

Interviewer: What other entities are here? Where are you getting the operating data from?

Participant: When we said [company], inside [company] the operating data came from the noun 60 data acquisition system.

Interviewer: Are we zooming in?

Participant: We're zooming into [company].

Interviewer: that's a new page. Comes from?

Participant: It comes from the noun 63. ... Level 3 are computers that talk to Noun 10 and do calculations. So when you request data out of a database which is Level 4, it ultimately came from a lower level.. but it might be derived data generated at levels 2 or 3. So we say the noun 63. ...

Interviewer: noun 63... We're going to zoom into noun 63 eventually. I can tell that this is going to be fractal.

Participant: The couple of major things that are hanging off of it. A database. Let's just call it a database. The data repository database.

Interviewer: entity? or internal?

Participant: Strictly speaking, physically it's external, network wise it's logically... just a repository for data afterwards. noun 64 database. I have no idea what noun 64 stands for. Noun 23making something something something.

Interviewer: noun 64 database to noun 63...

Participant: noun 63 to database is one way. We are sending Data for long term storage. Both raw and derived.

Interviewer: raw what?

Participant: Raw Data, would be an actual Noun 4 output.

Interviewer: and we're sending... derived...

Participant: Which is the output of calculations on raw data. Just trying to keep my context here. That's where that Data is coming from. Is actually from that database. We query it.

Interviewer: this is the [company] ...

Participant: In this context. Yeah. When we're trying to work out something, one of the things we'll do is data mining. Generally searching that database. That database is our primary place to start.

Interviewer: You say it's a primary place to start, so we need a you here. So where is this going?

Participant: So this is going into me

Interviewer: can we label me?

Participant: There's me as position 65. Me as position 65 is requesting Information from noun 64. But I can't write to noun 64.

Interviewer: What do you send it?

Participant: I send it an SQL query.

Interviewer: So you're sending SQL. This SQL is Data, Information, Knowledge, O?

Participant: It's other. I can't really classify it as any one of those three.

Interviewer: What can you classify it as that's some sort of category?

...

Participant: So the database contains a model of something. The purpose of a relational database is to model the real world. noun 64 is an position 58ing database so it's a hopeless model. So you're trying to convince that ... in the query, you know what the form of the answer

looks like, and you're trying to describe what you want. I say trying because we're not always successful.

Interviewer: the SQL is a representation of you trying to... what? Model....

Participant: Well, I'm looking for Data in a particular form. And I have what is a model of the real world in the database. And so by imposing some structure on the data that is selected...

Interviewer: so the form of your question is is that you're sending this form of data wanted to the database. What would you classify that concept as? Is it itself, Data, Information, Knowledge?

Participant: It's Information I guess. It's not Data. noun 97 of elimination. Is it Knowledge? It's not Knowledge in in of itself.

Interviewer: But you feel comfortable classifying it as Information because?

Participant: Because that's what's left. That's not the answer you wanted, I love doing this.

Interviewer: We don't have to classify something as Data, Information, or Knowledge.

Participant: That's why I said other originally.

Interviewer: But we do need to put a word in for other that I can add to the pretty colors. As opposed to other which I can't.

Participant: it's metadata. Data about Data.

Interviewer: That we can absolutely describe. So it's metadata. Data about data. The database then sends back, what?

Participant: Data. In the form of nicely formatted tables, hopefully.

Interviewer: Are we going to call the flow nicely formatted tables, or are we going to categorize the content of the flow? We've got SQL going there, so we can balance this with nicely formatted tables.

Participant: We're getting back our Operating Data.

Interviewer: So this flow back is operating data? Is this Operating Data flow the same as that operating data flow?

Participant: It's a subset of that. Because I can get some Operating Data that isn't in that database so I can get it from other places.

Interviewer: Can we differentiate this Operating Data flow Subset from Operating Data with a unique name?

Participant: noun 67 Operating Data. ... I also can query part of the noun 63 directly.

Interviewer: this query is in the form of?

Participant: This query is in the form of me selecting trends in a noun 68. A mouse driven noun 68. It's actually the noun 69the position 90s have for operating the noun 60. We have access to that read only. We

can't actually flip a valve, but we can get the valve's state at much higher frequency that's in the database for a limited period of time. …

…

Interviewer: You're sending – this is interesting. The noun 97 of you sending this to the noun 63 is clicking…

Participant: I'm selecting named variables.

Interviewer: Conceptually, what kind of container does that fall in? A set of selections of named variables is a? What kind of label data flow as? Because we can talk about computer events.

Participant: It's Information isn't it?

Interviewer: I want you to tell me this.

Participant: It provided the menu, I selected the thing from the menu.

Interviewer: Do we want to summarize that or do we want to go into… is it sending you menus and you're sending selections back?

Participant: It sends me a menu, I ask it to send me a menu, it sends me a menu.

Interviewer: Menu. This flow is a flow of?

Participant: Information.

Interviewer: Sending back is?

Participant: Data

Interviewer: Of what?

Participant: of my selection

Interviewer: Are there any other flows?

Participant: Having made my selection, it will then send me the Data that's in that tag. A live updating trend which I can choose to save. What I'm essentially doing is subscribing to a tag, and as new data comes into the system it keeps sending it at me. noun 70,

Interviewer: This live updating noun 70 is ?

Participant: Data. Conveniently.

Interviewer: You just said you send subscription as well.

Participant: When you've selected the menu, you're actually asking it to subscribe. you say "give me a list of things I can subscribe to." It then …

Interviewer: In that case, you're sending it selection, but at higher level you're sending it a subscription request. Is this subscription request as selection Data?

Participant: Isn't it Meta-data again? I don't think about this that often, I just do it.

Interviewer: So, subscription request.

Participant: The interaction diagram for the noun 68, oh god. That's what we're doing isn't it?

Interviewer: Yup. Now what?

Participant: Don't start drawing this bit, at any time within the Data buffer which is about 9 months for that system, I can slide back. Then it's no longer updated data, it's now just historical data, it's just that tag at that time.

Interviewer: So live or historical?

Participant: Yeah, that's good. Then there's other things that I'd be drawing on. noun 60 measurements, the noun 71 people taking noun 37, having chemical analysis done.

Interviewer: noun 60 measurements is another entity? Just noun 60 measurements?

Participant: We can do noun 71 for noun 71. That's one group I rely on.

Interviewer: Now, what kind of relationships – flows are we talking here?

Participant: There's their regular work, which is noun 72 Regular verb 73 that they do – so they have a program of work that they have to do every year. There's a specification of "you will noun 37 at these points for these noun 74."

Interviewer: That noun 72 goes to you?

Participant: It goes to a bunch of people. There are many recipients. There's a heap of people who use it for a bunch of different things. In this purpose, I'm one of the recipients. Let's just call it programmed verb 73. It's on their program

Interviewer: This programmed verb 73 is a flow of?

Participant: Data. I can make a request to them to do something special.

Interviewer: Let's label this request. What are you sending them?

Participant: Assuming I had approval I'd be sending them a program of work. noun 75 work? Anything abnormal is a noun 75.

Interviewer: noun 75 work is a flow of?

Participant: The request to them is Information. I'm asking them – telling them what to do and they're sending back Data again which is the results.

Interviewer: So they're sending you back Data. This is noun 75 results?

Participant: noun 75 results. Then there are noun 75s I carry out as noun 75 position 100.

Interviewer: So you as Data analysis talk to you as noun 75 position 100? Is a different bubble. You as position 65 send to you as noun 75 position 100 what?

Participant: I need Information on something.

Interviewer: You as position 65?

Participant: the position 65 needs to know something so the noun 75 position 100 needs to noun 95 a noun 75 to actually find that out.

Interviewer: Okay, that flow is Information?

Participant: Yes. And that's a request for Data.

Interviewer: And this request for Data is not meta-data because?

Participant: Um, Meta-data is... I'm saying SQL is meta-data. It's because I'm interacting directly with a model of the system as opposed to here where it's like "I need to know about those noun 61s that are over there, I want you to go and get some, get some analysis done on them."

Interviewer: So you send, to the noun 75 position 100 a request for Data which is Information. Then what happens?

Participant: Let's condense everything that goes on around that into "the noun 75 position 100 runs a noun 75" which is – actually collects the data.

Interviewer: The noun 75 position 100 noun 97 returns what?

Participant: In terms of – they're going to return Data and Knowledge.

Interviewer: 2 flows?

Participant: 2 flows, Depending on what they're ask. And then find something they wanted to ask, then they return that too. Depending on what we find, we have to go and return it to a person who didn't ask, who might not be me.

Interviewer: Data?

Participant: If it's a straight request to go and just get a measurement, that Data would be noun 60 Measurements.

Interviewer: The Knowledge is

Participant: The Knowledge is the higher level stuff that happened. So I made these measurements, but while I was there I was watching what was going on and actually though we measured those heights, you can see that they're all part of a flow. That kind of thing. The noun 75 position 100 imparts some useful abstraction about what's going on.

Interviewer: the problem with fractal diagrams is the conservation of inputs and outputs. otherwise I get very confused.

Participant: You realize that if you did this again tomorrow, you'd get a completely different set of diagrams, but anyway... Well, the core of it would be the same. There would be differences.

Interviewer: .. I'm trying not to ask to do this again tomorrow. ...

Participant: Next time you're down, you can do it. I'm actually relatively interested in this stuff.

...

Participant: what's the noun 75 position 100 doing? We're actually expanding the noun 75 position 100 out into their own...

...

Interviewer: That we need to model. Who receives requests for Data in this zoomed in model?

Participant: The request for Data, it would be going to... noun 75 position 100 is me but it's the noun 97 of ... having identified the need for Information separate for how we're going to get it. We have to empanel a meeting of the important stakeholders...

Interviewer: I want conservation of inputs... let's start with this flow. Which entity does that request for Data hit inside noun 75 coordination?

Participant: It does depend a little bit on exactly what was asked. It could hit a number of people. That's why I call it noun 75 position 100, because it's in the center of everything. When one or more of the people interested in this part of the noun 60 decides that the request has come and it's important, decided somehow that it's important.

Interviewer: So this has already been a validated request.

Participant: so we've already decided it's been important

Interviewer: which probably means we need someone to do validation.. who does validation? And is it at this level or this level?

Participant: There's the validation that self does, there's the validation that team that self is part of does, and then there's the validation that the wider people paying for it do.

Interviewer: Do any of those three levels of validation need to be differentiated from the Os?

Participant: Well, if they need to be differentiated, then we need to fractal noun 75 position 100. If they don't need to be differentiated, then we don't need to bother fractalling this guy.

Interviewer: is this of interest?

Participant: Probably not in this context. We're now at the stage where we're modeling how we run noun 75s.

Interviewer: I want more of this level than this happens then that happens.

Participant: There's stuff that the noun 60 meets by itself. There's stuff that people have to go and observe. And we use the noun 71 as an example. And there's stuff that we have to go and make happen. Those are pretty much all the important observables about the noun 60. Important external/internal observables. which this bubble is responsible for collating and trying to put into some coherent explanation of the universe. And that's why when [company] sits in this meeting, there's one person who's holding most of that, but there's a whole bunch of people who are peripherally involved. Like the people who were involved in the original noun 95 and adjective 96 actually hold a heap

of Knowledge. So they come to these meetings, but they're there to facilitate the meeting and provide background context. And go :" And the reason we did that was..." ...

Interviewer: Conservation of inputs and outputs. We need some. Operating Data emerges from [company] emerges from what entity at this meeting?

Participant: Officially it will emerge from one of two people because they've got to be permitted to speak on behalf of the company. Officially it will emerge from either the position 27 or the position 58ing position 41.

Interviewer: do they get bubbles or a bubble?

Participant: They get separate bubbles, I'm afraid

Interviewer: We've got bubbles, these bubbles send a dovetailed flow out of operating data?

Participant: They may coordinate or one may ask the other to do it or they may act independently. It depends. Primarily it will try and go through the position 58ing position 41 because he's a drinking buddy...

Interviewer: That actually matters.

Participant: It really matters, yeah.

Interviewer: How do we get that data flow out of this page?

Participant: How do we get the operating data flow out of this page? Information from the position 65 is being collated by one or both of those entities.

Interviewer: *1 goes to one or both?

Participant: Yep.

Interviewer: This *1 is what?

Participant: This would be research findings, technical reports – it's going to in the context of one of these sorts of we're going to ask a "question" in a few sentences. There's usually lots and lots of reports and studies behind it that are the technical basis for that question.

Interviewer: What do we want to label this flow as.

Participant: Knowledge. I don't know. It's the sum of our Knowledge which we acknowledge to be incomplete otherwise we wouldn't be asking the question. Incomplete Knowledge.

Interviewer: Now this is a flow of Knowledge which is labeled as incomplete Knowledge?

Participant: Yes.

Interviewer: Incomplete Knowledge as Knowledge is flowing from the position 65 to the tech position 41 and position 58ing position 41 at the same time?

Participant: Can, it's context dependent.

Interviewer: In this atemproal acausal diagram, it's going to both? Unless there's a reason to differentiate the flow

Participant: I should make your life easy and say that they're the one person for the purpose of this then.

Interviewer: You don't have to

Participant: When you explain it that way, it doesn't make that much value separating them.

Interviewer: Tech position 41 or position 58ing position 41 – or is there a header for both?

Participant: They sit in different business units. Let's just say, yeah.

Interviewer: One flow from position 65 to Tech or position 58ing position 25 is a flow of Knowledge of incomplete Knowledge. Are there other flows?

Participant: If we're physically sitting in the meeting together. I can directly tell the position 59 something.

Interviewer: Do we want to model that?

Participant: Usually I'm doing that in response to some other query so probably not.

Interviewer: From tech or position 58ing position 25, how do we get to that flow?

Participant: He controls the noun 97 by which we have this conversation. So either that flow has been provided to these guys, the incomplete Knowledge, which I've said, and this is Information and Knowledge. So there's the Knowledge, the Information which is the question part of it is either formulated by him assisted by him...

Interviewer: but it's a function of the incomplete Knowledge.

...

Interviewer: Operating Data is or is not a component of Incomplete Knowledge?

Participant: It is a component.

Interviewer: where does operating data emerge from this fractal zoom?

Participant: So the operating data – the connection to here. Because if those guys could write a query, they'd do it themselves, but they can't.

Interviewer: This is operating data qua data. Now, the Data + Information of the analysis from the position 59 hits whom?

Participant: It hits the position 58ing position 41/Tech position 41

Interviewer: then what happens?

Participant: Sometimes they sit on it. No. It's discussed, circulated, to interested parties. To which position 65 is one.

Interviewer: For purposes of this discussion, are there other interested parties?

Participant: No.

Interviewer: *1 is what?

Participant: This stuff coming back is their analysis which is data and Information.

Interviewer: so the tech position 41 is merely routing.. ?

Participant: He will also have his own interpretation

Interviewer: So therefore there's a...

Participant: a transformation could occur

Interviewer: Therefore this new Data & Information flow ... Is this new flow Data, Information, Knowledge, or O?

Participant: We'd love it to be Knowledge, often it's purely Data or Information. If we hit the jackpot, we get Knowledge. Let's pretend we hit the jackpot.

Interviewer: How are we labeling this flow?

Participant: Let's say it's Knowledge. Let's not model reality, it's too hard. That provides stuff back to you. It may be the answer to the question you asked, it could be the answer to a different question. That was subtly related... that people misunderstanding, yeah it's been doing it for years. Let's hope it's Knowledge. We want it to be Knowledge.

Interviewer: is it Knowledge, or just nonsense?

Participant: the purpose of us asking is to get Knowledge.

Interviewer: So this analysis turns into Knowledge? or is the explanation...

Participant: If they did a good enough job the explanation is Knowledge which we can incorporate into our model of the universe directly.

Interviewer: So you would directly get explanation as position 65?

Participant: We have done. It can happen.

Interviewer: Is there translation that the tech or position 58 position 41 performs?

Participant: He can do. It depends on what he was given.

Interviewer: How can we model this?

Participant: With difficulty.

Interviewer: Really what it means is that these flows dovetail. Because we're modeling all of the causal fiddly bits. So,

Participant: If they're able to tell me the answers to my questions, then you know, ??? is important. Stuff comes back and we deal with it

Interviewer: At some points of time you get Knowledge explanation back. You also directly get analysis back?

Participant: Yes.

Interviewer: other points in time, explanation goes through the tech position 41 – tech position 41 gets explanation?

Participant: Explanation and analysis, it's the same things.

Interviewer: Sometimes the position 58ing or Tech position 41 takes the analysis and or explanation goes "wow this is shit" and does what with it?

Participant: Okay, in the context of our noun 60 what that really means was...

Interviewer: Alright, so this is interpreted explanation? Who gets that interpreted explanation?

Participant: The position 65.

Interviewer: and that is Knowledge?

Participant: Yeah. It's Knowledge... and Data and Information. It's all three again. It depends on what it was.

Interviewer: So this flow of Data, Information, and Knowledge ... and or or?

Participant: It's and.

Interviewer: can be labeled as?

Participant: The internal? Feedback. "What feedback did you get from those guys? Nothing!"

Interviewer: Are we missing anything?

Participant: For purpose of this, we've got conservation of Knowledge.

Interviewer: Let's start with a discussion of the diagram you've looked over and that we were creating last time. We had in that diagram the noun 71, Database, and you as noun 75 position 100. And we did not iterate into any of those because of time constraints. Are there any aspects of that diagram that bear iteration?

Participant: Unless we want to go and start modeling the actual noun 75 planning noun 97 and things, I don't think so.

Interviewer: What I'm looking for is a difference in nature or kind. – Last interview focused on a more technical point of view. You're talking about interactions with databases and meta-data. Can we look at a scenario from your experience where there's been confusion over understanding of data, Information, or Knowledge? In a procedural sense, not just in a one-off, but in some sort of really nasty conflict of philosophies.

Participant: If I've got the right idea, I'm thinking of an instance where we have some noun 60 measurements of something, and they should be

– we're trying to use them to understand the physical conditions that are in the noun 60 and what actions we need to take, but there's more than one way of analyzing those. And there were a number of opinions and this evolved over a couple of years as to gaining consensus. And of last week we still don't have consensus.

Interviewer: That sounds like just the thing. Where do we start?

Participant: This is the noun 76 again. It has some noun 74points.

Interviewer: How do we want to model those?

Participant: There's 4 points and it's the differences between those noun 74s that we're trying to understand.

Interviewer: We'll start with your interpretation, and then I'll have you try to explain other peoples' interpretations. With your interpretation, what entities should there be?

Participant: There would be 6 in total ... You don't care what they are, you just care that they're there. ... The actual noun 74s that we measure at these points is a function only of the adjective 78 and the adjective 79. And the adjective 79 responds to the adjective 78. If we measure the noun 74s, we should be able to infer what's going on inside the box. We'd like this box to have all the same properties everywhere. However, since once of the purposes is a noun 80: if there's new noun 80 material being introduced ... and old noun 80 material ..., it's not going to be uniform .... We'd like to know about capacity and flows and stuff like that. So, if I look at this point and it's higher noun 74 than this point what does that imply about the system?

Interviewer: This point being an noun 81 versus a noun 82?

Participant: So I'm interested in that case between that differential noun 74.

Interviewer: In that case in terms of the data, we have noun 81 and noun 82...

Participant: My interpretation of this is that if we – this is time and this is noun 74 <drawing> – if everything was uniform then those two would be exactly the same. If over time this happens and it's the noun 81, we can have higher noun 74 at the top and lower noun 74 at the bottom. There are two ways to interpret this. One is, in my opinion, wrong. [explanation of model] And we justify this decision from numerical modeling. We can actually then independently create a numerical model which is consistent with that explanation. The alternate, incorrect view, is [alternate explanation] ... will give you the right answer most of the time. It's a nice rule of thumb. noun 83s just happen to be different. And the difficult thing for us was, until we constructed a numerical model and said look: this numerical model starts off with the ... equations about the universe and it's entirely self consistent and look it's the same result. That was something which helped us convince people that well, yes, that is consistent with the noun 60 measurements. And my explanation is not consistent with the laws of the universe. "OK,

you're probably right" say the Os. Where it goes wrong if, other things can also influence those noun 74s. Say if there's noun 61. ... And that's when I realized a lot of people were never really on board. Because you go to a meeting and they go, it's just like you're reading ??? trails. I don't necessarily believe your explanation because other things could be effecting it. You know how in any defined problem there is: these are the assumptions I've made, this is the framework in which I use to analyze the problem. If my assumptions are violated, then my explanation is no longer necessarily the only explanation. You go back a little to change the assumption: well we've got a special cause. A special cause is driving this. Naah – this is where we have some maths which are the noun 74 differences, we have Knowledge which is our – other people have studied the universe and we've implemented models based on their studies. The ... models which are based on the [name] equations. These are more well established. I'd class that as actual Knowledge. If I class that as Knowledge, I'm using that Knowledge to view the Data that I get from the noun 60. There are some assumptions embodied in that, and I have to be aware of those. If people reject that Knowledge or disagree with the assumptions then it becomes really hard to agree on what's the interpretation of what's actually happening.

Interviewer: This is a really useful discussion. DFDs are cute, but this is going to be more interesting. So we've got, Data, Information, and Knowledge. They're in a line or however you want it. Walk me through your mental noun 97 of coming up with some sort of explanation for this. The diversion where there's an argument. We'll be drawing lines in and out of these things, labeling lines.

Participant: Initially, we all started ... with the same Data. We are making these noun 60 measurements.

Interviewer: Data here are what? Data is or are? I'm not actually being pedantic here. I've found instances of Data as singular, so I'm interested in how you use Data.

Participant: Completely inconsistently. In this case, we wouldn't refer to it as Data. We may call them Data points.

Interviewer: So you start by getting Data points.

Participant: Originally the noun 63 didn't even measure it. It was originally a dial, and a guy had to go and read it once in the morning and once in the afternoon. Because it was a retrofit in the system. And later we connected it to the noun 63. Those are noun 74s I measured at known locations. At known locations and times. Initially we all agreed what that meant. Later on we began to fight over whether or not the Noun 4 is a good Noun 4.

Interviewer: So you're measuring noun 74 at a location and time. Is there a consistent definition of noun 74, and where does it come from? An operationalization, say. It's not a theoretical definition. When we use noun 74, we are referring to this.

...

144

Interviewer: So we're measuring gauge noun 74. The concept of Gauge noun 74 is what here, if anything?

Participant: That would be Knowledge.

Interviewer: So really, we have a Knowledge start of what? I'm trying to get at the really basic assumptions.

Participant: At the level of your standard position 58 they have to take a great many things in their education on face value. And it's an assertion.

Interviewer: You would assert that these assertions are Data, Information, Knowledge?

Participant: Knowledge.

Interviewer: So these assertions, as Knowledge, form what?

Participant: They form a context in which to view...

Interviewer: Context. Starting here from Knowledge, it produces a context for your Data?

Participant: Yeah.

Interviewer: Is there an incoming thing for Knowledge here? We have Knowledge which produces context for data. Is there anything which informs which context we produce?

Participant: We'll take the view people have: it is a modern western educational system. Background education, everyone should have that.

Interviewer: Background education produces Knowledge, Knowledge is or produces context?

Participant: It produces context.

Interviewer: What is that production – where are we getting context from in terms of production Knowledge? We've got background education and we somehow get context from that.

Participant: So what else produces the context?

Interviewer: Yes. What's the transition phase there?

Participant: That context together with what I learn about the noun 60 has to be applied to this particular problem.

Interviewer: So problem localization?

Participant: Yep.

Interviewer: What is problem localization?

Participant: That's your initial study/investigation of the issue at hand.

Interviewer: This study of issue at hand is which category?

Participant: That's collecting Information about it. The study of the problem at hand produces Information.

Interviewer: The Information, combined with background Knowledge, produces context?

Participant: yeah.

Interviewer: So we have, as the two inputs to this problem, we have background education and the study of the problem at hand. Are there any other inputs to this problem?

Participant: No.

Interviewer: now the study of the background at hand produces Information. Is there a transmutation of the study, or is this line the study of the problem at hand?

Participant: Not sure.

Interviewer: Is the study of the problem at hand Information or is compiled from stuff into Information?

Participant: It's probably compiled from stuff.

Interviewer: So the study is also Information itself?

Participant: It generates Information.

Interviewer: so we have the study of the problem at hand which generates Information. This Information generated is what?

Participant: In our worldview, the difference between Information and Knowledge, is that Knowledge implies some level of understanding. I go off and I get the noun 60 drawings and operational noun 116s, and I review all of that and I distill from that what is the important Knowledge which I hold in my head.

Interviewer: Let's talk about that. You're distilling Knowledge which you hold in your head.

Participant: I'm assembling a Knowledge framework in my head from Information.

Interviewer: What is the purpose of this framework?

Participant: To help me understand Data that I receive in the future.

Interviewer: So what you're telling me is that from the study of the problem at hand and from background education you're creating a contextual framework?

Participant: Yeah.

Interviewer: does the contextual framework do anything other than provide context?

Participant: Not really.

Interviewer: So this study of the problem at hand as Information plus background education as knowledge combine to create a contextual framework. This contextual framework is Knowledge and it is used as a basis for telling you what Data to collect? or for you to understand the collected data?

Participant: Both. there is lots of Data that I could be looking at, and my context is my filter, because a lot of it is not important.

Interviewer: Can you think of a useful nature analogy for this?

Participant: A molecular sieve?

Interviewer: A molecular sieve, fantastic.

Participant: It's just an ordinary filter.

Interviewer: So the Knowledge acts as a filter? And the stuff going through the filter is Data?

Participant: Yeah.

Interviewer: When you say the stuff going through the filter is Data, give me an example of something that's trapped by the filter and something that's passed by the filter.

Participant: The vibration amplitude of the fans that are feeding the system. It is part of the noun 60 and it is peripherally related to this problem, but it's not important enough to actually pass through the filter. Whereas the actual flow through that fan which is also an input of the system, this far more important input, will pass through the filter, and I will try to analyze it in light of ???.

Interviewer: In order for you to analyze something that's passed through the filter, is there some function of observation going on or has the observation happened upstream of the filter?

Participant: The observation happened upstream.

Interviewer: Context filter as Knowledge?

Participant: Yeah. I'd love for you to have an interview with [name]. You could have many interesting discussions about what is context and whether or not I've actually observed that or not.

Interviewer: So we have observations here. Observations of what?

Participant: In our world, they'd have to be observations of physical states.

Interviewer: Observations of physical things? Go into this noun 97? An observation is what?

Participant: An observation is a measurement of some kind.

Interviewer: So a measurement is generated and flows into the filter. The filter then winnows out irrelevant observations... and then we have relevant observations.

Participant: This is the important thing. Because people hold different contexts for these reasons. That includes what may or may not be superfluous around them. And chuck out the stuff that really matters.

Interviewer: So one area of conflict is the context filter. And that context filter conflict can come from either different priorities, a different background education, or a different study?

Participant: Yes.

Interviewer: Are there any other sources which perturb the context filter?

Participant: People's political agenda.

Interviewer: Where should I put the box?

Participant: I guess it's Knowledge.

Interviewer: What kind of Knowledge? Why is it Knowledge?

Participant: It exists by itself. It's not Information which you can tie down. A person may be... in all cases of position 58ing problem, we can't measure the things we actually want to measure. I want to know what the adjective 78 is. I want to know what the adjective 79 is there. Guess what, I can't measure either of those two things. Alright, what can I measure? I want to measure the noun 74. That's going to piss off that guy, because he's got to put a hole in his noun 60 and it's going to corrode at that point. His political agenda is to stop me. He wants to solve the problem. He works for the same company as me and we want to make money. But he doesn't want to put holes in his noun 60. So he's got conflicting things: help make money, make his life easier.

Interviewer: So the political agenda is basically an estimation of personal needs, and this is Knowledge?

Participant: For that actor, yes.

Interviewer: Now you said it can't be tied down like Information can?

Participant: Yes, because it's fluid.

Interviewer: Now background education is also fluid?

Participant: Well, people will learn more. But they could be misapplying some of their education. They could review that...

Interviewer: It is also fluid?

Participant: to an extent, hopefully less.

Interviewer: Whereas Information's not fluid?

Participant: Information should be more concrete.

Interviewer: When you say more concrete, what do you mean?

Participant: First off, Information is by more concrete – in the hierarchy. Data is concrete. It might be – I measured it via a defined noun 97 at a particular time. We might argue what it means, which is Knowledge, but it is what it is. But the Information... I've reviewed a whole pile of things. Those are Information sources, they are – it's clear how I obtained them. They may be wrong, their provenance may not be... It's less concrete.

Interviewer: Because the provenance may not be clear

Participant: It may not be clear.

Interviewer: is the nature different? Or is just a difference in reliability?

Participant: Between Data and Information? There's always some overlap, so it's not just provenance. ??? Data would be a nice discrete thing. That I could tie down. The Information ??? could be viewed as a collection of data.

Interviewer: So Data is something that could be tied down. Information is a collection of data?

Participant: It can be.

Interviewer: when you say tied down, what do you mean?

Participant: I can specify what, where, when. Probably what where when would define a piece of that. What the measure of it was, its location at the time.

Interviewer: When you say tied down, its tied down in terms of provenance? reliability?

Participant: Yep.

Interviewer: Information is a collection of Data. The noun 97 of collecting it is a function of something we'll get to. But the collecting it reduces its reliability?

Participant: It doesn't reduce its reliability.

Interviewer: what does it change?

Participant: It places its – it begins to place it in a context of other things. So you'll often have conflicting data. They're fine by themselves, but when viewed together in my collection of Information, now ??? Knowledge is going hang on,

Interviewer: Now let's go up to here. We've got observation through Knowledge filter giving us relevant observations according to someone's study, background education and political agenda. Relevant observations then what happens?

Participant: Then we are with that subset of relevant observations we try to produce a consistent context, a consistent view of the world. Yes I know I used context in the filter, but it...

Interviewer: are we using context in a different..?

Participant: This is a different context. Probably there's like – I have to decide what to look at.

Interviewer: In this case, context as what?

Participant: This is context as relevance.

Interviewer: when we're using one word in different ways, we can just do "as foo". So we've got context as relevance filter, winnowing out relevant observations. We... combine?

Participant: Combine those observations

Interviewer: We combine these observations into what? More Data? Information?

Participant: Into a Data set.

Interviewer: We combine Data points into a Data set. This Data set is Data, Information, or Knowledge?

Participant: It's still Data.

Interviewer: So we've got a data set of relevant observations? The fact that we have a set here. What produces the structure of the set? Why do we say that it's a set?

Participant: We can say it's a set, because we constructed our filter... It's a set of things related to the problem at hand.

Interviewer: So their only relationship so far is their relationship to problem...?

Participant: If this is the first time we're doing it, yeah. But later on it should be better than that.

Interviewer: But the Data points don't have any relationship amongst themselves?

Participant: They will, but I can't do this – it's hard to do abstractly. There will be subsets with multiple interrelationships.

Interviewer: So we've got sets as relevancy, but we also have sets as relationships. Set relationships are considered to be what? Data, Information, or Knowledge?

Participant: Information. Set relationship – if we know a priori yes, noun 74 is related to Noun 12, that's Information. ??? But it may have to be something I have to discover. It depends where I – where that relationship comes from. Sometimes I can just know where it comes from or I have to discover it through further manipulation.

Interviewer: That's relevant. If we know a priori, we get that from background education? which is then turned into Information? What is that transformation right there? We've got lots of background education, from that we extract Information?

Participant: Uh-huh

Interviewer: The Information is that there is a relationship here? So, we have relevant observations and data sets. Some data sets have relational context internal to the set. Os have context external too the problem. Well, all have context external to the problem. Some have internal context based on prior Knowledge turned into Information. Given that we have some that don't have an internal context that we have to discover, we could classify this as the position 58ing problem? Walk me through this.

Participant: In a concrete example?

Interviewer: Yeah.

Participant: This same problem? Ultimately, I'm interested in the adjective 78 and distribution of that noun 60. There are a bunch of things which I immediately know are directly relevant. adjective 79 rate, the

Noun 12. ... I can do all of that – any position 58 should be able to do that. There is something about the nature of the noun 84 that's probably important. That's probably going to have to do with: what noun 97es that produce the noun 84? What noun 97es that those raw materials were... It's not – I may suspect that they are related. I may not. There was an argument about that. I have to discover that relationship. In a nice case... the number of times when people don't have any idea about an position 58ing problem. They go "right. Here's the thing I'm trying to control. We currently have the noun 85 table, there are 500 columns in the noun 85 table, let's join that with that table and let's do bivariant plots to every one of them and get the r-squared. And we go through – I'm not kidding. Ask [name] about Noun 23 and noun 28. This is one of the only – if the only tool you have is a hammer, everything looks like a thumb. It's just like that. Can I do anything more sophisticated? Hopefully I can winnow out some of those 500 odd columns, because I know something. I have to be careful, because often – the thing that's driving this might be related to this. We don't measure that. We measure some other things that are related to that. Hopefully I can get a weak relationship between them and the thing of interest. And maybe I'll discover this thing one day. That's sort of the noun 97 where we're trying to work out what drives this. And there are quite a few things we have that where we'll find a subset of 5 things that can produce a decent linear model but allows us to describe, predict, what that's going to do. And then time goes by and they don't work so well anymore which tends to suggest that it was this thing down here that they were related to. And some other things that were also important are changing and we don't really know what's going on. That's actually the key we're trying to address: what research is trying to do and what the – technology department solve today's problems, solve next month's problems. Research department, solve /the/ problem. Not the same thing. That's the implication that we're actually getting to the – we're aiming towards laws of nature. What is actually the real reason? As opposed to "How can I get a useful working model that allows us to make more money?" That's more the technology position 58's job. Which means we're going to have overlap and conflict because we have different agendas.

Interviewer: You mentioned prediction.

Participant: Why do I want to predict?

Interviewer: a) why, but b) how do you predict?

Participant: If I had a causal model of the factors that control a variable that we're going to observe, the reason I want to know the noun 74s is that it's telling me about adjective 78. I want to control the adjective 78. if I know that these center 5 things have positive and negative influences on that adjective 78, and I want adjective 78 to be low, I will try to select those 5 things such that I make it tend to be low. That's why I want to be able to predict. So that when they say: "We're thinking of buying this new thing.." and I'm saying "don't do that because it's high

in this and this will make that bad and we'll make less money." That's why I want to predict.

Interviewer: A predictive model is what? Knowledge, I, Data?

Participant: It's the embodiment of our Knowledge.

Interviewer: This embodiment of Knowledge is what?

Participant: I can't be more concrete than that, it's the embodiment of Knowledge. It should be an algorithm.

Interviewer: Does an algorithm even feature on this spectrum?

Participant: An algorithm should be Knowledge.

Interviewer: Why?

Participant: The noun 97 of understanding is the noun 129 of Knowledge. Because when I understand the problem, and I can specify: "you take these things and you do these operations on them and the outcome of that will tell you something." basic way of looking at an algorithm. Will tell you what's going to happen in this other thing. The Knowledge is here, but I have to express it someway. An algorithm is a way for me to document the Knowledge I've developed.

Interviewer: An algorithm is an expression of Knowledge which is itself Knowledge.

Participant: I'm running out of words.

Interviewer: Feel free to make new ones. How does Information ... inform your ... It looks like we've got a couple types of Knowledge here. We've context and we've got predictive. Are they different?

Participant: In the view of the algorithm, when you're providing context to the algorithm, you're just providing the inputs into the algorithm. Which produces a prediction.

Interviewer: Do we have a type of Knowledge that's not algorithmic?

Participant: Yeah, stuff that I just know.

Interviewer: what is the relationship between stuff that you just know and algorithmic Knowledge?

Participant: I tried to say that algorithmic Knowledge... having developed a Knowledge framework about something, how do I write down, tell someone else, tell a computer about it.

Interviewer: Good point. What's computer code?

Participant: In a lisp sense?

Interviewer: Let's work from the trivial sense, back.

Participant: In the trivial sense, it's a series of instructions that you're asking the computer to perform.

Interviewer: so each of these instructions, does any instruction fall under this chart?

Participant: An instruction tells the computer to operate on Data in the computer.

Interviewer: So it's data operation. Is the instruction itself Data?

Participant: Well it certainly can be, depending on what your code is doing? Self-modifying code.

Interviewer: So code can modify itself, this modified code is Data?

Participant: Instructions.

Interviewer: This modified code is instructions, is Data?

Participant: Are we talking Von Neumann architecture, are we talking LISP? Or are we going to say a procedural language on a Von Neumann arch? I'm being a pain I know.

Interviewer: First question is: do we need to differentiate architectures?

Participant: Well, if we're going to talk about something that's a purely functional system. Where's the state of the system? Well, it's in the current state of the functions. They're all and the data is immutable, and it's a different way of thinking about the problem. If we leave it in a nice procedural world where Data resides in either memory or in registers and I carry out operations by moving Data from place to place or combining two pieces to produce another piece of Data. Then I have instructions and I have Data. In a functional architecture – let's say LISP. Let's not talk FORTH. I've had limited experience with stack based languages.

Interviewer: The real question is: when does stuff in a computer become not-Data? Where does computer based stuff escape the classification of Data?

Participant: Well, that's a meaning we ascribe to it. At the end of the day everything is a stream of bits. So this bit stream is Data. It depends on the level of abstraction you've chosen.

Interviewer: At what level of abstraction does that change? Is there a point where we get more and more abstract and suddenly those bit streams cease become data.

Participant: At a certain point, I'm able to attach meaning to it. When I'm looking at a character on the screen, from the screen's point of view, some bits came along and it does stuff. And you have an A. Great. I look at those bits whether they're on or off and say that's an "a" At that point, I've got some Information.

Interviewer: So your perception of a character on a screen is not Data, but Information

Participant: It can be. If I was looking at DNA sequences, that "a" is just one of the base pairs. And that would be a representation of Data.

Interviewer: This representation of Data is itself Data, Information, Knowledge?

Participant: It's just a representation.

Interviewer: is the representation significant?

Participant: How something's represented is the significant act. Data is represented by something which is significant.

Interviewer: So you as position 90 looking at computer screen see Information, but not Data, because you're looking at representations of Data?

Participant: I will call it – it will depend on how people decide to look at it. if they're just looking at a number in a text box, And that's changing with time or they're looking at a trend line – it's the same underlying measurements of the noun 60. They'll react to them differently.

Interviewer: Is how they react a function of something?

Participant: It will depend on all the things that are personal drivers on them. If it was something they were trying to control and they're just looking at the current number, and they go yes, that's less than my current set point limit. I'm happy. They're living in the now. If they have a trend plot of it and they see that it's under the set point limit but increasing, I'm worried. As opposed to previously: "Oh, fine. Is it going up or down, I don't know, it depends on my context."

Interviewer: How does Data become Information?

Participant: When my theories?? get classified into data sets, collections...

Interviewer: You have the Data sets, and the Data sets with relationships are Information?

Participant: Yeah.

Interviewer: These Data sets that become Information, what happens to them?

Participant: In an ideal world, I'd write the perfect technical report which would form part of the institutions Knowledge and everybody would read it and we'd all have pushed back the frontiers of Knowledge a small amount.

Interviewer: Let's unpack that. Candide's universe, we've got el dorado. You have Information this information is your Data set.

Participant: Yeah

Interviewer: You then write a technical report. Is this technical report Information?

Participant: It probably is Information.

Interviewer: Does it represent Information?

Participant: It should represent Knowledge. It depends. In an ideal world, you've carried out a study, you draw some conclusions about the universe based on that study. You're trying to impart that Knowledge to someone else along with the supporting Data and Information. We

would say Data but it's really Information, because measurements were taken and we use them in some way and that is Information from which we derive some Knowledge which hopefully is self-evident and we're trying to impart that to people.

Interviewer: This Knowledge is a change in prediction modes, is a change in context, both?

Participant: Both.

Interviewer: So Information can cause changes in Knowledge, or is a change in Knowledge?

Participant: It can cause changes in Knowledge. It could always be rejected or it could be wrong.

Interviewer: We've got Knowledge filters Data, produces Information, produces Knowledge. Are there any other metarelationships here?

Participant: I think we've had enough

Interviewer: Nonsense, I'm only mostly confused.

Participant: ???

Interviewer: What kind of Knowledge does Information cause a change in?

Participant: I'm taking change to include simply: we've increased our Knowledge. So, all Information increases my contextual Knowledge. ??? If it's more Information about something I already had Knowledge about it will either reinforce or cause me to do stuff: well, I've got 2 pieces of Information in conflict, it means the Knowledge worker is trying to resolve that conflict so that this bit might get weaker in this area.

Interviewer: When people you communicate with disagree with you, do they diverge from this?

Participant: Sometimes yes.

Interviewer: Lets look at both cases. When they diverge from this, where is their point of departure?

Participant: The annoying case is: you're wrong, I don't have any evidence but I know you're wrong.

Interviewer: What does that annoying case stem from?

Participant: Some past politics. A lot of problems just aren't really clear cut. Because you have a limited understanding, a limited data set about a thing, you can't always... well I've developed a model about its behavior, and I've got maybe 20

Interviewer: Is that a problem with contextual filtering?

Participant: Yeah. their contextual filters are different than mine.

Interviewer: Stemming from all these three sources . That's the annoying case, what are the other cases?

Participant: Where lets say I didn't have the right contextual filter and I say I've gone through and I've done this and here's my case, this is great... "Yeah, you didn't consider what happened in 2002" yeah, I didn't put that in the Data set, I didn't think it was important. "Yeah, it is." I look and go: "Actually it is important." So my contextual filter was incomplete. So they provided me with new Knowledge or Knowledge that was new to me, which I then have to adjust my contextual filter, and in light of that, the Knowledge coming from my study of my data points may not be correct.

Interviewer: Where do "shoulds" fall here? Either incoming or outgoing?

Participant: Should should be an outcome.

Interviewer: Specifically, if you're producing normative assertions, what are they here?

Participant: For me, I would say that they're Knowledge. For the recipient they're probably Information.

Interviewer: When you are the recipient of normative recipient of normative assertions, you consider them to be?

Participant: I consider them to be Information. I don't trust other people. When someone says you should do blank, you should go and measure this. Well, OK, but I have to decide myself if I'm going to actually believe you.

Interviewer: So these are untrusted assertions?

Participant: Until I do some vetting, everyone's assertions are untrusted.

Interviewer: And an untrusted assertion is Information? Is it different from the Information from the study of problem at hand?

Participant: It can be Information that's coming from the outside, yeah. While studying there will be other stakeholders who will tell you stuff. If they're powerful, you have to at least pay lip service to them.

Interviewer: How do you vet these pieces of Information?

Participant: Who do you trust? If it's somebody who I respect technically, then I will take their advice. If it's someone who I think is a <bad person> then, unless what they've said is obviously true by inspection, then it gets far more weighting down. If the person is really powerful it's no longer about trust. If someone really powerful in an open forum tells me to do something, I ignore that at my peril. It's about weighing up the Information that's come to me. Do I have to act on that? That will depend on the original relationships?

Interviewer: Where does meta-data fall here?

Participant: Sometimes it's Information. You didn't want that answer. The Data about Data.

Interviewer: In this flow, what is the meta-data? Is there any? Is it flowing past?

Participant: It's part of the filter, yeah. To get the specific example of those noun 74 tapings, they measure noun 74. What noun 74 do they measure? The meta-data is how the measurement is actually taken.

Interviewer: and so that meta-data can be a ??? is Data itself, is Information, is Knowledge?

Participant: It's Information in the main. I make the distinction in that case so when we're discussing the meta-data of those noun 74 tapings, there's a pipe with X cells and has a hole in each cell. But the pipe... we have one Noun 4 at the end. I've told with the other dimensions: you know how many holes there are, you know where they are, you know what the separation is. That's all informing you about the measurement I'm making at the end of the pipe which is my original piece of Data. So I've given you a bunch of Data about the meta-data and you have some Information which is maybe "this isn't reliable" or maybe its something measuring this thing or the other thing.

Interviewer: Is Data atomic?

Participant: In this example, we pretend that Data is. Because in every case leaving out quantum mechanics, these are continuous variables. But we noun 37 them at discrete points in time and space. So we treat them in all analysis pretty much as discrete observations.

Interviewer: So there are no sub components of observations? You can't have half an observation?

Participant: That's true. We would like to think... it's a thing.

Interviewer: Is there anything past Knowledge?

Participant: An overall Knowledge framework?

Interviewer: Is this overall Knowledge framework Knowledge?

Participant: I consider the whole thing is Knowledge. We're just growing. I think Knowledge is just the boundary.

Interviewer: Is Information divisible?

Participant: Yes. A piece of Information, if I take the noun 74 difference between two things is foo, that's a piece of Information about the system, versus the difference between two things is immediately decomposable to two things.

Interviewer: Is there cases where it's atomic, or is it always divisible into Data?

Participant: It could be atomic. It depends on what it's Information about.

Interviewer: Give me a case where we have a piece of atomic Information.

Participant: The diameter of the pipe on the drawing.

Interviewer: The diameter of the pipe on the drawing is Information not Data, because...

Participant: If it was a measurement of the actual pipe it would be Data. But this is Information. It may not be correct. This is a drawing, it's not the actual thing. It's a thing about the thing. It's not meta-data.

Interviewer: It's a thing about the thing that's not meta-data. Is Knowledge atomic?

Participant: If I could only point to something and say... we do do that and point to something and say "That's Knowledge." except that that's not – it depends on the context we said that in.

Interviewer: Can we point at something and go that's Knowledge?

Participant: Not realistically.

Interviewer: When we point at something and go that's Knowledge, what are we pointing at?

Participant: In the abstract sense?

Interviewer: What category of thing is it that we are pointing at? I point at this and go this is Knowledge. When I point at this and go this is Knowledge. Is this Knowledge? Is this Information? is this Data?

Participant: The physical thing that you're pointing at is the physical representation of the Knowledge that hopefully both you and I hold in our heads.

Interviewer: So it's a Knowledge representation that's not necessarily Knowledge, Information, nor Data? Representations of these categories do not necessarily require that the representation itself be in that category.

Participant: I agree.

Interviewer: Is there a category for representations?

Participant: Yeah, meta-data.

Interviewer: When you try to communicate these representations to people and an error occurs, where are those errors, what are those errors, and why?

Participant: The errors can be my communication of that. Faulty analogy, assumptions about recipient, that kind of thing. There's also reception errors where the person doesn't listen or is missing something that would help them that they require to receive that Knowledge.

Interviewer: Errors can apply on either side. The noun 97 of communication is what? Errors occurs on the sending side of misreading of what?

Participant: Faulty assumptions, deficiencies in transmission.

Interviewer: So these transmissions are all representations of these three categories?

Participant: If we're talking about transferring Knowledge to someone else, it's all a representation.

Interviewer: In transferring Knowledge to someone else, the sender can have errors of assumptions or of representation. Can you transfer Information to someone else?

Participant: Yes, that's what we defined in those examples.

Interviewer: Sending Errors there are what?

Participant: Straight out mistakes. Transcription errors.

Interviewer: Sending Errors in Data?

Participant: Transmission errors or corruption

Interviewer: Reception errors of Data are?

Participant: They just didn't listen. That would apply to all categories.

Interviewer: Intentional deafness? What other errors are there receiving?

Participant: understanding errors. They got all that was transmitted but their own contextual filter rejected it or parts of it.

Interviewer: Contextual filter errors.

7.5   Interview 5

Interviewer: What would you say – can you differentiate into projects? Or has it just been one stream of continuous research.

Participant: The parts of the noun 118 can be broken up to some degree. So a lot of time has been spent looking at how to put the materiel into top of the noun 118. How you feed a noun 118. ... And the bottom of the noun 118, the hearth, ... there's a lot of work that happens on that. Because that is the blackest of black boxes for the noun 118. What happens there. Stuff goes in, stuff goes out, and you can measure Noun 12s around the side. And that's it. And you cannot probe it. You just see it once every 20 years.

Interviewer: Because that's when you flush it out?

Participant: That's right. That's when you do a reline. And even there, the mere act of cooling it off completely changes what you can look at. It's vaguely representative of what was happening.

Interviewer: So the ... experiments ...

Participant: when they verb 119ed a noun 118 and cut it in half basically.

Interviewer: Doesn't really model...

Participant: Even that. You'll get Information out, but it's still not perfect Information. It's pretty cool, but it's not perfect.
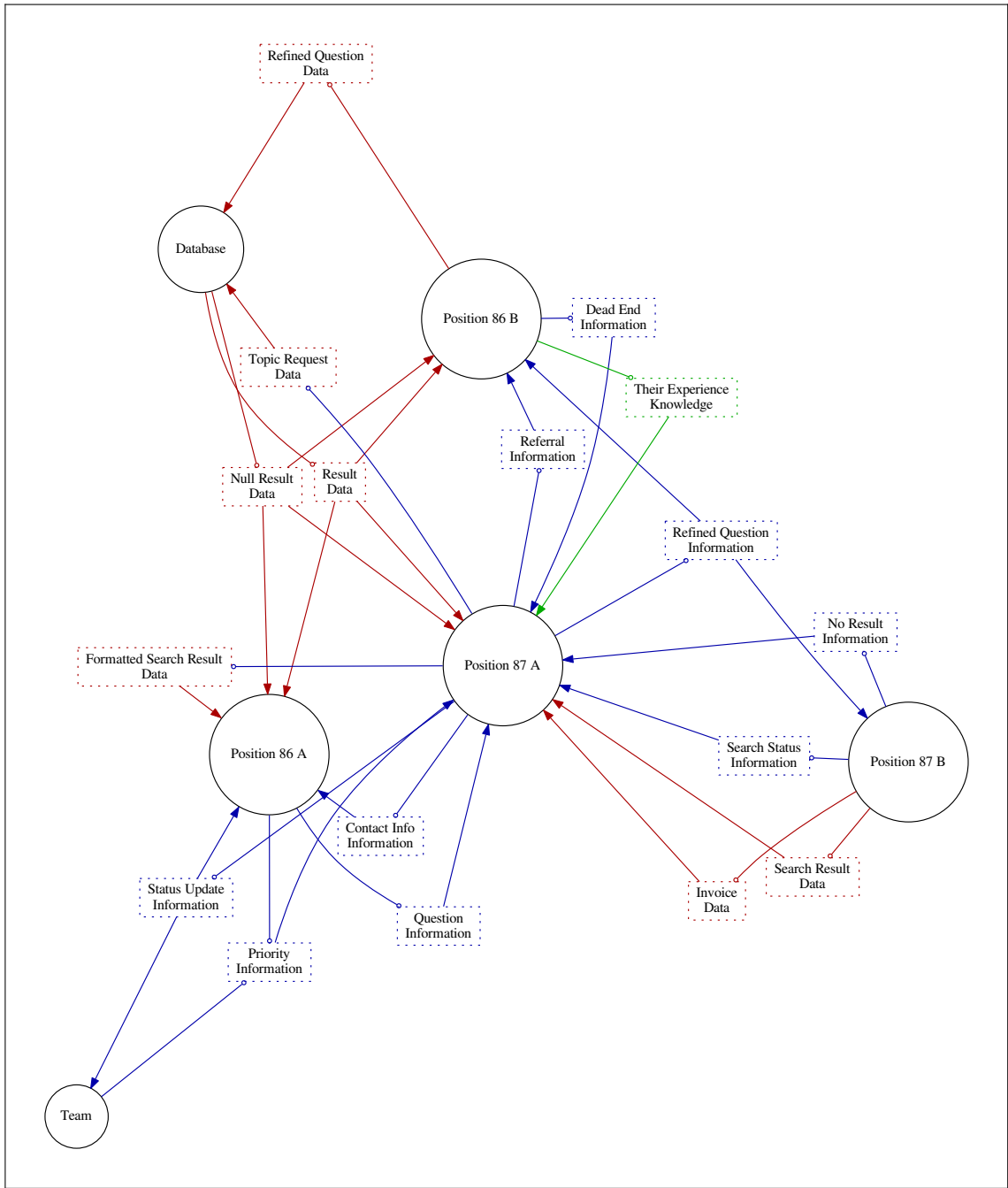
Figure 7.7: The SDFN Diagram for Interview 5

Interviewer: When you say we get Information out, This is what I'm looking at. What do you mean by Information?

Participant: You can dig down, in a controlled way. You can excavate the noun 118, taking noun 37 of materials and analyze their chemistry

or their micro-structure or their metallurgy or something like that. And you'll be looking at the physical shapes of what's where. ... Plus also their micro-structure and things like that. So, as you go down through the noun 118, you're looking for zones: "What's the material structure in each zone of the noun 118." Before things have melted, at the zone things are melting, below that point. And you're trying to estimate what the physics or the chemistry was that's happening when the noun 118 was running at that point in space.

...

Interviewer: So, we start with a nice trivial flow. What is one of your roles that you play in this project?

Participant: ... position 56, doing research experiments.

Interviewer: Are there other ... position 56s?

Participant: No.

Interviewer: So we can just call it ... position 56. At some point, you either get a flow from somebody or send a flow to somebody. Who is your most common person or thing of interaction?

Participant: Specific names, or are you looking for the entities?

Interviewer: I'm looking for the roles. If names help you think, feel free to use them. I edit them out during the transcription.

Participant: During the experimental stage, it was the position 127. ...

Interviewer: Do you send stuff to him or does he send stuff to you?

Participant: Both ways.

Interviewer: Let's model one of the things that you send to him. I'm using stuff here so I'm not going to prejudice you in terms of ... what we're going to be doing is we draw a flow like this, we label the flow with the topic of the flow, and then we label the flow with category of the flow. Whether it's Data, Information, or Knowledge. And if you say O, we go into depth with what you mean, and what the category implies. And I'll keep asking you this question until we're done. There's no temporal model here, everything just happens at the same time. Because dealing with temporality is just... too much of a pain. We have a flow from you ... What do you send him?

Participant: During the experiments, I'll send him observations from the results.

Interviewer: Observations? Or what shall we label this? Results?

Participant: I don't send raw Data. I send interpreted Data. I won't send raw measurements to him, but I'll send an interpretation of those raw measurements.

Interviewer: Let's go interpretation of measurements. Is this flow Data, Information, Knowledge, or O?

Participant: What would I call it? I'd be hazy on the difference between Information and Knowledge. Can we define it, or do you have a better definition in your head?

Interviewer: Here's the thing. If there isn't a difference, then there isn't a difference. I'm putting the terms out there. You can use them as complete synonyms if you wish. Or you can differentiate them. I'm trying to model what's in your head. Unfortunately I've got difficulty putting probes in.

Participant: It's a black box. I'd probably call it Information then. And I'd call it that because it's not just raw numbers. I'd plot an interpretation to it by some means. But it's not Knowledge because it's the experiment. And it's a very contrived environment that it's coming out of, so it's interpreted Data about this contrived environment. Saying when we do this on the experimental rig, we see this effect happening. And I'll often leave it fairly plain like that for them because I'm not familiar enough with their day to day observations of the noun 118 to know if this is a relict observation of our experimental rig or if it's really a phenomenon of the noun 118. It's not Knowledge yet. We don't want them to do something at the noun 118 based on that yet.

Interviewer: what does the position 127 do with this flow?

Participant: What they have done, in the past, considered it. Often it will be verbal or written. They'll consider it, and then they'll tell me whether it's consistent with something they observe on the noun 118.

Interviewer: How would we label that flow back to you?

Participant: Again, I think it would be Information, because we both realize, me and the noun 118 position 58 realize we're not talking about the same things. I've got a model as a chemical noun 120, and so we're talking about observations that we're making and an interpretation of what we're seeing, and we're looking to see if there's a correlation between the two.

Interviewer: And so this Information flow back to you. What topic is it? He's transforming your interpretation of measurements how?

Participant: I'll give a comment about a specific phenomenon that I see. Let's say it's a variation with flow rate and time. And I'll see, under a set of circumstances at a particular point. And that's what I tell him and he'll come back and may say: "that's interesting, but we have situation x,y,z. We see a similar effect."

Interviewer: So we would say that's correlation with noun 120?

Participant: He'll tell me if my Information correlates or doesn't.

Interviewer: So really, it's a confirmation of correlation. And you would say that that confirmation of correlation Information because of it interpreted but not situated?

Participant: It's interpreted, but it's not immediately relevant to the noun 118. We're not going to change anything. We're not going to

change what we do at the noun 118 based on that yet. It may change what I do with my experiment.

Interviewer: Let's look at that. He sends back a confirmation of correlation, whether or not it's positive or negative confirmation, he's giving you something back. Then, what do you do?

Participant: It depends on where the experiments are going, and how critical it is with these experiments. Often I'll say, "That's interesting" but I'm going to keep going with my experimental plan because we're only partway through or it was a side issue that's not an immediate concern.

Interviewer: Is your experiment an entity here?

Participant: Yeah, I guess it would. If I dropped dead, I would hope that somebody would keep going with the experiment. The experimental plan doesn't depend on me.

Interviewer: Do you send Data, Information, or Knowledge to the experiment?

Participant: I don't. I build the experiment.

Interviewer: You send a flow of Data to the experiment and what would you classify this Data flow as?

Participant: It's Data. I'm specifying the equipment geometry, the operational set points.

Interviewer: So you're sending specifications to the experiment... Do you send anything else to the experiment?

Participant: The experiment itself is not a conscious entity. So I don't think I could send anything else to it, I could only send it Data.

Interviewer: Does the experiment send anything to you?

Participant: Data.

Interviewer: It sends you Data.

Participant: I get numbers back from this.

Interviewer: Do we want to call that numbers or do we want to call that something else?

Participant: Data. I'll get.

Interviewer: Sorry, the flow, the label for the flow. It's Data, but what topic is it? Is it just numbers, is it results, or is it...

Participant: I'll get numbers from the various instruments we have, be they cameras or weights versus time, that sort of stuff.

Interviewer: So you send a data flow to the experiment of specifications. The experiment sends you a data flow of numbers back. Then what happens? What other interactions do we have with these two entities?

Participant: We'll also, for this particular project, we'll also have position 122s. Who are noun 95ing new equipment.

Interviewer: What interactions are they having?

Participant: For there we're, it's a stronger influence than the Position 22's. I'm actually wanting to give them Knowledge.

Interviewer: So you send them a flow of Knowledge or do you send them a flow of something else that they take...

Participant: They will get the same information as the Position 22.

Interviewer: So they get that flow?

Participant: They get that flow. They also get a flow from me about "I think you should do this."

Interviewer: And you asserting: "I think you should do this?" is? Let's first go: how do we label that?

Participant: I've done this experiment, I've asked the Position 22 and it correlates with noun 118 performance, therefore I think the new noun 95 should have these features.

Interviewer: Can we summarize that?

Participant: I'd say that's Knowledge.

Interviewer: That's Knowledge. Can we have a 2-3 word description of this Knowledge that's not that sentence?

Participant: Equipment noun 95 recommendation.

Interviewer: So you send the position 122s your interpretation of measurements and you send them a Knowledge flow of equipment noun 95 recommendations. Do you send them anything else?

Participant: As part of the Information flows to those, I may include a little bit of raw Data, but not very much.

Interviewer: Does this raw Data ...

Participant: usually photos or a graph.

Interviewer: So you would say that photos are raw data.

Participant: Yes.

Interviewer: Would you say that the raw data is contained in the Information? i.e. you send them interpretation of measurements. As part of that interpretation you have to send them some of the measurements that are really interesting. Would you say that the Data is inside the Information flow, and we can just label this as Information? Or would you say that it's Information + Data?

Participant: I'd keep it inside the Information flow, because if it was just raw Data. They could very easily reach what I think is the wrong idea – misinterpret.

Interviewer: Therefore you're not going: "Here's the Information, here's the Data." you're going "Here's the Information, here's some Data inside the Information to back it up."

Participant: Yeah, that's right. But with just the raw numbers and no context, that's Data.

Interviewer: You send them this Knowledge flow and this Information flow. Does the position 22 send them anything? or are there any interactions that way?

Participant: As far as I know, yes there were. Part of that was direct transfer of people. They actually got some Position 22s into the noun 95 team.

Interviewer: Would you say that's a flow of Data, Information, Knowledge?

Participant: Knowledge.

Interviewer: So there's a Knowledge flow, from the position 22 to the position 122s of personal expertise?

Participant: Yes. Personal expertise and personnel, physically personnel. But personal expertise.

Interviewer: Personal expertise of Knowledge. Would you say that personnel transfer is a flow of Data, Information, Knowledge or O? Should we model a personnel transfer as some sort of conceptual flow?

Participant: No, it's a convenient way of doing the Knowledge transfer.

Interviewer: So it's basically sneakernet.

Participant: it's better than just writing it down in a book and just posting the book and having them read the book. It's much more convenient.

Interviewer: So personal expertise as Knowledge flow from the position 127 to the position 122s... do the position 122s send anything back to the Position 22?

Participant: They'll send Information about their current noun 95.

Interviewer: so would we say current noun 95 or is there something more specific?

Participant: Current noun 95 and intent of operation. No, it has to be all together.

Interviewer: So you've got a flow of personal expertise and they return current noun 95 and intent.

Participant: and I'll say that's Information as opposed to... because of the intent side of it. It's not just raw Data. Data I tend to think of as just numbers.

Interviewer: If it was just current noun 95 it would be Data?

Participant: yes

Interviewer: But, because it's current noun 95 and intent, it's Information. Would you say that the Data is encapsulated as this one is?

Participant: Yes.

Interviewer: Would you say that there's no Data, there's just Information?

Participant: No, because the intent is embodied in the noun 95. But you need to know it in case they've misunderstood an operation and the Position 22 can correct, can feedback Knowledge, and say no no no, your noun 95 won't meet that intent.

Interviewer: And that feedback of Knowledge is the personal expertise?

Participant: Yes, that's right.

Interviewer: So they send this Information flow of Here's the Data + Information around it which is Data and Information, not just Information. Do you as position 56 get anything from the position 122s?

Participant: Yes. I tend to get that same stream. the current noun 95 and intent. But not so openly. I'll get it filtered through because the noun 95 is much bigger than what I do. There's much more and there's things that I don't care about.

Interviewer: So therefore it's a different flow because it's filtered. Do we just want to put filtered in front of this?

Participant: Yes. For example, in my project, I very rarely get electrical Information. Because I was looking at material flows, not electrical flows.

Interviewer: This flow of filtered current noun 95 and intent is... Information, Data, Knowledge, O?

Participant: Information. Information because it's not a real thing yet. It's not Knowledge about a real device. It's still a hypothetical.

Interviewer: Do they interact with your experiment?

Participant: Yes, I may change, just as the Position 22 feeds back, I might change the experiments because of that, likewise from the position 122's side. I have changed experiments because of what they've said.

Interviewer: Is that just a function of changing specifications or is there another flow to your experiment based on their Information to you?

Participant: I've changed noun 95, the noun 95 of my experiment.

Interviewer: Is there a noun 95 flow to your experiment?

Participant: I'd say that's part of the specifications. The geometry of it.

Interviewer: So geometry is part of specifications and all this is just Data.

Participant: It's just numbers and stuff. Based on that, I'll say I need to change dimension X by this or curvature Y by that.

Interviewer: Given this set of 4 entities, are there any other interactions?

Participant: We did have people building equipment. I don't know if you'll want to include that.

Interviewer: Are there stuff flows with these position 123?

Participant: Sometimes it was much more convenient for the noun 95 team to talk to the position 123 rather than me.

Interviewer: So the position 123 are interacting with the position 122s in your experiment?

Participant: And with me. So: experimental position 123 as opposed to people actually building the real noun 118. They were somebody else.

Interviewer: What flows are there from the position 122s to the experimental position 123?

Participant: I would give them plans.

Interviewer: Labeled as plans or drawings?

Participant: Drawings. Information. Wait, No I would say it's Data, it's numbers. I want a sheet of steel cut to these dimensions, surface treated in this way, joined in this way, join it in this way to this other piece of steel here.

Interviewer: And that's just Data.

Participant: That's just Data. They didn't tell them the intent of what I was going to do. The position 123 is not expected to interpret anything, they just manufacture.

Interviewer: If they were expecting to interpret, the position 122s would be sending?

Participant: Information, we would like the device to do this, please make sure that it can do this. But usually, if the position 122s were talking to them, they didn't say that. They said: here's the drawing, make that.

Interviewer: The experimental position 123 then do what with that Data flow?

Participant: they'll build equipment.

Interviewer: Does building of equipment represent some sort of flow of Data, Information, or Knowledge?

Participant: It's the embodiment of the Data.

Interviewer: Is embodied Data a flow of Data?

Participant: Yeah, I'd say it would be. After that, the equipment position 123 walks away and you have a device there which has the embodiment of Data in it. But you don't know where it came from.

Interviewer: So there is a flow there.

Participant: Yes.

Interviewer: Where is that flow here? Do we need an equipment entity?

Participant: No, I'd put it into the experiment. Because everything there was included in the experiment.

Interviewer: So this embodiment of Data is a flow of Data?

Participant: Data, yes. So the piece of equipment is Data.

Interviewer: How do we want to label this flow. Piece of equipment or embodiment?

Participant: Piece of equipment.

Interviewer: The experimental position 123 send a piece of equipment to the experiment. This piece of equipment is a Data flow because the piece of equipment embodies Data.

Participant: Yes.

Interviewer: Do the experimental position 123 send anything else to the experiment?

Participant: Not to the experiment. They send a bill to me or someone else, the position 124.

Interviewer: Do we want to model the position 124?

Participant: No.

Interviewer: So they, to you, send what? A bill? They can send multiple things.

Participant: Usually the piece of equipment, we either pick it up on a vehicle or it gets delivered here. They don't pass anything to me. What I get from them is the piece of equipment. It doesn't have to come to me, so long as it ends up at my experiment, that's all I care about.

Interviewer: So is there a flow from the position 123 to you?

Participant: No.

Interviewer: Ok. Are there any other entities?

Participant: There is a flow from me to them. And that will be of the same form as what they get from the position 122.

Interviewer: So you send them drawings?

Participant: No. From me they're more likely to get an intent. They're going to get Information. Because I'm not a position 122 or a mechanical position 58. And I know the function I want it to have, but I may not know that you need to make it out ... [of.]

Interviewer: So you, to the experimental position 123, send intent. Is this intent an expression of Information, Data, or Knowledge?

Participant: That will be Information.

Interviewer: What else do we have? Are there other entities that are interesting?

Participant: Involved with other position 56s. There are other position 56s who have helped from time to time.

Interviewer: And they're interacting...

Participant: They interact with me.

Interviewer: position 56s do what?

Participant: They'll, I'll show them... sometimes they actually help doing the experiments. What they'll do, they'll provide a – I'll happily show them Data from the experiments.

Interviewer: So you send them numbers?

Participant: Yes.

Interviewer: Are these numbers the same numbers, or do you transform them somehow?

Participant: No, if I'm going to talk to them, I'll send them real numbers.

Interviewer: Functionally you don't perform any other task than switching.

Participant: Correct. It's as if the experiment was sending them numbers. Not all like.. but for a time, a season... there's a fellow position 56 and I'll need to double-check or get a second opinion or "can you help me with this kind of thing?" and they'll get the raw numbers and we'll have it.

Interviewer: So the numbers from the experiment go to the position 56s. This I assume is a Data flow?

Participant: Yes

Interviewer: Do you send anything to the position 56s?

Participant: Yes, I'll be talking, in the system I've worked out here, Data is the raw numbers. Information is an interpretation of those numbers. And Knowledge is an embodiment of something, a real physical operating device of the company. So it's something that we're going to write down and keep for posterity and keep using. With the position 56s, ... I'll stop modeling the Data for a while and just talk about it. It'll be where... I see something interesting, but I want to make sure that it's actually a reasonable conclusion. That it's a sensible thing. That another person in my shoes will also reach. So I'll do the interpretation of the Data myself, and I'll also ask them to review it, to have a look at it. And see if it's a sensible conclusion. So they often would have gotten some raw numbers, but they'll also get the Information.

Interviewer: Is it the same Information that you would be sending to the position 22? Or is it different?

Participant: It'll be different.

Interviewer: What shall we label it?

Participant: It varies, from time to time. It'll be Information or it may even be Knowledge about an experiment.

Interviewer: Are these different flows?

Participant: Yes.

Interviewer: What is the Information flow about the experiment that you as position 56 send to the position 56s?

Participant: I'm trying to thing of specific examples here so I'm not vague.

Interviewer: Vague is fine.

Participant: If it was an Information case, it would be, I have seen ... when we do x, y, z, I see behavior a, b, c in the rig. And I'd call that Information. And I'd usually ask them to review the Data, look at the Data, and say do you think that's reasonable observation? However, if I've seen a class of behaviors or a systematic response to the rig to whatever it is I throw at it, then I would say – then I'm talking to them about Knowledge. And it will be Knowledge about either the rig itself or about some phenomenon we may see on the noun 118. So it's something to ...

Interviewer: So you send them an Information flow and a Knowledge flow. Let's label the Knowledge flow first. It's a Knowledge flow about phenomena or the rig?

Participant: It'll be as though it's a physical law of something or another.

Interviewer: So potential physical law?

Participant: Yes.. Potential physical law? It's something on which I can predict the future.

Interviewer: whereas the Information is... correlation? request to confirm correlation?

Participant: It's a request to confirm correlation. I think I'm seeing behavior x, y, z. Do you see it to?

Interviewer: But, because it's only a correlation, it's only Information?

Participant: Yes, that's right, it's not a predictive tool. I can't... rather than saying, I think I understand the mechanism behind this thing that we're seeing, and I think that this is the explanation for it, and if that explanation is correct than I can predict the future. If you throw x, y, z, at it it's going to produce this effect. That's probably the Information and Knowledge difference.

Interviewer: They have three flows going into them. Then what happens?

Participant: They'll hate me because I've given them work. They'll laugh at me...

Interviewer: Just a second, is that a flow? Is that expression of resentment?

Participant: Actually, no, it's often more interest. Particularly when it's the Knowledge kind of thing.

Interviewer: Is that a flow?

Participant: Their immediate response before anything else happens?

Interviewer: Yes.

Participant: I guess so, a positive response is always encouraging.

Interviewer: What it is a flow of?

Participant: Peer support.

Interviewer: There's a flow back to you of peer support sometimes.

Participant: Peer/professional support. This is going to get more vague and nebulous as we go on. I'd say it's just Information. It's very ... it's not numbers, it's not ... although it could be... no, ... this morning, we talked about the case where [someone was] saying something ridiculously beyond the laws of physics. Even a first year undergraduate would recognize that it's just a loopy kind of thing to say. And how do you correct that kind of thing? It's more than just Knowledge, the professional support of a positive response to a hypothesis is kind of acknowledgement that things are... that you're on the right track. So it's Information there. Because the person has noun 97ed it at least slightly. Ever so little bit. And it hasn't knocked up any red flags to say that this is stupid. So it's Information. It's an immediate interpretation of what's in there. They haven't looked at it in detail and found specific problems or errors with it or things like that.

Interviewer: Now are there any flows beyond that "Hey, that looks cool."?

Participant: Eventually, I hope that they come back with a technical comment about it, that they've thought more deeply about it. That would be a critique.

Interviewer: So this critique is: Data, Information, Knowledge?

Participant: Information I would say. And there will also be times where they will also come back with Knowledge. Often based on the Data themselves.

Interviewer: So they can send to you Knowledge. This Knowledge flow is?

Participant: They look at the raw Data, and they think they can see systematic behavior which can be used in a predictive way to predict the future.

Interviewer: Would you say it's the same Data flow that you send to them?

Participant: It's of the same nature, yes.

Interviewer: Does it have a different label?

Participant: No I give it the same label.

Interviewer: Really, it's a bidirectional data flow.

Participant: And that's really good when that happens. In part, that's the kind of the embodiment that I'm replaceable. And that we are actually doing science. The fact that somebody else can look at the raw Data and come up with that Knowledge is good. That's how science operates. It should be independent of the observer.

Interviewer: One of my other hats is philosopher of science. I would say that you guys are engaged in a Lakatosian research programme

<explanation of Lakatos>

Interviewer: From what you described, the refinement of theories and the research of theories is a research programme. Because you're not really questioning your inner assumptions of "this is Noun 23, this is what it does. We know it, we're not going to go: 'hey, I've discovered a new property of Noun 23' we're going 'hey, here's this refinement of the outer shell.'

Participant: that's right. ... we don't question the laws of physics. We try to find out how they're expressed in a very specific set of circumstances.

...

Interviewer: Do we have any other interactions here? Perhaps a couple of Knowledge interactions? Interactions with ...

Participant: I guess we never got to the end of it, the other roles or things like that. Do we want to go onto other roles of me?

Interviewer: Absolutely.

Participant: Because then we have time... during the first adjective 96 we were involved in, we would actually take measurements and giving immediate advice on those. It was convenient for me to be in that role, but it didn't have to be a position 56 as such.

Interviewer: But it was part of this noun 97?

Participant: It was part of the project, yes.

Interviewer: So what shall we call that role?

Participant: Whether the person in the position 56 hat to be that person or not?

Interviewer: That's the beauty of just labeling it as a role. You as the position 56 can send flows to you in that role. We don't have to instantiate roles really.

Participant: I'll talk around it a bit first to help clean my mind. The person who fulfilled that role needed to be in charge of a small team. Who would take measurements using temporary pieces of equipment on the noun 118. This isn't using normal noun 118 Noun 4, we actually put extra cameras and extra this and extra that up there to take measurements of what was happening when the noun 118 and bits of equipment were being run for the first couple of times. They weren't being run with the noun 118 being all operating, So it was slightly contrived. But it's part of adjective 96 to make sure that everything works properly. that you push button a and things happen as they should.

Interviewer: "If we push button A, does pipe A actually flow?"

Participant: Does it do the thing we expect it to do? During that there needed to be a bunch of measurements taken so that they could be

analyzed to confirm that the noun 95 intent was being met. So that it could be confirmed to be the noun 95 intent as best as we could tell. And you needed somebody in that role who was familiar with the noun 95 intent and with the equipment that was needed to take those extra measurements.

Interviewer: Let's label this role.

Participant: That's adjective 96 ... are we separating the role from the actual measurements? I would say that's that way. It's different from experiments. noun 126s is what they were.

Interviewer: These noun 126s, what entity sends them flows?

Participant: The main purpose of the noun 126s is to find out that when you press button A that things happen as they should So it was the position 122s, they send to the noun 126s, I guess it was Information.

Interviewer: Information about?

Participant: Not the noun 126s themselves. They're not a self-aware entity. So they just receive Data.

Interviewer: So here's the thing. It depends if the noun 75s are

Participant: There wasn't somebody who wasn't on the noun 95 team or in the Position 22 camp who was running the noun 75s.

Interviewer: So if you want to put that person in the noun 75s you're welcome to, or you can have them be a separate entity.

Participant: Yeah, position 125s.

Interviewer: do we want to make an entity for the position 125.

Participant: Yeah, separate the people from the actual noun 97.

Interviewer: So the position 122s to the noun 75s send what?

Participant: Data, numbers.

Interviewer: These numbers have what... what numbers are hey sending?

Participant: It could be embodied Data in physical equipment. So the actual noun 118 top.

Interviewer: Equipment.

Participant: Equipment and that's what they provided to the noun 75s.

Interviewer: What are they providing to the position 58?

Participant: They're sending him Information. And that was the designs and the operating manuals.

Interviewer: Flows from the position 58s to the noun 75s.

Participant: From the position 125 to the noun 75s?

Interviewer: All of the above.

Participant: The position 125 also got Information from the position 22.

Interviewer: And what is this Information?

Participant: It would be a request. They tell them, when we usually run the noun 118, this is how we tend to operate the equipment. I want you to make sure we cover that kind of operation in your noun 75s.

Interviewer: So request for noun 75s maybe?

Participant: Request for performance is a better description. For example, there's no need for the position 125 to run the conveyor belt by startstopstartstop. You don't need to test that because we never do that. It actually is physically capable of doing that, but we don't care about that kind of testing. It's not going to happen.

Interviewer: To test it the way we want the performance.

Participant: The way we usually do it and the reasonable boundaries that that...

Interviewer: and these requests for performance are Information. Do the position 125s receive anything else from anybody?

Participant: Yes, they've got Information also from me as position 56. And that would be Information of a similar nature, but to say this is the new behavior that we're looking for, please make sure that the noun 75s test that area. We want to measure the behavior under this circumstance. Please make sure that circumstance happens during that noun 75 and that we have the ability to measure it.

Interviewer: So request for measurement.

Participant: Request for measurement. And that was just from me though. I was the position 56 so everything was coming through me, or that role. And then the position 125 wouldn't give feedback to any of us. They said what was possible. So they told us what their plan was.

Interviewer: Were they talking about plans or feasibility?

Participant: Plans.

Interviewer: Who were they sending plans back to?

Participant: Everybody who told them things. ... And then it was the same Information came back to all of us.

Interviewer: So they're sending their plans as Information to position 56, position 22, and position 122s.

Participant: They found it convenient to send everything to everyone. Because then everyone could say: "No no no, you've missed X, please change your plans."

Interviewer: Is that no-no-no you missed X a different Data flow?

Participant: no, it was reinforcement of the previous Information. ???

Interviewer: Are there any Knowledge or Data flows to the position 125?

Participant: Well, there would be some Data included with the Information. It's usually interpreted Data. You don't want to give raw numbers coming from anybody. We're expecting them to generate or their noun 126s to generate raw numbers. Then they'd run the noun 75. So they send Data to the noun 75. They say the set points.

Interviewer: So noun 75 specifications?

Participant: Yes.

Interviewer: These are Data?

Participant: Yes.

Interviewer: These noun 75 specifications as Data are sent to the noun 126s. These noun 75s take the noun 75 specifications and the equipment, the Data embodied in the equipment, and do what?

Participant: It just operates. When I think of noun 126s, I think of the piece of equipment that's over there set up in the way it should be plus people who are there who might be actually operating or pressing buttons. But they're not necessarily... these are my instructions. I must press this button at this time. Or talk to so and so. And we do the noun 75. Just do this.

Interviewer: These noun 75s send what to whom?

Participant: We get Data. The position 125 gets Data from the noun 75. Everybody who's going to get it. In actuality, all four, the position 125s, the position 122s, the position 127s and myself were all there.

Interviewer: Basically the noun 126s send a flow to *2 which is the distribution.

Participant: Yes, that's right. As well as back to the position 125s.

Interviewer: That and this flow is?

Participant: adjective 96 Data.

Interviewer: And this adjective 96 Data is Data?

Participant: Yes.

Interviewer: Do the noun 126s send or get any other flows?

Participant: No.

Interviewer: Are we missing anything?

Participant: So the noun 126s are in a way similar to the experiments in the way that there's Data going around. I think that'll do, because then you get into noun 118 operation. And that's not part of the... most of us are gone, it's just the Position 22s who are doing their thing at that point. Again, from the noun 126s the data that say I get as position 56, I'll feedback Information to the position 127s, as Information. But I may also feed back Knowledge at that point.

Interviewer: So you're sending interpretations of measurements, but they're just different measurements. But it's the same flow.

Participant: It's the same flow. At that point, I may be sending Knowledge to the Position 22s. And that will be a different thing. That'll be a further refinement and converting the Information to give it a predictive capability.

Interviewer: how do we want to label this?

Participant: Operational understanding.

Interviewer: Does anyone else get this operational understanding?

Participant: Various people in the technical community. That may also go to the other position 56s, but it's not really of use to them. It's more politeness to them.

Interviewer: Fundamentally speaking you're sending the flow to the Position 22, though other people may receive it

Participant: It gets recorded in the library, that kind of thing.

Interviewer: Let's talk about the library and books (which are different.) does the library have a role here? And do books have a role here?

Participant: Oh yeah.

Interviewer: Let's start with books. What interactions do these entities have with books?

Participant: Internally generated documents? Or any kind of recorded Knowledge?

Interviewer: Do we want to differentiate those?

Participant: Probably not. Myself and the position 122s and the Position 22s. But mostly myself and the position 122s. We'll get Knowledge from books, from sources.

Interviewer: Call it books?

Participant: Call it books, but it means books and printed matter or internet things. Not necessarily Wikipedia. Trusted sources of Information.

Interviewer: These books, as trusted sources of Information, are sending ....

Participant: Trusted sources of Knowledge.

Interviewer: These books, as trusted sources of Knowledge, are sending these three entities, although these two in particular, what?

Participant: It'll be things like the laws of physics as applying to something close to our area, as best as we can apply to our area.

Interviewer: These laws of physics are Knowledge?

Participant: Yeah.

Interviewer: Do you, or any of these entities, send anything to books?

Participant: I'll record. I try and write reports and documents of my experiments once they're complete.

Interviewer: So you as position 56 send documentation.

Participant: I'll create documents, I'll create books. Not very big ones.

Interviewer: But you're sending to the entity books...

Participant: Information and Knowledge. Sometimes...

Interviewer: In the same flow or different flows?

Participant: In the same flow. Sometimes I'll be able to deduce some Knowledge.

Interviewer: So you send a flow of Information and Knowledge... This flow of Information and Knowledge can be labeled...

Participant: What we did we call the data coming back to me?

Interviewer: Numbers.

Participant: Can we change that label to experimental Data?

Interviewer: Yes we can.

Participant: The experiment sends position 56 and position 56s experimental Data. And my Information and Knowledge that goes into books or documents is, call it experimental interpretation. Analysis? Maybe analysis. Cross out interpretation. Experimental analysis. I use the term analysis in a lot of titles.

Interviewer: Do we want to generalize this documentation or books?

Participant: Documentation.

Interviewer: So you send Information and Knowledge in your experimental analysis to the docs. The docs send, to these three people, applied laws of physics, or pertinent laws of physics.

Participant: Yeah, pertinent. ... We differentiated here between the library and the books.

Interviewer: Is this here a worthwhile differentiation?

Participant: The library is merely an internal convenient store for it as opposed to, we can get into the local library down at [town name.].

Interviewer: So documentation is location independent, so therefore the library exist as an entity outside of its documentation?

Participant: Yeah.

Interviewer: How do you request documentation? Is that a flow, or is that a flow or do you just get appropriate flows from them?

Participant: They'll construct catalogues and things of what they've got.

Interviewer: Is the catalogue pertinent to this project?

Participant: No, not really. For the sake of this project, we don't generate that many books. Putting the library in it would be overkill.

...

Participant: I'm thinking that the overall project is for the noun 95 of the equipment of Noun 2. And I'm thinking that we've pretty much covered everything in this project.

Interviewer: Let's have a theoretical discussion. What happens in programming?

Participant: Computer programming?

Interviewer: Yes. What kind of Data, Information, Knowledge flows are there?

Participant: Do you want me to think of a specific computer program that I've written?

Interviewer: When you're programming, what do you send into the program?

Participant: It depends if I'm writing a quick little thing to do some analysis of Data, there'll be something I want to do to the numbers, so the program is going to be fed some numbers, and it's going to spit out some numbers, or print a graph or something like that. But I want to transform those numbers in a particular way, If it's quick like in an excel spreadsheet, if you're doing a little formula, it could be just some Information. But I say noun 97 these numbers in this way. Sum it over the ??? something like that.

Interviewer: and you would say that instruction to the computer is Information?

Participant: The intent of what I'm doing is Information to it. Because I don't know how it's actually going to do it in the CPU. But I'm telling the device... you fill out these numbers in this manner and tell me the numbers at the other end.

Interviewer: Whereas a more sophisticated program, what are you sending it?

Participant: Now, let's say one of the models, I've written a Noun 2 model or something like that. I usually think of them as an embodiment of Knowledge. There are laws of physics and the laws of chemistry and all of that that we can express as equations, and we can put them in in a generic kind of way then we have this model that has our Knowledge in there. And then we feed it numbers in and numbers will come out and they'll be transformed according to those laws of physics and the laws of noun 109 in there as well.

Interviewer: Physical laws.

Participant: Yeah.

Interviewer: So you would say that... your transmission to the program of physical laws is a flow of Knowledge? which the program then embodies?

Participant: No. I tell it Information.

Interviewer: You tell it Information... it then embodies...

Participant: I watch it transform, but then that Information is an embodiment of Knowledge. I know what the laws are, but the computer is just a machine. It doesn't know that... it doesn't know that the law of gravity is actually useful for saying that things fall down and break. We know that. It gets Information. I won't even tell the hardware. You're talking to the software that's running on top of the hardware. "I want you to do whatever you do to manipulate these numbers in this way."

Interviewer: So you would say that code is Information.

Participant: Yes. Based on Knowledge, but it's just Information.

Interviewer: But what you're sending is not Knowledge, it's Information.

Participant: That's right.

Interviewer: You would say that the ability to make predictive assertions about the universe is a function of?

Participant: That's Knowledge. That's an understanding of reality. So you need to have consciousness and self-awareness and that kind of thing. So my laptop, I don't know about yours, but mine definitely doesn't have self-awareness as far as I can tell.

Interviewer: Let's look at this. Self-awareness is a vital prerequisite for.... sending flows of... Information and Knowledge?

Participant: Knowledge in particular.

Interviewer: Without self-awareness, Knowledge cannot be transmitted?

Participant: You can transmit it into a book, but then that book is frozen communication.

Interviewer: So it's just an embodied... frozen communication as you said. An entity is talking to another entity via the communications medium of a book?

Participant: That's right.

Interviewer: So the book itself does not have awareness, therefore it cannot operate upon its Knowledge. Even though it's the vehicle for Knowledge transmission.

Participant: That's right.

Interviewer: So how does ... it doesn't actually sound like you're describing a hierarchy here in terms of Data, Information, and Knowledge. Or does it? Tell me at any point if what I'm saying is incorrect. You would say Data are... would you say that Data is or Data are?

Participant: Data is plural. Datum is singular.

Interviewer: I've found different professions have evolved different plurality rules.

Participant: [name]'s very demanding about that.

Interviewer: So these Data are raw numbers?

Participant: Yes.

Interviewer: When you say raw numbers, what do those numbers represent?

Participant: They could represent anything, they're just numbers.

Interviewer: So any numbers are Data? or are some numbers not Data?

Participant: There's some Data that's not numbers.

Interviewer: What Data is not numbers?

Participant: Qualities of things.

Interviewer: So qualitative and quantitative Data?

Participant: That's right.

Interviewer: and qualitative Data are textual?

Participant: Bigger smaller higher. It's something that you haven't been able to put a number to, but you think it's important or that it's worth noting. It's some... Information that you want to convey. That's not the right word. It's something that you want to convey, a quality about something. And so you'll try and give it some numbers if you can, but otherwise it'll be qualitative.

Interviewer: Would you say that qualitative Data are Data or Information? Despite the label?

Participant: It's probably getting into Information then, because you have to know the context. And ways to understand the terms that are being used. So say it's a hot day. What do you mean by that? Do you mean you're an Englishman or an Australian when you say that? Or an Eskimo. An Eskimo and a Hawaiian are going to have very different ideas of a hot day. So it's Information. Context is going to be something. But if I say it's 42 degrees Celsius, then except for the Hawaiian everyone else probably thinks that's hot. But it's actually put into a ... it's not a comparative ... You've given it in some kind of absolute term.

Interviewer: How does Data relate to Information?

Participant: Data itself is even useless without a context, you need to know what it's about. but Information has a richer context and it's less precise.

Interviewer: The maximal precision is in Data?

Participant: Yes.

Interviewer: And the minimal precision is in Information? Or is there something outside on the precision graph?

Participant: There's more room for interpretation in Information rather than Data. Data is meant to give you very little room to wiggle in. It's meant to be precise. But then it's precision kind of restricts its usefulness to some extent. There's nothing hanging off the sides of it.

Interviewer: So there's an inversely proportional relationship between precision and scope? So we've got Data. Data is ... I don't think we actually have a definition there.

Participant: Data is numbers.

Interviewer: And these numbers are? What do they represent?

Participant: They could represent anything. Any time you write numbers down it's Data.

Interviewer: So Data is merely written numbers?

Participant: Yeah.

Interviewer: These written numbers can be turned into Information?

Participant: Yeah.

Interviewer: They can be. So there's a relationship between Data and Information?

Participant: Yeah. There'll be a context to interpretation.

Interviewer: So Data + context + interpretation?

Participant: Data + context is enough for Information.

Interviewer: equal Information.

Participant: Yeah.

Interviewer: What provides context?

Participant: Metadata on the side.

Interviewer: So metadata provides context. i.e. other Data, juxtaposed with Data. Does anything provide the addition? Or is it just Data + metadata = Information?

Participant: IST may come from the thing that's generating the Data itself. So adding a title string to a file or something like that gives you a context. It helps you remember about it.

Interviewer: So there's metadata which provides context, but there's nothing which attaches the context to the Data save for the context.

Participant: Yeah, I guess so.

Interviewer: Can Information become Data?

Participant: No, not really.

Interviewer: So there's a flow that way but not that way.

Participant: I guess it could. But you have to change what you're looking at. Information can become Data but not when you're talking about

the initial thing the Data was about. You have to kind of change topics. Looking at a different scope of things...

Interviewer: Is there a backwards ...

Participant: No, not within a system.

Interviewer: It's data + Context becomes Information. and Information is irreducible to Data within the confines of a given scope.

Participant: That's how I think of it, yes.

Interviewer: This is what I'm getting at. Can Information become anything else? Or is Information the upper bound?

Participant: In the same way that Data flows into Information when it's given a context, the Information can become Knowledge.

Interviewer: With the addition or interaction with?

Participant: I'm trying to think about what's the thing you have to do to the Information to generate Knowledge. It's stronger than interpretation. Plus understanding?

Interviewer: Information plus understanding.... Understanding of the Information? Is understanding a function or a separate thing?

Participant: I don't think I kind of... a mental model is that there has to be some kind of consciousness looking at the Information to generate Knowledge.

Interviewer: So you go, Information and understanding are arguments to consciousness?

Participant: Yes.

Interviewer: Might as well use familiar terminology. Consciousness, taking in arguments of Information and understanding produces Knowledge.

Participant: It needs Knowledge. The understanding is actually Knowledge. So Information plus other Knowledge generates new Knowledge.

Interviewer: Can Knowledge become Information?

Participant: When you constrain it.

Interviewer: How do we want to represent that?

Participant: It's applied Knowledge.

Interviewer: So Knowledge + .. - maybe? What metaphor is it?

Participant: Knowledge applied to a specific situation can give Information. So, for example, the law of gravity is Knowledge. And I apply it to the earth-moon system, and the Information I can get is the tides or something like that.

Interviewer: It's not Knowledge, it's Information.

Participant: Yeah. So that there's going to be a high tide in 2 hours or something like that, that's Information.

Interviewer: Are there direct interactions between Data and Knowledge?

Participant: Yeah, because the Knowledge is the law of gravity or something like that. And you'll throw some numbers into that and it'll spit out some numbers.

Interviewer: So Knowledge + Data = Data. And this new data is what, predictions?

Participant: Yes. So, Data + Knowledge = Data.

Interviewer: So Data + Knowledge is what kind of Data?

Participant: That'll be Data. So Information + Knowledge will give Information. So like the earth-moon system, that's probably more... So now we know that there is an earth-moon system. And we're going to apply the law of gravity to apply the Knowledge to it, and it'll become some Information about the earth moon, so I can make general comments about it.

Interviewer: Information, there is a solar system. Knowledge, here is law of gravity, Information comments applying law of gravity to earth-moon system. Deriving LaGrange points would be Information?

Participant: No, because you have to put in the masses of different things.

Interviewer: So that is Data + Knowledge is Data. So you're going these numbers plus this way of manipulating these numbers produces other numbers.

Participant: It's a number where you can put a rock and it'll sit there in space.

Interviewer: Are there any other directions that Data can take here? Can Data become other things outside of Data, Information, or Knowledge?

Participant: I can make up words, but it would just come back to the same idea.

Interviewer: Is Data atomic, divisible?

Participant: Well, Data, except the plurality... I'd say it's atomic. It's as simple as it gets.

Interviewer: So you can't get below a datum?

Participant: No.

Interviewer: Is Knowledge ... is there anything that is to knowledge as knowledge is to Information?

Participant: No.

Interviewer: So we have bounds. Data and Knowledge. Nothing outside of the bounds.

Participant: I don't think so. Physical reality, but I'm not sure where you can philosophize then. Well, maybe that's below Data. Physical reality is reality. Real is, that's it.

Interviewer: So therefore, we have reality on this?

Participant: Yeah, put it in. Yeah, because numbers aren't physical reality, numbers are numbers. They're a quantification of physical reality. So it's even more underlying than Data.

Interviewer: We have what applied to physical reality to get Data?

Participant: Usually a device of some kind.

Interviewer: Do we want to say device, or is device doing a verb?

Participant: OK, I'm going to read the Noun 12 in the room at the moment. So we get a thermometer. But it needs to be calibrated and I need to know that thermometers tell me about Noun 12, so I need to have Knowledge about the operation of them.

Interviewer: So we have physical reality, but in order to get Data out of physical reality,....

Participant: We've got physical reality, but I actually need some Knowledge to quantify physical reality.

Interviewer: The act is quantifying physical reality via Knowledge to get Data.

Participant: Yeah.

Interviewer: But that's a one way function. And then Knowledge floats around being Knowledge.

Participant: Yes. It's frozen in books. And if conscious beings want to make their existence comfortable in this reality then you're free to pick it up do with it as you wish.

Interviewer: Final question. Normative assertions. When you say you should do that, would you say that you're passing them Data, Information, Knowledge or O?

Participant: When I say to my kids, you should do your homework, it's Information.

Interviewer: When you say to an position 58, is that Information?

Participant: How many sentences am I allowed to add at the end of it? If I just say: "You should increase the blast Noun 12 of the noun 118." that's Information. But they know how to do it. If they want to say why. They'll probably come back with "Why" or "No." But that's the same thing.

Interviewer: If they come back with no, how do you respond, do you give them a flow of Information or Knowledge or Data?

Participant: It depends on the case. I may give any one of those three. It could be Data, because I might say to the Noun 20 and say "you should increase the Noun 12" why? "because your blast is at 900 degrees." And

it doesn't need to be said to them that that's ridiculously low. That you need to up it to at least 11 preferably 12 hundred degrees. So I could give them a piece of Data and they have enough Knowledge to go, say "Oh yeah. That's bad." they know that it's in error. Or a Knowledge situation is "I think you adjust the angle of this thing by .5 degree." and at that point I may give Information or Knowledge according to how much Knowledge I think they have or need. Or don't have.

Interviewer: But the initial statement is always Information?

Participant: Well, I may talk to them about a Knowledge thing, I assert a normative assertion, it'll be, a you should... I may say Data, I may say Information, but if I'm saying you should, it's not a Knowledge statement.

7.6 Interview 6

Participant: Out of that, I became really fascinated in data and knowledge and information because it's – that project – people had different views on it and that's why I'm interested in talking to you because you can say "about Noun 1" but it means something different to everyone else there. Even the form that it takes.

Participant: It becomes an inherently personal thing in the noun 129. I was developing this thing in a very personal view, and it wasn't shared by everyone. "You don't do that, that's not what Noun 1's about." but, to me it was. To some people it was about data. To me it's about knowledge. Other people it's about information. I guess we're here to ???. ...

...

Participant: Anyway, you might want to lead me through the noun 97.

Interviewer: Fair enough. ... You mentioned everyone has their own definition of Noun 1.

Participant: I'd say their own view. Probably not at the definition level.

Interviewer: Understanding? View?

Participant: I'd say definition because they don't actually describe it. View is they articulate a vision. A definition would be they actually write something meaningful.

Interviewer: Tell me what your view of a Noun 1 is, and what you think their views are.

Participant: My view of a Noun 1, it really should be about capturing different sources of Knowledge from different aspects, and that can be from different people, different approaches: You can have production thoughts, research thoughts. Mathematical models. Other motels, like data mining. ... And so you have disparate sources of how things work,
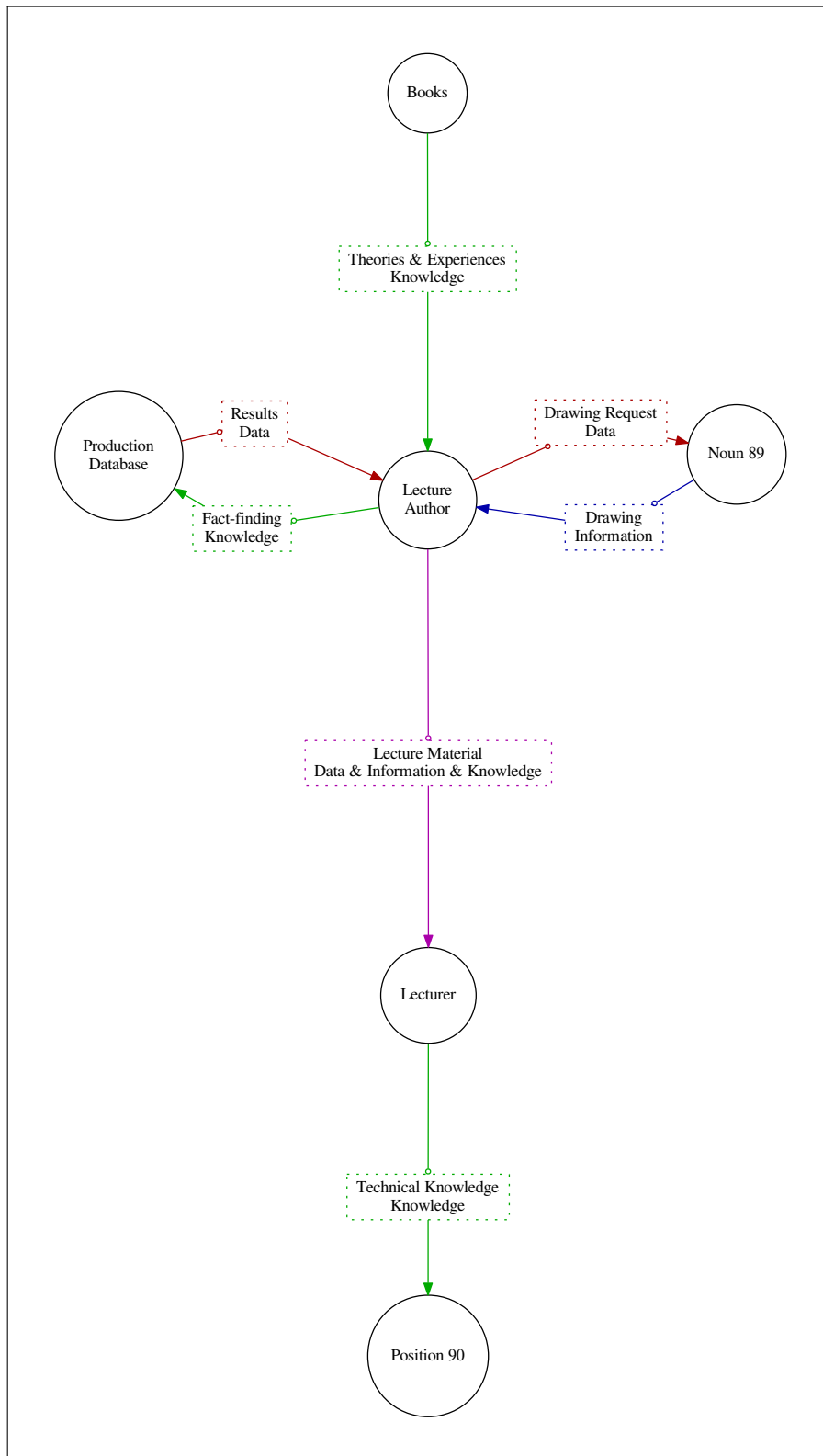
Figure 7.8:   The SDFN Diagram for Interview 6

and they all relate to making the Noun 2 work better. They think you should do it this way, and the computer model says we should do it this way. They don't always agree, but they look at it in different perspectives. And they say: "oh the Noun 1 is to grab those Knowledge and we refer to them as knowledge-bases and that's an extraordinarily old reference. Knowledge-bases together. And then to make a sensible input to provide a single piece of advice to a person. That was my view. A single piece of advice and I've implemented the Noun 1 as such. ... we have another person who has a different view on a Noun 1 and he see as much more as a Information system where you provide, as we might refer to: Information to people. And this is commonly, in our field, or philosophy, or where we come from, our research? as well as our [people] in other parts of the business see as Noun 3. So you have a valve that is open or closed, a 1 or 0, as a function of time. We have lots of Noun 4 with calculated outputs and ... A very other common view is to say that they should be provided that information to make their own... A Noun 1 doesn't tell anyone how to do anything, it just gives them stuff to look at. And maybe it's just a guided look, maybe you don't show them this valve, maybe you show them that valve. Nonetheless, it's about showing them things they want to look at.

Participant: So that's really the opposing, for me, that's sort of the two main approaches we talk about taking.

Interviewer: So advice versus providing information. Tell me about advice. What do you mean by advice?

Participant: With advice, I mean a specific action. You actually have to tell someone to do something. Advice is like: "Brian, catch the earlier train." Information to me would be: "Brian, here's the train schedule. You choose your best train. I'll just tell you the schedule and then you'll go on and sort it out for yourself." Where mine is if I want to help you make a decision about what train to catch I'll gather knowledge from the train master, from people who catch the train and present you: the 1:00 train is the best one for us. In comparison to other people saying "I'll collate the train timetable information, maybe I'll have indexes relating to crime levels. So you may then choose, 'I want that train, but it's got a high crime level, maybe I'll choose another train. It's a bit opnoun 71ue, you don't know how people made the decisions, you're just giving them information to look at.

Interviewer: When you say knowledge capture, is that what you meant by Noun 4, or is there a different term you mean by...

Participant: In terms of knowledge. Well,it's funny. I use it in a different way, but it's probably pretty similar. When we talk about the bubble [DFD people] concepts as inputs, we refer to that – like a Noun 4 as data as such. And Knowledge would be someone which would be putting that data into context. You might say: "When the valve is above [Noun 12] Celsius, I would turn the valve off" that, to me, is Knowledge, where the value itself is Data. It's like a hierarchy. Wierdly enough I set my system up in a hierarchial fnoun 137ion. I didn't want to do it, and

I really love case-based reasoning. Well, we can talk separately if it comes out of – the difference between case based reasoning and that form of knowledge compared to other expressions of knowledges, I refer to it as.

Interviewer: Data as Noun 4, Knowledge as the Noun 4 stuff with context with memories, you'd say? What do you mean by context?

Participant: You might have such as limits, there might be arbitrary statements like: "When the data is trending up" It may not even be specific. Sometimes you might say: "When a valve is a Noun 12 above [Noun 12] Celsius"

Interviewer: So it's generalized statements about a Noun 4?

Participant: Yes, that's right. That has been a reflection of a compromise how we've set up the work that they're doing over the last 5 years.

Interviewer: What about information? Is it different? is it the same?

Participant: In that hierarchial concept, Information sits between data and Knowledge. Again, A classic form of Information that I describe it as, it's multi-dimensional Information. That's the difference I draw here too. We're sort of getting into multi-dimensional information as we describe it. You should talk to [Person X] about this as well. You might have data versus time. It's two dimensional, it's a line. And we like making three-dimensional graphs of things so we have this versus that versus that. you can have a peak. We use a mathematical model to generate an optimal point and you talk to [Person X] at the university of [X] about this n-dimensional space which means nothing to me. ...

Participant: Information is really, its like... Data is something you can work with. Information is an expression of the Data in an almost trivial form. It can just be plotting it as a function of time on a graph or contour plot. Something like that. I hold Knowledge as the hierarchy of it all. Where it has that line in context of other things, or people's experience. It's where I come from. It's a personal sense, that it's a reflection of, that – I believe in mathematical models. They're absolutely excellent. but they need to sit alongside experiential knowledge. Mathematical models are, to me, information. Because they take data in and they can transform it into something is maybe trivial or non-trivial, but you might get a predictive Noun 12. And that's good, but, them, themselves, in my experience, are not actually good on their own. They need knowledge. Someone says: "I only look at that when this happens or that happens. Or it's no good today because that thing's broken. We just can't use it at the moment. But if you looked at it solely in itself. Well, that's very important. it can or not be, that is people's experience the Knowledge can help transform the data into information into knowledge.

Interviewer: Is there anything on either side of Data or Knowledge? Does the hierarchy continue even as not a hierarchy?

Participant: It's certainly a very [company] view. We often come to that ontology – It's sort of a common ontology that's developed within 10 years around the technical areas about these sort of things.

Participant: Do things exist on either side? Well, they do. and I think they would, and it's almost like other problems. To make my life easier, I ignore them almost. ... it's commonly accepted that I can top and tail that. ... Because I'm working within a known space. You could go out there, but I just don't. I can see its there ... it's good enough for me. ...

Interviewer: You said atomic. Do you believe that data is atomic, in the traditional sense, i.e. indivisible? Or is there there...

Participant: For our [company] ontology, yes it is. As my professional life, yes it is definitely. As a philosophical perspective, I can appreciate its not. But as far as philosophy in a [company], no its not. Even when we talk to the instrumentation guys, we still talk at that level, we don't go through it. You can sometimes touch on it. People sometimes allude to it that there's stuff there. "Well, we've got [number] minute data, and we want more data or different Data", but it's not commonly accepted"

Interviewer: The second part of this, as you know, is the drawing of the little circles. Let's identify two or three work related activities that we can diagram where you can illustrate what you meant in what you just told me. What do you think they should be? Here's a better question. What do you think your 2 or 3 most important or interesting work activities are, preferably ones involving small groups?

Participant: Developing rules, or knowledebases.

Interviewer: Noun 1 noun 129 you'd say? No, you said developing rules for the knowledgebases.

Participant: Yeah, that's specific...

Interviewer: Okay, developing rules, what else?

Participant: Well, the O, the fascinating thing for me which I argue is the debate about the description about the new ... system. called Noun 5: ... The Noun 5 system re noun 129. And then those two things are the most interesting related to data. The other things I really hold strong opinions on .... so they are probably the best place to start.

...

Interviewer: So let's identify the entities.

Participant: I never know quite which level to start at. There's always me, it it just me, or is it Os? I am the center of my own universe, I guess.

Interviewer: So this is you as what?

Participant: As a Noun 6.

Interviewer: So you as Noun 6. Are there any other roles that you play within the developing rules context

Participant: There are a few there. You can have me as the Noun 1 developer. There is me as ... it's a little bit of my general day job? It's

what I might commonly do? Me as a Noun 6? ... if I'm setting there today looking at my system, weirdly enough, they don't always happen to match up. Because that's, there's drivers driving me there. There's go and do this and I put on that hat. And I'm not quite in the same space. I can quite often self-justify anything you can do. Maybe I'll leave it like that, otherwise it gets a bit sort of hairy.

Interviewer: hairy isn't necessarily bad

Participant: But there are two main roles. I'm building systems and I'm typically using them as well, or using other systems. Those are the me roles.

Interviewer: Who do you talk to?

Participant: There's other position 56s, sort of group members. If you talk to a few people, there's some sort of common ontology of who these people are will come out. ...

Interviewer: how do I label that?

Participant: Exactly as I said it. It won't be personally identifiable. Can I put [name] there? There's Position 7 members. These are a very specific terminology. We have Position 8. These are different from Position 9. They're a different role in the controllers. These are various [company] specific positions. ...

Interviewer: We're talking about Noun 1 noun 129? No.

Participant: Noun 1 rule noun 129, Developing rules in the Noun 1. Interesting thing I do with data or how I look at these things.

Interviewer: Where do we begin?

Participant: Me as the center the universe, of course.

Interviewer: You as what?

Participant: Me as Noun 1 developer, because that will be my primary job.

Interviewer: Tell me about the flows of stuff.

Participant: Actually, another bubble there: it's a computer system I'm working on! It's definitely in there. We had that configured in [last interview] but – we'll have the Noun 1 as such, let's even say it exists. It goes weirdly through that metaphase it may or may not exist. As you're developing it it's sorta there there and sorta not.

Interviewer: It will have had existed?

Participant: Yeah, it might exist for a bit of a period during a noun 75, then it disappears. But it exists now, then it continues on. Then there's Noun 5, we talked about Noun 5. That's another computer system we deal with. I'll draw the line there, Noun 5.

Interviewer: Well, you're the Noun 1 developer, so let's do the nice simple start with Noun 1.

Participant: That's right. I would stay – I have my computer system.

Participant: I send it, what I would refer to – as a developer – I'm encoding knowledge. I'm sending it knowledge as far as I'm concerned. It's not easy to try to express humans as an if-then-else statement. Maybe we can talk this as some sort of psychological session later.

Interviewer: You're sending it knowledge...? Can we label this knowledge, or are you just...?

Participant: It's computer code. I'm sending it computer code. It's knowledge expressed as computer code, for sure. It actually sends me back Knowledge expressed as textual advice which is often been... ??? over the years.

Interviewer: It sends back knowledge as advice?

Participant: Yeah, as textual advice. I don't know why I make that distinction., but I've always tried to put the stuff it sends back, even to me, because I sit there and read it as textual things that sort of are vaguely sentences I best can do in 32 characters. Its like its text, it's words it might have a noun and it might have a verb or ...

Interviewer: Is this medium? Or is this part of the message? The text?

Participant: It's the message?

Interviewer: I don't see what you mean by medium there?

Participant: Is it sending the advice through text? Or is it sending text advice through something?

Participant: It's sending text advice through a range of different mediums. Because you can access it ... we didn't sort of go there, how people access it, but there is interfaces to how it goes out. And it changes a little bit depending on the interface, but there is sort of one interface which I consider which they really use in the Noun 20. I sort of focus, when I think about it, I think about that. That's my little box with little words on it. Noun 11 which are bit more descriptive. I've actually got a Noun 11 part which has a lot more stuff in it which you can do, no-one ever uses that so it's almost like it's ... I don't really support it.

Interviewer: What else do we have? So you send it code, it sends back to you advice. What else?

Participant: I'd put Noun 5 over here, because it sends information to the Noun 1. Data! It sends it data, for sure.

Participant: and that's expressed as values – it can be both real values such as a Noun 12 it can also be derived values such as an Noun 13 like that. We sort of draw distinctions a little bit between that.

Interviewer: So derived values as in calculation results?

Participant: Yeah. So you might calculate an Noun 13. So Noun 5 calculates some – an efficiency of something. I take that as well as I take raw data.

Interviewer: Do we want to say that as a different string?

Participant: Yeah we can. I certainly do derive differences between it.

Interviewer: What are we labeling this string as, these calculated results?

Participant: Derived data, we refer to them as. So it would be "Data, derived." As much as these are data like real ...

Interviewer: what else?

Participant: Out of here... My next primary focus is Noun 14.

Interviewer: What are they talking to?

Participant: They'll be talking to the Noun 1 and to Noun 5, and to a few other of these guys as well. This is sort of getting complicated already.

Interviewer: you have no idea. And tell me about this?

Participant: We're sending, Noun 1 is sending text advice as knowledge to those guys – to their Position 8.

Interviewer: and what else?

Participant: we've got to put the Noun 2 here. It doesn't make any sense without having a Noun 2 here, because they have to send stuff to the Noun 2. So that's an object.

Participant: I'm going to put the Noun 2 in the corner which may be very limiting in a way by cornering it like that.

Interviewer: we've got those little star escapes

Participant: oh, those little stars.

Participant: So it's sending – those guys send back data in the form of set points. So they'll want to change a valve position, say valve 7 to 45... it will want to return stuff to them.

Interviewer: it will or will not?

Participant: It may or may not.

Interviewer: is it relevant?

Participant: no. We can say here that this is a really good line... these guys send... it sends data and we term as real values. it also sends data as calculated values. These guys send, really they send stuff I refer to data as derived values back into Noun 5. Which is a bit weird. It actually was super important because these guys have a value they operate the Noun 2 by.

... Participant: ... They will sit there, they'll do calculations on a computer. they'll get some – they get a reading from the Noun 15 and they calculate a derived value which they'll operate. This is a new derived value we've got to operate

Interviewer: so do they send data to themselves?

Participant: Wierdly enough, a key part of the project early on. They have a whole heap of information – I feel like I'm in front of a barrister, with the terminology. These guys send information back – these guys have information which they didn't used to send to Noun 5, but

I needed it over here because I was telling them to do something, but they said "well dude, it changed." but I said "I don't know about it." so I went through a noun 97 with them entering derived data back into the system so I could feed it into my system to feed it back into them... there was definitely a loop there. There used to be a bit of a disconnect, because it was mostly data coming here, data going there. These guys would just spin around themselves....

Participant: Let's call that Noun 16? they only noun 97 Noun 17, I can give you an example of one. It might be nice for your thesis

Interviewer: I would love artifacts. ...

Participant: Yeah. They will send derived data back to the daily noun 97 Noun 17. They'll also send data as real values as well. But that's a dead end. That's a sheet of paper. Previously, there wasn't this link, it didn't exist before...

Interviewer: this link being the Noun 14 to Noun 5 derived data? You but that link in so you could reverse position 58 what they were doing...

Participant: Yes

Interviewer: How does the Noun 1 get that data from Noun 5 ... as derived data? And so, Noun 5 also sends it real values... where is it getting the real values from?

Participant: From Noun 5 as well.

Interviewer: Where is Noun 5 getting the real values?

Participant: That's why we needed the Noun 2 in here, because the Noun 2 obviously is doing stuff and it's sending real values. The Noun 2 doesn't really calculate anything. It sorta does it sorta doesn't. As far as I'm concerned it doesn't. But it has real values, that's just Noun 4, composition and things like that. And it will send that real values – it's weird, it's where you sit. Now we're talking about the philosophy of data, where you sit and things, because – you're right, you can pin me down and say well: "Why when it calculates the top down composition that I see that as real value data when it's not really if you think about it, because it's a instrument that's measuring atomic percents and then doing some calculations and reports its as percentage to me. And obviously, percentage is not a real value in terms of what that is, but to me, and as you flow through the chain of data, it transforms. It becomes into different forms. Maybe it's sort of related to the ontology. Well its not really, if you pin me down, maybe it's not. But it sort of is to me. It's percentage. Well, it's not real, but it becomes real. It's a Pinocchio thing. They want to be a real boy. Eventually, after a while, you get these bizarre analogies from me. After a while, once you pass it around long enough, I guess this happens – Probably well documented, Chinese whisper style, it can become real after a while. And this stuff goes from a non-existant calculated value to something that's real, measurable, and people love it. You look at its source, and often we have these problems. If you look at the source, the source may not actually maintain the love of over here, but, as you stretch out and

try to transform people's opinion of it changes. It does sort of – it can be loved, but if you have a look at it, and say no, "that machine broke ten years ago. Why are you still looking at the value?" and they say "Well, we love that. That's a real value." It's not doing the right thing, but through transformations and noun 97ing through systems peoples' experience becomes real when its sort of not.

Interviewer: And by real, you mean, a representation of a real thing?

Participant: That's right. It becomes a something that we can use. It comes back to data as an object or as a piece, a discrete sort of thing which is slightly different, as I said, to information, which isn't as discrete knowledge – which is weirdly explicit but not discrete.

Interviewer: wait, what is explicit but not discrete?

Participant: Knowledge. I'm being probably a bit loose in my terminology. But its like – Knowledge is explicit because I have to, in the Noun 1, as the Noun 1 developer. I have to code knowledge. So it has to be explicit. I've got to tell it to do something.

Interviewer: is the code Knowledge? Or is the Knowledge embedded in the code?

Participant: Code is Knowledge. For us, that's the case. But the reverse is obviously true. You can't have one without ... The code is the expression of the Knowledge. The code is the expression of the Knowledge that I have gained through various means.

Interviewer: Does this system output to you besides that textual analysis?

Participant: As a Noun 1 developer? Well, yeah... I'd draw that one in, I guess. That's a tough line to describe. This is between Noun 5 and the the Noun 1 developer. That line between Noun 5 and the Noun 1 developer – this is where it's complex, it is there because, as part of the Noun 1 developer, I'm interpreting.

Interviewer: we can have more than one line. It might be helpful to just do all the trivial bits and see what's left.

Participant: At the moment, there's almost – there's a data flow diagram for how shit happens. I get this and shit goes there and Knowledge goes there. But then there is, weirdly enough, the noun 129 noun 97 of making such a thing has its own different data flow diagram because this is where there are layers upon layers. Because I'm looking at Noun 5. Noun 5's feeding me stuff, and I'm feeding that. And those guys are talking to me which is the Noun 14s to Noun 1 developers. But maybe we're getting ahead of myself. Because we haven't quite – this sort of DFD that we've got here with Noun 1 developer, Noun 5, Noun 14, Noun 16, Noun 2 is describing a bit of the – maybe a different problem that we started off with. Maybe we can label this one as "Noun 2 operation optimization" which is really the purpose of the project and it's above. And it's obviously different from developing the Noun 1. Obviously,

we'll find similar things. Maybe it's hard to find where to draw the lines.

Interviewer: do we need a line there?

Participant: I'd just take the line out between Noun 5 and Noun 1 for the time being. Depending on the context, I guess. In Noun 2 operation optimization, I'd just..

Interviewer: in that domain, are there any other flows?

Participant: That's what I'm trying... maybe I can get a photocopy of this and go again? And we have that and we have Noun 5 and the Noun 2 has Noun 5 data and we send it to the guys. And here you have more – we left out a lot of circles here that we really should fill in. And we'll just stick to the Noun 2 operation optimization which is different from the Noun 1 noun 129. Maybe this is a more tractable problem. And interesting, and it's something interesting – if we can get you in the Noun 20. No promises there. If we can get you over that hurdle? This is sort of like their experience and how I interact with those guys. It would be good to do Noun 1 rule noun 129 because that's how I interact with the university people which don't – I can't quite – I can squeeze on here. I'm holding my hands above each other. As the layers of information. That's a column of information which then something may sit above. It's not two-dimensional as ???. Obviously me as the Noun 1 developer is dealing with multiple spaces. Maybe I'm playing it out maybe too two dimensionally.

Participant: So now we're going to go to. These guys talk to the Noun 9.

Interviewer: these guys being?

Participant: The Noun 14s

Participant: The Noun 14 and Noun 9 often interchange roles. But they see people with vastly different experience. But often when the Noun 9 is away, the Noun 14 has to step into his role in terms of Noun 2 operation. They send each other – this is really a tough one. They talk to each other. So I'd describe it as they do literally talk to each other. It's so hard to capture. And I had a go at trying to capture verbal feedback and things like that. These guys, they talk to each O, and they'll say: "Oh man, the Noun 18's broken." "Yeah, OK, change this." And my systems blithely – because they get a phone call "Yeah, something's going on. Yeah, backed up Noun 19 system" "Okay, yeah, we're going to make this change." Just lost, it's out there and thing. People have sorta become accepting of the limitation of this, over here, but it's really an issue. It's a really strong data flow I guess that I'm talking to... So they talk to each other.

Interviewer: is this talk: Data, Information, Knowledge or just talk?

Participant: I'd describe it as just ... there are a couple of different levels, but let's say on an operation optimization ... I'd describe it as Information. But they can actually. But no, no, it's true. We can put

another line in which is really they transmit Knowledge. I can call that experience. And that's absolutely true. They transmit experience in this direction sort of going out into there and often these guys [Noun 14] are younger guys without a high degree. Maybe they haven't been to uni. And [Noun 9] will be telling them experience and how to do it while interacting with my system as well. But they don't quite – often these are. I don't know how we can... maybe I can give it [the data flow] an AA. They're almost duplicate lines, but I'd like to acknowledge them as they're different. Because there are two people in the Noun 20. There's the position 8 and the position 9. They're both looking at my box, but they aren't separate people. They deal with things separately. It would be too trivial to combine these guys together, because they're different not from my experience. But for your purposes, I'm copying things now, like the Noun 1 provides the same knowledge by text advice to these guys as it does to them.

Participant: And similarly, you know, a lot of these flows where they're putting set points and getting derived data and calculated data between Noun 5 and the Noun 14 are the same as between Noun 5 and these guys. Similarly, they sink data into here, and so the Noun 9 sinks data into the Noun 16, the same as the Noun 14. And that's "real value" and "derived value." Maybe we need to get to the bottom of why that is different.

Interviewer: Why is that different?

Participant: It's literally, again, it's a hierarchy thing I think. ...

Interviewer: Position 8 are the boss of...

Participant: I didn't draw any information back from the Position 9 back to these guys because it doesn't flow that way. These guys aren't going to say, ... they might say: "Should I do something?" They won't change that. While those two guys are in the room together, this is not an absolute rule of course, but organizationally, that guy won't change – won't do Noun 2 operation optimization while that guy's in the room. He won't change those set points while he's there. But if he's out of the room, then it's his job to step up to make those changes. This is why I draw. Well, he tells him experience and knowledge, and he talks to him about ... "and so I'm leaving and I've got to go somewhere, have a look at that, if it goes up, put a bit of Noun 21 on. sort of stuff." it doesn't come back the other way. The guy doesn't sit there and say "ohh, I'm going to make this change without your permission, it doesn't happen like that. So then, that's probably why they're a little bit different. Organizationally, this guy is higher than that guy. He looks after the machines, equipment, primarily. This guy's after the operation.

Interviewer: do any of these systems get Information? We've got one Information flow here. Are there Os?

Participant: Well, we have all these replicant lines between the Noun 5 ones.

Interviewer: we've got all these data lines. And obviously these replace that... Are there any Information flows?

Participant: Yeah, there are. That's where we have. We almost have the data flowing away from the place we started.

Interviewer: this is a very useful discussion.

Participant: So we've got the Noun 14s just overlaid on the top. We have the Noun 9s... This is why I'm interested in your topic. This is real stuff for me. Wierdly important. And it's terrifyingly difficult to wrap your mind around. Especially when someone new starts in my sort of role. And they've just got to understand all of this.

Participant: We talked about this one, The Noun 2position 22. The Noun 2 Position 22. He is sending him information. It's weirdly enough, I'd raw that distinction between stuff. These might be stuff like daily targets. I want to make so many thousands of tons of Noun 23 today. I want to make 7 thousand tons today.

Interviewer: I really love living in a world where you can say that...

Participant: It's true. And I wanted it to be at this Noun 24 rate. And I wanted it to be at this chemistry. And I do draw the distinction here because...

Interviewer: so page 3 is overlaid on page 2?

Participant: It's sort of off to the side. I'm bringing the extra characters that are probably sitting up here, so maybe I could just ... do replicant boxes so we can just join them up together later as well. So they'll send them Information as daily targets. And we've got the Noun 1 down there. How do these guys interact with here? It's really weird to explain. And maybe I think I'm crazy. Now why did I describe that line here ... text advice and knowledge between here... We'll label that line AA. These guys influence the ability of these guys –

Interviewer: wait, the Noun 1 is getting the AA line?

Participant: It's just the same line. There is a line in here as well. maybe that's the best way to describe it. There's a loop in here which is – it's not a direct line ... that's not true either. That's definitely true. He sends this – The Position 22 sends that Knowledge into the Noun 1. He doesn't send a daily target, that's going this way.

Interviewer: He doesn't send that?

Participant: He does send it. He doesn't send it to me / he doesn't send it to the Noun 1. He just sends it straight to the ???

Interviewer: He's sending Knowledge to the Noun 1?

Participant: That's right. That's an SOP, for your purpose, SOP is standard operating procedure.

Interviewer: He's sending this as code?

Participant: No, as like a written document

Interviewer: To the Noun 1?

Participant: That is not true. He is sending it to me as the Noun 1 developer, – He sends me knowledge in the form of the SOP. And I send it back as Code. So that line is the same as that line. When you put them together, you start seeing that commonality

Interviewer: That would be BB, and this would be BB.

Participant: So that's true. But the point was – it's almost like: Can you describe the strength of data as well? As part of the philosophy?

Interviewer: Sure! but what are we talking about?

Participant: The ability for this guy to accept my Knowledge is influenced by this person.

Interviewer: Say again?

Participant: The ability of the Noun 14 to accept textual knowledge from the Noun 1 is influenced by the Noun 2 Position 22.

Interviewer: Where is that influence? Here? Or is that influence not on this diagram?

Participant: That's what I was trying to...

Interviewer: That's what I couldn't quite, I was drawing the lines and I got a bit technical.

Participant: Maybe, it's best to describe now in terms of: "he's sending him" as I see it, "knowledge." which is really Knowledge about how to use the Noun 1. Because, you remember, I'm sending it all the time. It never stops. Like the mail. It keeps coming. He's his own person as well. But this person reports sort of vaguely to that person, and this person definitely does and [Name] tells these guys – he influences as a form of Knowledge about what to do with my system. Whether you listen to it or you don't listen to it. Whether you pay attention to it...

Interviewer: Because there are certain times when you should and certain times when you shouldn't? according to him?

Participant: According to him, yes, that's right. – He really is sending them Knowledge. Because it, again, to me is not just about information. that Data/Information/Knowledge hierarchy – He's telling people how to use the Noun 1 based on his own experience, based in some sort of context. And there is a group 130 context as well. [Name] doesn't like that, don't do that. ... There's no rational sense in anything. That's why it's Knowledge, it's not Information. Because there's a context around it.

Participant: So that's the Noun 2 Position 22. So there's a little loop there. And that's true, because they won't send me – so the Noun 2 Position 22 to the Noun 1 developer. He won't send me much else apart from the procedure. Because this guy writes a procedure and says "These guys must follow that procedure." And he's also – we must acknowledge that he is sending them Knowledge in the form of SOP. That's the CC line, same line. I just get a copy of it... and I reinterpret

it. And even send it back to them. Which is apparently one of the benefits of – perceived benefits of these systems. Maybe I just – I haven't got a good context of a system noun 75 within there. Similarly and functionally, these guys don't talk to these guys because these guys report to someone else who is the Noun 2 position 25. He doesn't appear in any of these diagrams. Because he's more of a <grunt> what are you doing? <Grunt> sort of relationship. But similarly, but as the chains get a little bit weaker as they are going up we have ... And here we're actually progressing weirdly enough, we seem to be progressing through the hierarchy of these organizations. So now, maybe it's best to describe it as a hierarchy. as the Noun 2Position 22, we've got the Noun 2 position 26. We've got feeding through... we haven't quite covered all the boxes, but

Interviewer: we don't need to,

Participant: I might just... because it's always about me... Noun 1 developer here. So we've acknowledged the line between the Noun 2 Position 22 and the Noun 14 which is the CC line, the DD line, the EE line.

Interviewer: Since we're labeling, we have those three lines as a set? FF... because Noun 5 also sends FF to the Noun 9...

Participant: That's right. So these guys send Knowledge to these guys – the Noun 2 position 26 will send Knowledge to the Noun 2position 22. I'd say actually this is is where I don't know ... it certainly does influence the Noun 1. Maybe I'm just trying to go back through the chain as I see it. This is just my personal opinion. It would be great to see what they think is different. As it's getting further away from me, it's getting more chunked up and not even true anymore. Well, they'll send the Knowledge back such as like, production conditions and like monthly targets, "We want to make so much Noun 23 month" and these guys will say "we want to make so much Noun 23 per day." And these guys will send Information back.

Interviewer: Information or Knowledge?

Participant: I used it, and I meant it. Info such as getting pretty general there. Noun 2 noun 97 condition.

Interviewer: what is that?

Participant: Maybe a health? How healthy is it? Like a person. It's weirdly enough, because its... because I did say Information, because there must be something special about that. because I perceive this as – Those guys send those guys Information. They'll say, we're going OK. And may not have that much contextualized nature. It may be a graph and they'll sit there and look at the review at last week's worth of Information as a time-series graph. They'll say, "We're hitting our targets" weirdly enough I didn't use Knowledge, I used Information. Noun 2 noun 97 conditions like how are we traveling? Here is some graphs which describe. Traveling is a colloquial term of how the Noun 2 is performing. Because if the Noun 2 is not performing well, and the Noun

2 is down here somewhere, it can effect those guys. The production conditions and the monthly target it. Those guys might acknowledge and will here rightly acknowledge problems in production, or in the Noun 2 will affect his monthly target. So it will be the other loop which is getting really hairy for me up to the Noun 23 Making position 41, he's like "Yearly targets." we're making X million tons a year. That's sort of, we're making X million tons this year. We'll make that target. ... But I didn't put the position 41 of Noun 23 making on here, because he does – This guy sends me as the Noun 1 developer, he really sends me Knowledge in terms of like a philosophy. We have regular meetings with this guy [Name.] We'll sit down – we had a showdown meeting just before Christmas. And as the Noun 1 developer and him we sit down and he expounds his philosophy on what we should – I should be doing as the Noun 1 developer. He's definitely sending me Knowledge about what he thinks should be happening. It is a philosophy saying, "you're going about this the right way. We want to look at this aspect" of it. To be honest, I don't send anything back. It's the hierarchical nature of this ... I just take it like a man. My project got shut down over Christmas. He decided, he expounded the philosophy that we don't want to do this anymore. He sent me an e-mail one month later, saying "how are you going with the project? Are you completing that work?" I wrote back: "Currently it is not a priority and I will do it when I have time." Dude, man, you told me to stop doing it and then he's asking me why I'm not doing it? As a hierarchy, he's there and he's above me in the organization. He tells me what he wants and I'm there to implement a vision. That actually draws an upper limit boundary for us!

Participant: Wierdly enough, if you'd describe sort of from there to there, to here, we sort of captured that – the business chain of this? It's really the business side of things. Now, we haven't delved into the other side. It really is the operations perspective. Which is a funny one for me to do. In terms of the business , it's where the money's at. It's the important one as far as the business . I wanted to describe it how I developed the Noun 1. Notice I've left out the Noun 2 position 27. He hasn't appeared in here because he's not part of that business chain in terms of operation optimization. But he's very important in terms of rule development. Because he'll have a big influence over how I – on another section of the diagram which we haven't filled out which is maybe on, as you look at the page, on the right hand side with Noun 5 at the bottom, there's a little... He's over here and there's a whole other side which we haven't dealt with which is – and maybe it's weird that I've drawn that distinction. That I've traveled off one side of this diagram. There's a whole other side here which we haven't dealt with which is Noun 2 position 27 and Position 7 including other position 56s. ...

Interviewer: Before we stop, can you give me three examples (not on here) of your interactions, your data interactions, your Information interactions and your Knowledge interactions with your fellow position 56s? Just a little fluff...

Participant: Data interactions: well I sometimes get tidbits of Information (haha). That I send onto them. I'll find out about that there's a spreadsheet somewhere that there's a spreadsheet full of data that I might choose to share with people. So these guys have. People keep their own little personal records. I've been working on a task force about Noun 23 and noun 28, and there's a whole record about the Noun 23 levels in noun 28, which isn't recorded in any database, but which some guy with a spreadsheet who works for another company... And he sent it to me and now I've got it. And all of us have these little bits and we share it. "Have you got any data on that? Yeah I've got a little bit of data." that's the data side of things.

Participant: Information side, We certainly share reports and things like that. It probably comes to presentations and reports and things like that. I often describe them as Information that we'll share. Someone will say look at the data about the noun 29. "Who will look at the data? Who will write a report?" To me that will contain Information about stuff that I can learn about. Not necessarily Knowledge, because maybe I don't have the context.

Participant: And the Knowledge component, I see it with, as my other position 56s, it's that true sort of sharing of experience and things like that. It won't happen in reports. It'll happen in discussions and conversation. Maybe after a presentation, someone will present some information. This is this, this is that. And then we'll sort of flip into a Knowledge sharing mode – there will be a bit of back and forth about clarification or context which will help then build your Knowledge. Which you can go away with because Information itself is OK, it doesn't... it's for now. But Knowledge, i guess, is something that's more general. It needs a bit of context. That's true now, but if I know the context I'll be able to generalize it for the future, which then will make it Knowledge.

Interviewer: and so you would say a report's true now?

Participant: Yeah a report's Information because it's true now. If you looked at that same report a year later, I don't necessarily know its true anymore, because conditions might have changed which aren't documented within that thing. We try to document our assumptions or conditions, but it's not terribly important to us. Mostly we work like that and we look at it and we say that that was true then. And we go on and say: "is it true now?" Sometimes we don't even do that. We have a report and "are we going to confirm that's true? No, we'll just ignore the report. It's just Information. It's not true anymore. we can just dismiss it. And we're just going to do something now."

Participant: Data would be like values and bits and pieces to share with people. Information is like "we can share reports" again derived that often... sometimes...

Interviewer: Time series stuff?

Participant: Yeah. And then the Knowledge part we share would be the conversations, the contextual stuff: "Yeah, you need to think about this, and yeah, that was happening at the time" and things like that.

7.7 Interview 7
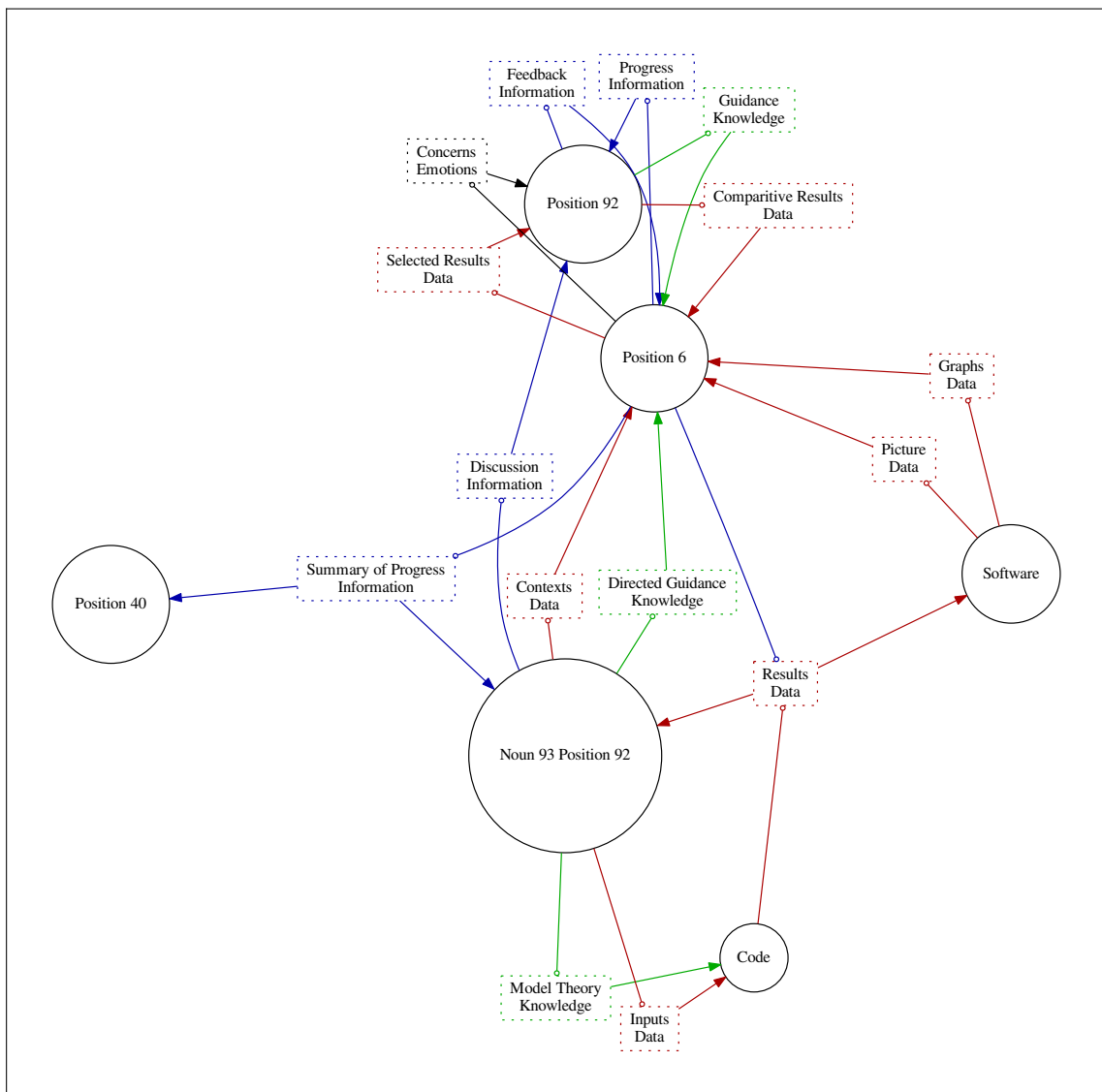
Figure 7.9: The SDFN Diagram for Interview 7
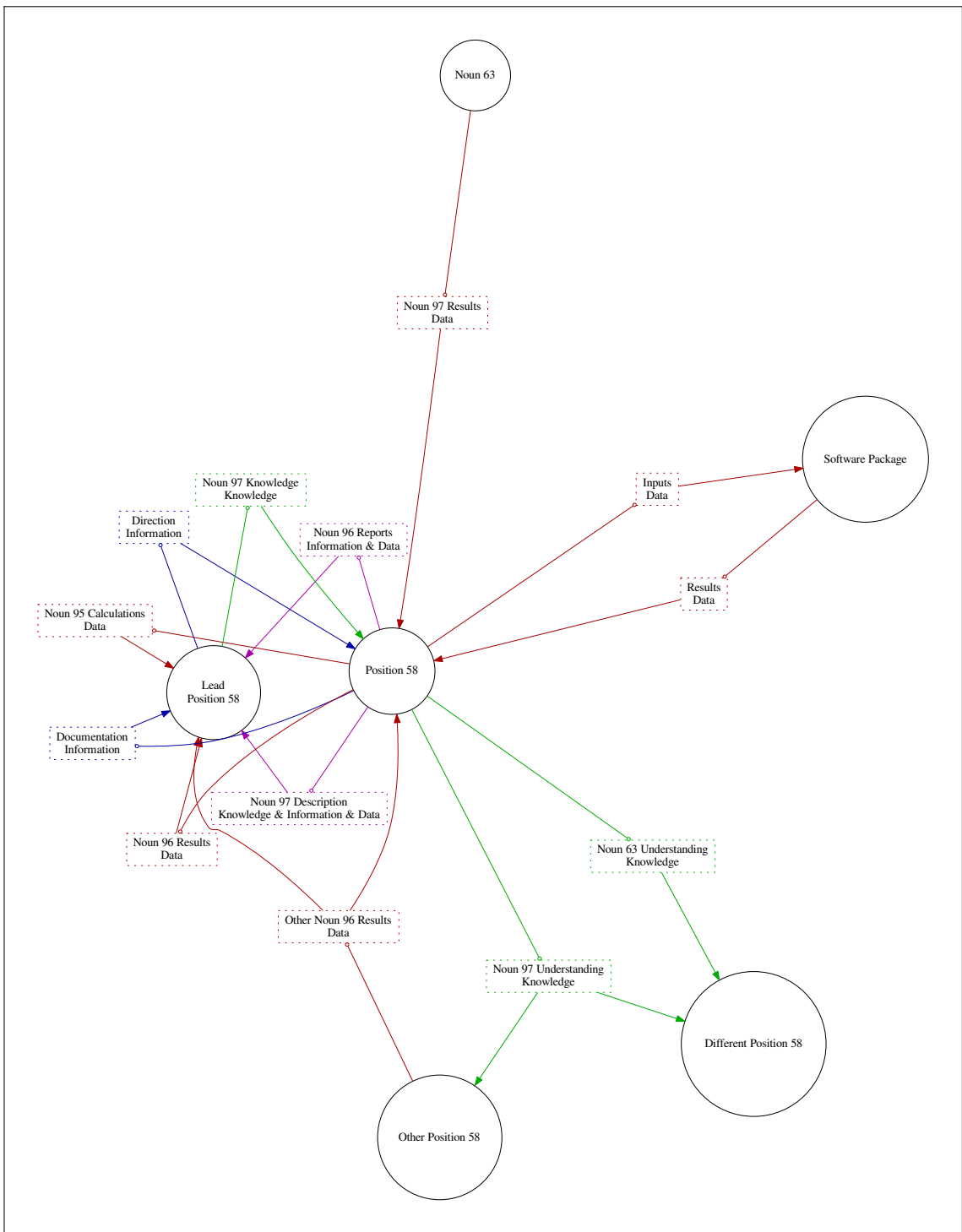
Figure 7.10:   The SDFN Diagram for Interview 7

Interviewer: ... I think a good way to segue into the entity diagram is
to tell me what project you think you handle – that are data-important

or that are knowledge important.

Participant: Do you want me to distinguish the two?

Interviewer: Feel free to...

Participant: Clearly data is important – not clearly. Data is important in a project such as the one that [Name] is working on. Data is also important in a project such as [Name] is working on. In both cases I'm not talking about any project I'm involved with. Data is also important in a project that I'm involved with which is related to changing the Noun 30 Noun 21 blend in the noun 60. Now you asked about knowledge, and information, as well?

Participant: Information is important in each of those. In my definition of knowledge.

Interviewer: and knowledge as well?

Participant: Oh yeah, yeah.

Interviewer: So let's do the one you're involved in to start with, and if we have time, lets do your view and your role in the one that's [name]'s involved in. But let's do the one you're involved in. And you mentioned a mathematical model.

Participant: That's what I was involved with, yeah. It won't impact on this one though.

Interviewer: Hopefully if we have time after the modelling, I'd love to talk about how you use data to inform and to create and then to use inside the mathematical model. Just to have that simmering in the background. Let's start with an entity dictionary.

Interviewer: How would you characterize yourself? Here, what we're looking at is not just you as you, but each of the roles you're in. So, for example, for me, at the university, one of my roles would be noun 91. student, and that would be separate from teaching assistant. So that's the kind of roles that I think are best to start with when we're working on our entity dictionary. ...

Interviewer: We begin by looking at a trivial flow between something and something else in the project you're participating in.

Participant: The project involves noun 35 with an external organization. So we have an entity called the noun 36. I can name it if it helps you.

Interviewer: Not really. So, what stuff is flowing to and from these guys?

Participant: noun 37. Noun 21.

...

Participant: So that's a physical noun 37. They also send noun 37 characteristics, so that's data. noun 37 characteristics or properties. Chemical properties if you want.

Interviewer: would it be better to say noun 37 characteristics?

Participant: Characteristics is probably sufficient.

Interviewer: This is Data?

Participant: That's Data.

Interviewer: Would you characterize the physical noun 37 as Data, Information, Knowledge, or just?

Participant: It's the means, really. It's not data. It's not Knowledge. They're just physical noun 37.

Interviewer: So we have our initial flows. What other flows are there?

Participant: So here's the position 32 over here, that's me. We haven't split – this is the research organization, so there are people within this organization – well, there's probably one person if you want who's the position 100 here. So you could actually say "position 40" So, flow from here to here is the noun 42 experiment. Over here we have position 39 who help to define the noun 42.

Interviewer: so what do they provide?

Participant: They provide information to assist in noun 95ing the noun 42.

Interviewer: So the noun 42 is information? or Knowledge?

Participant: It's information. It's information provided to the position 41 here to allow that person to undertake test work of a specific nature. It's guidance into what it is they're going to test. We've sent them noun 37 and a noun 42.

Interviewer: How can we label this flow from position 39 to you of information? So you're sending them a noun 42, and the position 39 are sending you Information to assist in the noun 42. Do we label it as Information to assist in noun 42?

Participant: I think information to help noun 95 noun 42.

Interviewer: What else?

Participant: Within here there's also some work going on in terms of communication between this guy and their team. position 43. Information, noun 37, across there. Information,

Interviewer: Information of the noun 42? Is this modulated by the position 41? Or just distributed?

Participant: It's likely to be modulated. They're likely to have their own standard operating procedures.

Interviewer: How do we want to label this differently than noun 42 to indicate that modulation?

Participant: Along the lines of noun 42 conducted according to SOPs.

Interviewer: And they're sending Information, and they're also sending noun 37? Which are just stuff?

Participant: a lot of test work goes on in here. A lot of analysis goes on.

Interviewer: Do you know or care about the instrumentation they use for analysis or verb 73? Or is it just a blackbox thing?

Participant: We do care. In part, that's covered in the noun 42. Is the question about: do we care about what sort of instrument they are using or the type of tests that they are doing?

Interviewer: Both. Because. We can treat this as a black box. And stuff will flow from this black box to these. Or we can treat this as part of your understanding of the project, at which point we're going to label their external manipulations of your noun 37 according to the SOP.

Participant: I'll explore one area. So these are physical noun 37. They actually have to manipulate those physical noun 37. They actually have to take and crush those noun 37 and prepare the physical noun 37. These are physical noun 37 from our noun 60... In order for the experimental team to do their experiments, they need to actually do some further ... and classification of the noun 37. The way they do that depends on the noun 42.

Interviewer: So, is there an entity – some technological device that is a useful catch-all or a useful...

Participant: Let's call it a properly sized Noun 21 noun 37

Interviewer: So they're sending properly sized Noun 21 noun 37?

Participant: No. They receive the physical noun 37, then they have to generate this properly sized Noun 21 noun 37. So they have to do some crushing and sizing.

Interviewer: So this would be a noun 44? So this goes physical noun 37 in, and out comes properly sized and crushed... where does that go?

Participant: It goes into an experimental noun 45. Crushed and sized Noun 21.

Interviewer: Are there any stuff flows, Information, Data, going along this route?

Participant: No, because it's being done by this group.

Interviewer: So they're not setting any settings on the noun 44.

Participant: Yes they are. That setting will be done to be consistent with the noun 42 that's been provided.

Interviewer: Are they setting the setting? Here's the tricky bit. That setting of the setting, i.e. the transmission of the settings into the machine, is that Information or is that Data? Or is that something else?

Participant: Well, it's data from us, but it's Information into this machine, isn't it? We tell them – this was Information... Information can be Data as well, right? You don't want so say?

Interviewer: The thing is, I can't tell you what you think.

Participant: The noun 42 has Data on the size distribution. So that's a set size distribution. So that's Data that we've provided the people so that they're informed so as to how to set.

Interviewer: So, in the Information of the noun 42, which is compiled of multiple protocols

Participant: this is one of them

Interviewer: This is one of them, and you would say this individual protocol is Data? So in this packet of Information it goes through the position 41, the position 41 adds SOP, and that packet plus SOP is still Information. Do they perform something interesting to this packet to extract out – or can it even be extracted out – the settings for this noun 44?

Participant: No, they just take Information and make the setting.

Interviewer: So that would be transmitting information of size settings. – Now, they've got an experimental noun 45. How are they interacting with the noun 45?

Participant: They're running the noun 45.

Interviewer: Are they transmitting stuff to the noun 45? or is it

Participant: Are they transmitting stuff?

Interviewer: Are they transmitting components of the tech plan to the noun 45?

Participant: No, because that's what the SOP is about. The SOP guides them as to how they should run that noun 45.

Interviewer: So there's no data flow or Information flow into the noun 45.

Participant: No.

Interviewer: What happens from the noun 45?

Participant: Produce new noun 37 of noun 46 which need to be analyzed. It goes to a Laboratory.

Interviewer: And we've got new noun 37 of noun 46 to the lab: Who follow standard procedures as transmitted from him or them?

Participant: Probably by him.

Interviewer: So we've got SOP. And these SOP are Knowledge, Information...

Participant: Also, according to the noun 42

Interviewer: but it's not that? Is it this flow?

Participant: In the noun 42, there is Information as to what lab tests are to be done. It's like that. You could take it out of there if you wanted to?

Interviewer: What reflects reality best?

Participant: The reality is that it probably... as well as SOP, it would be noun 42. In fact, it's a bit similar to what you've got there. noun 42 conducted...

Interviewer: Is the noun 42 conducted also Information when it's sent over here? So noun 42 conducted according to SOP goes to the lab along with noun 46 noun 37 which have been prepared through all that. What happens then?

Participant: Then Data from the lab comes back through to this person here.

Interviewer: So this is Data. What is this Data?

Participant: The Data is the results of laboratory testing of noun 46 noun 37. I'm being Generic, if that helps you. I could be more specific about what exactly the Data is, – I'm happy to do that.

Interviewer: I would love an artifact... like a verb 73 of the results? But unless you think that it explains what you think Data is more, I don't think we need to go into any more detail.

Participant: I think the emphasis is that it is a result from the test.

Interviewer: Because that's the second time you've used Data in this, which is why it is significant.

Participant: So this person prepares reports, and then the reports come back to the position 32.

Interviewer: so you get reports on the results

Participant: Correct. Via e-mail.

Interviewer: is this significant?

Participant: No.

Interviewer: It's interesting. Some people would say that e-mail is significant and that it actually changes the way the other people go –it's e-mail. So reports on lab results. Is this Data Information Knowledge, O?

Participant: At that point it's Data. They are providing us Data. They haven't made any – Even in context of comparing other results, they would have compared Data from previous experiments that they've done with ours.

Interviewer: And are these comparisons also Data, or are they privileged in some way?

Participant: I would have thought that would have been Information which they are not privy to disclosing to us anyway. As far as I'm concerned... we get the Data back.

Interviewer: Are there any other entities or steps we haven't described?

Participant: Here there's another one. Then we have the discussion internally, with position 39. Obviously, we each, individually analyze

the data ourselves and then we do an excel spreadsheet analysis, if you want.

Interviewer: Let's stop there, let's diagram that. How can we diagram that? We've got this line going to the position 39, what's this first – do you perform any operations on these reports before sending them to the position 39?

Participant: I do. I will probably summarize key points in an excel spreadsheet. Differences between successive tests.

Interviewer: What do we want to label that, and is that Data as well?

Participant: It's certainly data. "Results Analysis" Summary, I guess, of results analysis – Results Summary. I guess I use the word Analysis there as well. ... and Analysis.

Interviewer: Then what happens?

Participant: Then we have our discussion.

Interviewer: How can we represent this discussion? Should we represent this discussion?

Participant: No. There's all sorts of discussions going on elsewhere as well, but we haven't presented that necessarily. You can leave it as that. And then, there's a bit of recycle loop. This whole noun 97 then out of discussion are conclusions based on the current set of results and then we may well review the noun 42 and start this all again.

Interviewer: So they send conclusions to you?

Participant: No no.

Interviewer: you all generate conclusions

Participant: Yeah.

Interviewer: How do we want to represent these conclusions, or do we? Or is that stream the conclusions, from the earlier iteration?

Participant: Yeah, that's fine. This was specific to the noun 42.

Interviewer: So we add another stream going here...

Participant: well, if you want, you can leave it as it is, then this discussion. The entity is that the position 39 – the position 39 and myself get together – therefore that stream ....

Interviewer: so we have position 32 sends position 39 and position 32 a stream of Data which is results summary and analysis.

Participant: Therefore, we can go from here with an updated noun 42. From then on we can follow the loop.

Interviewer: And the updated noun 42 is?

Participant: It's Data, Sorry it's Information. I mean, It is both, right? I'm providing data, but it's primarily information.

Interviewer: Unpack that for me.

Participant: So I'm providing numbers. So the numbers, in isolation, mean virtually nothing. So attached to the data is some Information. Particularly in this feedback loop, I would present, we said we were going to crush the Noun 21s to this specification, but based on the current results, we'll change it, so here's a new setting which we still refer to as Information anyway.

Interviewer: is that correct?

Participant: I think so. In making that judgement, or in providing that Data, we are informing them of why we are changing the specifications, changing that Data.

Interviewer: In a sense, this size setting is composed of numbers plus ... explanation?

Participant: Almost Intellectual Property, some explanation.

Interviewer: IP is a good word. We've got numbers plus IP. And you'd describe the numbers as?

Participant: As the data.

Interviewer: And you'd describe the IP as the Information. And that entire packet of data + Information, of numbers + IP, that entire packet is also Information.

Interviewer: Where does Knowledge happen, is there any Knowledge transmission in this whole noun 97?

Participant: Yeah. The Knowledge transmission probably takes place in this discussion.

Interviewer: in the discussion of the position 39 and the position 32, what – where's the Knowledge interaction there?

Participant: Where is it, or how does it...?

Interviewer: All of the above.

Participant: So it's a meaning??? Data's on the table. Information is exchanged: what do I know, what do you know? What does the Data, what do the results mean?

Interviewer: And that meaning of results is... Information?

Participant: The meaning of results is Information. But, that Information may be new. Or that's an unexpected result. There is potential for Knowledge, for new Knowledge to be generated.

Interviewer: And the new Knowledge which is generated comes from

Participant: From the analysis of the results and the discussion of the results and the implications that are derived from the analysis of those results. Am I confusing you?

Participant: Data, Information, and Knowledge. In my books, you analyze data. And generate new Information. New Information can generate a new... So, I have a model in my head, and this is what a noun 46 should be. And that model is Knowledge. That model could be shifted

slightly. A shift in some model (physical or whatever) is due to new Knowledge.

Participant: Within that discussion

Interviewer: we have that iteration. I need to figure out how to represent that iteration. In here we've got Knowledge being generated, does it go anywhere?

Participant: For a start, it comes back – the Knowledge is then transmitted back via information to these guys. We won't provide that to them. There may be a little bit, but you're not explicitly... It's in the form of Information returned to the external organization via an external noun 42

Interviewer: By them analyzing the updated noun 42 they see their own new Information, which means they can generate their own Knowledge?

Participant: No, I don't think they'd do that necessarily. The distinction here is these are research position 41... they're an external service provider. If you have a contract for wanting you to do some work on your home for example, a plumber, you won't necessarily tell them about what it is that you're going to do with this device, you just tell them I want that fixed, and I want that service provided. and that's whats happening here. We're asking them to perform tests and asking them to provide us Data. We'll assess the data, there's no interaction with the position 43. They're just providing that service.

Interviewer: Therefore do you sink your Knowledge somewhere else here so that other people can take it?

Participant: Yes.

Interviewer: So let's name that.

Participant: We've named these position 39 but we could probably name them more specifically as position 47 because over here because the next offshoot here, the other entity is the position 48.

Interviewer: You, to them, send...

Participant: Information. We provide them with Information, but they might ask for explanation as well, so therefore we give them Knowledge.

Interviewer: First step is you're sending them Information. You're sending them what Information?

Participant: It will be something – a summary of the summary that has just been produced. It could be a presentation. Executive summary if you want.

Participant: now, I put in this category, generally speaking what happens is when you get to this point you will have different levels within the organization. For example, you might have a general position 41 of the operations, a position 41 of the operations, and then a... So what

we refer to as level 4, level 3 and level 2 position 41s. Let's call them position 41s for a moment, I've clumped them all together under position 49.

Interviewer: So you send to these level 4,3,2 position 41s...

Participant: We would probably send them Information first.The alternative ... we may send them Information. More than likely we'll sit down with them and present the Information.

Interviewer: So there's some communication of Information, and then what?

Participant: Then that Information is discussed in a meeting, not that meeting, with these guys. And decisions are made as a result. The decision might be to continue the noun 35, continue the noun 42

b; So they send back to you ...

Participant: Remember I said we'd probably do it as a meeting

Interviewer: so this meeting sends back to you...

Participant: In here position 49 + position 47

Interviewer: and this dovetailed flow goes back to?

Participant: Ourselves.

Interviewer: And this dovetailed flow is the recommendations?

Participant: Recommendations, yeah.

Interviewer: And these recommendations are?

Participant: Recommendations of Information via definition...

Interviewer: So these recommendations are information. Do you transmit Knowledge from this meeting – the position 47 and yourself to the position 49? How can we render this Knowledge transfer? So you're sending Knowledge...

Participant: It's that IP. It's like this. You say here is new Information, here is what we interpret as the new Knowledge, we're providing that new Knowledge.

Interviewer: So we could say that this is the IP... What do they do with the executive summary plus IP besides send back recommendations to you regarding the experiment?

Participant: So these guys in this project would then – these are position 49 people, they would communicate that Information, these people may well be in the same meeting, by the way, but let's assume they're not, to noun 51 department.

Participant: Because what this is about is the purchasing of alternate Noun 21 noun 51.

Interviewer: so position 49 are sending Information

Participant: Information? So this would be a recommendation to verb 52 another Noun 21. It is Information. And these guys would then talk

to by phone or whatever, transmit in some way, that information to the Noun 21 noun 51ers. verb 52 and negotiate noun 53 and that sort of thing.

Interviewer: and that's just an Information flow? What do they do?

Participant: They will send stuff back.

Interviewer: Do you care about the stuff they send back?

Participant: Yes, we do. They will send specifications back to these guys here.

Interviewer: And these specifications?

Participant: Noun 21 specifications

Interviewer: and these Noun 21 specifications are?

Participant: Data.

Interviewer: And the noun 51 department then

Participant: Then they would provide that noun 51 to the position 49

Interviewer: Are they modulating this or are they just passing it on?

Participant: They're just passing it on. And that's where it sits, because these are the guys that actually...

Interviewer: What else? Are there any other flows in this noun 97?

Participant: There are only other flows over here, but they're outside... Probably there is, but that looks pretty good for me.

...

Interviewer: This diagram is your how does one turn into the O, which means that you're stating that one can turn into the other. Which is important because not everyone believes that.

Participant: Mind you, having gone through that, I was thinking all along: this is no different to any other activity you might end up doing anyway. The fact is that there is an position 43, but in the end it could be anybody.

Interviewer: It's just someone who applies the SOP. they're a clearing-house.

...

Interviewer: Let's spend this time chatting about your philosophies of Data, Information, and Knowledge. Let's begin by going: "Can you give me your own definition of these three terms?"

Participant: Of Data? I guess I was sort of doing it here, in a way. It's what you do the Data... What is Data there? Data in itself, if you look at numbers on a sheet of paper, multidimensional, these are numbers in this case. An example using just numbers. Looking at that assemblage of Data points, it's hard to come up with some mental model. Particularly if it's multifactorial.

Interviewer: So one of these is a data point?

Participant: All of them are

Interviewer: Each of them is a data point and collectively they are data?

Participant: Yeah. But it's the interpretation of the data which is required. That can be done by your own mind, but it can be done in a more systematic way: steadfastly, there is tools. You have excel tools, excel software. You have other statistical analysis tools.

Interviewer: So data + analysis through tools creates Information?

Participant: Creates Information, yeah. Ultimately, where this is going to end up is, I've got a model in my head: this causes that, for example. I'm looking at the Data and I'm analyzing the Data, and I want to understand whether the Data indicates that that is true or not. That there is a cause and an effect, that there is is a...

Interviewer: So the Data is informing Knowledge of causal relationships in your model?

Participant: Yeah. One of these may be a dependent variable, and the rest might be independent variables. I'm trying to understand if these dependencies are there?

Interviewer: And this understanding is Knowledge?

Participant: That's part of the Knowledge, yes.

Interviewer: And so we've got analysis... and so the analyzed Information in this... Where does Information fall in this causal... ?

Participant: The Information is almost an extraction ... Data + analysis provides Information. and Information then can be used to generate new Knowledge.

Interviewer: Information + what? Is there a plus?

Participant: Well, I was talking about models here, so Information + model....

Interviewer: And a model is?

Participant: A model could be anything. It could be an equation, it could be a physical model.

Interviewer: and the class the model falls in is Information or is it Knowledge?

Participant: It's probably old Knowledge in this definition.

Interviewer: And so, in this definition, Information which is the combined Data + analysis of these data points, which are numbers.

Participant: They could be numbers they could be... You could also.... Data + analysis, but over here it could be Contextual Information. There may be things around... you could have a set of data, two sets of data, and the difference between those two sets of data is the constraints

that were imposed by group 130 or by some other entity, and that's the explanation. Data is first, then constraints, analysis, and so on.

Interviewer: And these constraints and this analysis are? Analysis is really a verb. These constraints and analysis are data?

Participant: No, constraints they're... other data. I jumped to that because you were talking about the data points. So that's just the data points. The constraints are, you've got this objective function, there are relationships that define X = Y, help to define that. But then bounding that are constraints. Maybe a way of thinking about it is: if you're trying to optimize production rates or something like that, you could say OK, production rate = A+BC, but you know that C is bounded by something else, by Data. Therefore, it isn't just the data, it's Data + Constraints.

Interviewer: And these constraints are other data. You mentioned a very important word to me: relationships. Relationships are?

Participant: A relationship may such as that?

Interviewer: So this function is a relationship?

Participant: We call it ... you're distinguishing mathematically now?

Interviewer: What I'm trying to get at: do you classify a relationship as Data, Information, or Knowledge?

Participant: I think a relationship, like that, is probably Knowledge.

Interviewer: And that would be the model, basically.

Participant: Yes.

Interviewer: you mentioned context?

Participant: So the context is that you could have a relationship like that unbounded, but the context is that's knowing what the constraints are.

Interviewer: So context is the application of constraints.

Participant: Constraints, yeah. That's certainly one way to interpret context in this example.

Interviewer: This application of constraints as context is also Knowledge? Or is it...

Participant: The application of constraints? I think that's probably Information. It's either Data or Information. It's not Knowledge.

Interviewer: Can we have arrows going the other direction here? We have this new Knowledge from this analysis noun 97. What happens then?

Participant: What happens back this way? New Knowledge is always, or should be, communicated and then – on the practical side of it is: communicated, argued, and then agreed or discarded.

Interviewer: So you've got the Knowledge which is bound into or discarded from your knowledge-of-world, since we can't have impressive

German terms.... At some point in this chain, you issue normative orders: "You should do this." Those normative orders, here, are function of your noun 42?

Participant: They're represented by the noun 42, and they're represented by that Information.

Interviewer: Your classification of these orders, these "You should do something" is it part of this model, or is it part of a different model?

Participant: I see what you're getting back to this... which is a flat... I think it's, in practice, I think it's separate. particularly in practice –

Participant: Let's say we've got scientific or technical Knowledge. And think of it back over here, it's sort of like technical Knowledge is sort of in here.

Interviewer: Technical Knowledge is where?

Participant: See, you have learnt new things. Sorry, these are results plus summary and analysis. You're generating, performing

Interviewer: You're performing the data + analysis into Information into Knowledge.

Participant: So, now, it's expressed here as IP. I guess in that case it is coming back up, it's Information. So, in effect, we described it as Information.

Interviewer: We've described that Knowledge generation noun 97...

Participant: has provided new Information to be provided to the noun 60 position 90s.

Interviewer: and they can then take that information and put it into their own cycle?

Participant: Yeah. They can. In fact I'm sure it happens like that, too. The reason I'm ??? is that there's probably a filter here before it gets to noun 60 position 90s.

Interviewer: So we've got Technical Knowledge, and this Technical Knowledge is this IP? No.

Participant: It's that IP. Put through a filter, so that these people will understand better.

Interviewer: This filter is?

Participant: What it is is the executive summary. What becomes is – that's the executive summary what it is over here a detailed summary, if you want, but it's more than that.

Interviewer: Detailed summary and stuff. And then it's winnowed into this executive summary.

Participant: It's sort of like the questions you want – you'll have some questions in your mind that these people will want answered. so they've got objectives: do I, or don't I change Noun 21 noun 51? That's the question. You're anticipating that question, so therefore you generate

216

a summary of or a distillation of this detailed stuff to present to them to address that.

Interviewer: that distillation is?

Participant: Is – over here we've got information. And we've said IP, in part. but over here the distillation is the application of – no, it's quite separate. So you've got new Knowledge, which is then put through a filter which is – its nothing to do with the technical. it's more about, let's call it, a social. It's non-technical. It's about how to explain to you what I know? So I will try to put it into ... I'll try to work out what you're likely to understand, so I'm thinking: "Well, he's not a social scientist, he's a IST person, so therefore he'll understand some noun 109. So I'll put some maths or a relationship there. But if you were a salesperson with a marketing degree, probably forget about that." The filter is a verb 112 of – it's a device that permits more effective transmission of information between one entity and another. And in this case it's probably a filter around "What will that person understand?"

Interviewer: I think one way of talking about that filter is a local language. That we evolve that we know, that when we're speaking to each other that we have these specific terms that are locally true but not globally true.

Participant: It's certainly jargon, if you want to put it that way. It's more than that, it's... the way that Information can be communicated is based on level of people. Meaning time, availability. But it's also, what I was saying earlier, about Knowledge.

Interviewer: The Knowledge of the listener?

Participant: Of the listener. And their level, and Knowledge of the listener which includes their understanding level. The other aspect is just they usually. It's synthesis. It's the synthesizing of all the Information. It's the bottom line. If this, this, and this, are all true, what does it mean? What's the bottom line? It's being able to: that executive summary is being able to go that plus that plus that to get to the bottom line. This is the ultimate result. The ultimate piece of information that they require.

Interviewer: And that executive summary, as Information, is functionally a normative recommendation to change Noun 21 noun 51ers or not to change Noun 21 noun 51ers.

Participant: Yeah, it's a recommendation as opposed to a decision.

Interviewer: That distinction is important.

Participant: This doesn't necessarily provide – that's Knowledge, because it's the IP, but the executive summary is Information via recommendation. But that's really a decision.

Interviewer: And that decision is an order to change or nothing's sent.

Participant: That's right.

Interviewer: and you would say that this decision is Information in the same style that the executive summary is Information or the noun 42 is Information. Or is the difference sort of Information?

Participant: Well, the decision is almost really Data? In a way. It's Information, but it's also ... no, it's Information. I was thinking whether or not it was Data. you're just saying, it's yes or no. In that context "Do I or don't I?" and you're saying "Do", to the noun 51 person. It's 1, go.

Interviewer: And that Boolean decision is still Information?

Participant: I think that's Data, that's what I'm saying.

Interviewer: Perhaps this decision ...

Participant: It's informing the person to do something like the noun 51er to do it

Interviewer: It's a container for Data? So, perhaps you've got the D: 1/0 change, don't change. But it's transmitted in some sort of information?

Participant: Yes, yeah. It's yes or no, but behind the yes or no is this other Information. It's what is supporting the yes or the no. What Information is supporting that decision. A yes or a no.

Interviewer: In this other context, In order to undertake something new , say you're ordering someone to go do that research. An order is an example or instance of what?

Participant: An order is – you've made a decision

Interviewer: and you're transmitting that decision to me as the orderee. And that transmission of decision is? Is it on this line?

Participant: The transmission of the decision is there. It's Data.

Interviewer: Because it boils down to a 1 or a 0?

Participant: Yeah.

Interviewer: If you told me, Brian, go research that new IST system for CRM, that is also Data?

Participant: Yeah.

Interviewer: The go research component is? Constraint?

Participant: Yeah. It's probably Information. You're saying context of... It has to be Information, surely.

Interviewer: You have a pretty good sense of what I'm trying to tease out now. Do you have any parting words on ... conclusions I should come to in this analysis that might not be obvious from the diagrams

Participant: I'm not sure... I think the discussion about Data which is what you said there anyway, the Data, Information, Knowledge it's quite confronting in a way for me, because I ... but I think it's worthwhile for you to pursue it. You've probably held back your own views about what is and what isn't data. And what is and what isn't Information. I'd - I'm not sure. They are distinct. To a degree. At times, they will overlap.

I think what I've been struggling with, through this interview if you want, is how to separate them. To make them more distinct. But maybe there isn't a need for that. Maybe there are different flows here that are both Information and Data, and both Information and Knowledge.

Interviewer: Thank you for this time.

Participant: It was a pleasure, and I say that sincerely, actually. What I do appreciate is being challenged. Not necessarily by Brian, but what he was trying to do.

7.8   Interview 8

Interviewer: Let us do the training. Because we're short on time, I'm going to skip making an entity dictionary. Functionally, an entity is any person or thing or noun 97 that can manipulate Data, Information, and Knowledge in some interesting way. You, yourself, can be multiple entities depending on what your role is. When I'm at university for example, I can have the role of noun 91 researcher, but I can also have the role of Teaching Assistant. And they're different. My computer doesn't necessarily have a role. If it's acting like the pen in my hand, I don't think about the pen I'm using, I just write. But if I suddenly, go "wait a minute, no, the pen is important because it's doing some sort of transformation of what I'm doing, then suddenly it's an entity." We'll start with you diagramming a role that you play in this noun 97. And then we'll go from there to just a trivial flow of Data, Information, or Knowledge, and just build out. What would be a simple role that you play?

Participant: One is, of course, if you count the training, One would be putting together the material.

Interviewer: What could we label that role as?

Participant: I suppose its ... author? I'm lecturing.

Interviewer: We can call it author. Perhaps lecture author?

Participant: Everyone that has some training, they don't go and get a book. We should sometimes, it would be a lot easier and take less time. It seems like everyone, every time there's some training going on you start thinking about 'what are we supposed to be telling them?' what subject? What do I feel is important from a theoretical view? And also what's important to the noun 60? You have to look at both sides. From that I produce a lecturing material.

Interviewer: Let's model that. So we've got Lecture Author... now, you as Lecture Author, you tell me that you're producing lecture material. Let's start with ...

Participant: I usually base – I use Data a lot. Because I start with a theory and then I actually use noun 60 Data – that theory is not just something I came up with...
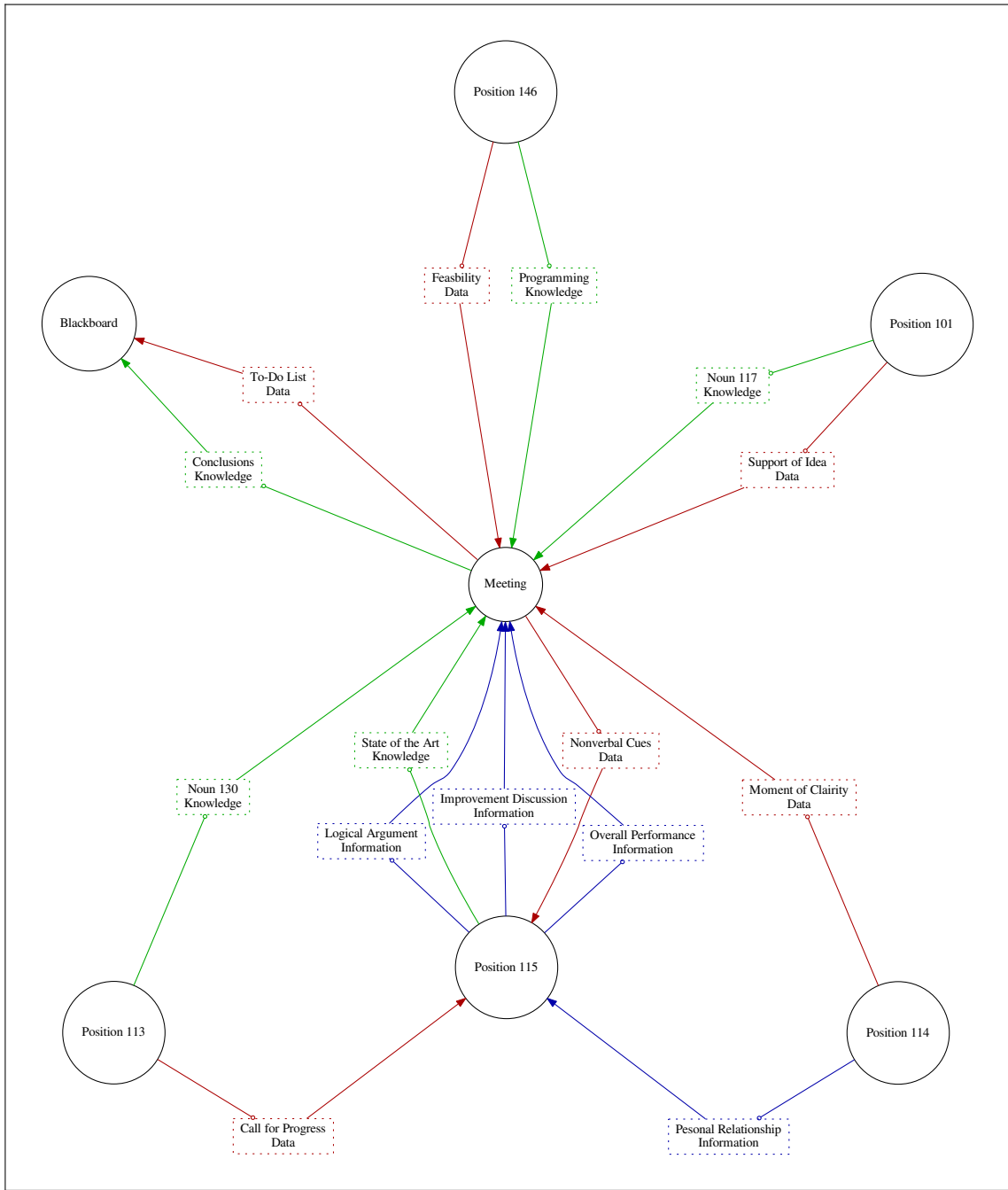
Figure 7.11:   The SDFN Diagram for Interview 8

Interviewer: Where do you get Data from?

Participant: From different production DBs so they're collecting –

Interviewer: Production DBs? So we'll call … production…

Participant: You're sort of really looking at numbers.

Interviewer: Are we looking at numbers?

Participant: No, we're also looking at relationships between Data. Which in the theoretical case would be coming from books. Fundamental chemistry, I suppose. Characteristics.

Interviewer: So we've got the production DB and we've got books. What do you communicate to the production DB to get Data back? Or do you?

Participant: I just do the ordinary thing. I just use our tools to extract our Data. Manually we sort of – that's what we have to do all the time. We don't have any kind of simple statistics. We normally don't have any problems. We just have them SQL script, and just extract with whatever tools our IST department noun 51. And it could be different programs for different DBs, because in a noun 60, we have sort of a mix of different types of DBs from old to new and depending on system levels too. We normally work with 3 levels. One planning level, noun 88, and we certainly have the level 1 which is the noun 60 sort of system, whatever that controls the noun 60. Level 1, and 2 is more ... calculation. And 3 is the planning. So you might get Data from all of them. And they all have different DBs.

Interviewer: Should we generalize your request to those DBs? Is it possible to put one request to any of those databases under one heading?

Participant: Yes, I suppose it is. We call that fact-finding or something. I wouldn't sort of treat them differently. Because they're sort of just numbers anyways.

Interviewer: This flow of fact-finding to the DBs, is it Data, Information, Knowledge or O?

Participant: It's just Data, numbers, from measurements.

Interviewer: So you, to the DBs, are sending Data?

Participant: No, I'm collecting Data. The equipment is...

Interviewer: We've got a flow from you to the DBs of Fact finding. What is that flow? Is it Data, Is it Information, is it Knowledge, or is it something else?

Participant: I'd just ask...

Interviewer: So you asking to retrieve – are you sending the computers, Data, Information, Knowledge, or something else? Or nothing?

Participant: Or nothing. no, no, I just ask for – I want a list of numbers.

Interviewer: So that statement of list, you don't categorize as Data, Information, or Knowledge?

Participant: No, I haven't. But I suppose it is Knowledge, because I put together the question encoded to my Knowledge. I would be sending that Knowledge, even if the DB doesn't acknowledge it.

Interviewer: You are sending it Knowledge because it's a representation of your Knowledge?

Participant: Yeah. I choose, I suppose, And expand beyond that, to extract the Data, think on a little bit more would be someone else doing the same job might extract some other Data

Interviewer: because they don't have the same Knowledge?

Participant: Yeah.

Interviewer: Now, you send a Knowledge flow of fact-finding to the DBs. What do you get back?

Participant: I would get back numbers, I suppose? Again the facts. Hopefully the facts and not fiction. I wouldn't know. I have to examine the Data. And that's sort of – when I retrieve the Data, you go through the Data, and you sort of quality control, I would think. Based on my Knowledge. That's no fun. The scrutiny of Data.

Interviewer: Let's set that aside for a second. You get back from the production DB Data, yes? It's not Information or Knowledge, right?

Participant: No, it's just Data.

Interviewer: The Data you get back, just like you're sending fact-finding flow. What do you call a flow you get back from a DB?

Participant: I suppose it's some Knowledge too, I suppose. Or results or ... It would be results. It's the result of something.

Interviewer: Of your fact-finding

Participant: Yeah.

Interviewer: This is what we're looking at. You as your lecture author, send fact-finding Knowledge flow to the production DB. They respond with a Data flow of results.

Participant: Yep.

Interviewer: Now you mentioned books. How can we build something like that to books?

Participant: The books would have a big part in the fact finding. In what Knowledge I would sort of send...

Interviewer: The first question is, if we're treating a set of books as an entity, is there a flow from you to the books?

Participant: At this point, I couldn't think of one. No, I wouldn't be able to tell the books anything.

Interviewer: Would you consider the selection of a particular book from a set of books a sense of communication from you to the entity. Or is it just access which isn't communication?

Participant: It's just access.

Interviewer: From the books to you, however, is there a flow?

Participant: Of Knowledge. A Knowledge flow.

Interviewer: so there's a Knowledge flow. This Knowledge flow, what can we label?

Participant: Theories. And I suppose Theories and experiences. Depending on what...

Interviewer: so the books can provide a Knowledge flow of experiences?

Participant: Yeah.

Interviewer: Is it separate from the Knowledge flow of theories?

Participant: No, it would be probably one. I think books would be either/or. If it's not something published where somebody did investigation. But if you look at printed books, they're either theories or both. Even in the books, they would use sometimes examples based on production data.

Interviewer: You as lecture author use books and the production database to generate your lecture. Do you use anything else to generate your lecture?

Participant: It would be more illustrations.

Interviewer: Where do you get the illustrations from?

Participant: It could be drawings... or if its equipment, that would be from our noun 89.

Interviewer: What entity should we label it as, noun 89?

Participant: Yeah.

Interviewer: now this noun 89, do you send it anything as part of this noun 97?

Participant: Not more than a request for a certain drawing. I just usually go to someone who has to print out the... I would ask...

Interviewer: What do you do? You ask the archive...

Participant: for a particular text.

Interviewer: how can we label this flow?

Participant: Drawing Request.

Interviewer: This drawing request, is it Data, Information, Knowledge, or O?

Participant: O, I suppose. Or it's the same with the Data. We'll ask for something specific because of what I want – I suppose it's a little bit of Knowledge in that. It's based on what I want to show so it might just be – I would relay when my word?? comes back, it's more like Information.

Interviewer: So you send a drawing request, which is an expression of expertise? i.e. Knowledge? Or is it something else?

Participant: Something else. I mean, I don't know anything about it.

Interviewer: you just say I want such and such.

Participant: Because I want just show something practical or have need or practical use of it together with some other data or some event or whatever.

Interviewer: So that request, if we're putting it in the other category, I need to label it something. What would you label it as, category wise?

Participant: So you have Knowledge,

Interviewer: Information, and Data.

Participant: I think I'm sending it Knowledge, I suppose, because I know something. I have Knowledge about something, that's why I need this particular drawing.

Interviewer: Because you have Knowledge, you're sending that Knowledge to them?

Participant: I never thought I was sending something more than the request for Information. Or get back, it's just – that's a little bit harder.

Interviewer: Let's have the flow back, what's the flow back from the drawings?

Participant: That's Information, yeah.

Interviewer: Now, what is this Information? What is the drawing archive sending you as Information?

Participant: what would you call a drawing?

Interviewer: You said illustration earlier. Would it be illustrations, would it be something else?

Participant: It would be a scaled drawing of some equipment.

Interviewer: And this drawing is Information?

Participant: Yeah.

Interviewer: But the drawing request is not Information?

Participant: No, I suppose it would be...

Interviewer: I'm not trying to put words in your mouth. Don't use the exclusion... don't go: "Well, it's not Data, and it's not Knowledge, so it has to be Information"

Participant: Of course I have a reason why I want this particular...

Interviewer: But you're not communicating this reason, are you?

Participant: No. You go to the computer "I want this drawing. and send it to this printer."

Interviewer: Is that meta-data, is that communication?

Participant: I suppose I'm sending it Data, because I'm just punching in a few numbers.

Interviewer: That's different from the fact-finding Knowledge, yes?

Participant: Yeah.

Interviewer: So you are sending the noun 89 Data, because the drawing request is Data.

Participant: Yeah, that's right.

Interviewer: But you get back Information as the drawing.

Participant: Yeah.

Interviewer: You then take results from the production database, theories and experiences from books, and drawings from the noun 89 and what do you do with them?

Participant: I produce the material.

Interviewer: Now, what do you do with that material?

Participant: I would produce sort of lecture materials which I would – would give us communicate with the position 90.

Interviewer: Who would the lecture materials go to in terms of roles?

Participant: It goes to me as the lecturer.

Interviewer: so the lecture author sends to the lecturer lecture material.

Participant: Yes.

Interviewer: This lecture material, is this Data, Information, or Knowledge?

Participant: I would say all three.

Interviewer: So this is Data and Information and Knowledge.

Participant: I would say that it's Information and Data.

Interviewer: It's Knowledge?

Participant: It's Knowledge. I suppose it makes up your Knowledge if you say Knowledge is based on Information and Data.

Interviewer: Is it?

Participant: Knowledge... you learn from something. Experiences which can be either a specific experience Knowledge... but. Because if you're just sending Data, and the receiver is the person who has to come up with the – make up the Knowledge part.

Interviewer: So the author is sending Data, Information, and Knowledge to the lecturer.

Participant: That's the material container.

Interviewer: Does the author send anything else to the lecturer?

Participant: Not that I can think of.

Interviewer: Now, what does the lecturer do with that lecture material?

Participant: Communicate to the receivers. Now it's the case of position 90s, or the class.

Interviewer: Lecturer sends to the position 90s, what?

Participant: I would say mainly the Knowledge part. That's the big thing. That's the purpose.

Interviewer: So the lecturer sends to the position 90s Knowledge of, what can we label this?

Participant: It's Technical Knowledge.

Interviewer: Since we have 4 minutes left, let's do a very brief theory session. What is Data?

Participant: For me? It's Information. It's some kind of recorded readings, values of ??? discontinuous or continuous flow of events.

Interviewer: Now you mentioned the term Information. Can you define Data again for me?

Participant: Information I suppose... is a measure of the Data, as opposed to the measurable property.

Interviewer: So Data is a measure of a property?

Participant: Yeah, for me.

Interviewer: What is Information?

Participant: Information for me, would be more of the numbers.... but more sort of verbal text. For me, Information is text, and Data is this numbers. But even a text could be Data.

Interviewer: How is text Data?

Participant: Because sometimes text could describe a measure better than a number, I think. Depending on what you want to use it for.

Interviewer: so a measure of a property can be either qualitative or quantitative. Whereas Information is what?

Participant: I would say it's a quality, I think. Information, I suppose, can mix with all Data, Knowledge. Something that you relate – it goes from one source to another. Any kind of Data or Knowledge is some kind of Information.

Interviewer: So would you say Information is a container for Data and Knowledge? Some sort of communicative thing?

Participant: Yeah, I would think. Or, ??? when you're just looking at numbers... Information I suppose is a continuous flow of Data and numbers and Knowledge sort of ...

Interviewer: Okay. What's Knowledge?

Participant: That's harder. Knowledge could be your perception of something. If that's what you think you know, anyways.

Interviewer: Knowledge is what you think you know?

Participant: For me, yeah. But I suppose someone has some other Knowledge which I could sort of receive too, and then it becomes my Knowledge.

Interviewer: How can someone give you Knowledge? Do they just give you Knowledge, or do they give you Knowledge through something else?

Participant: You can receive Knowledge, I suppose. Text, verbally, visual.

Interviewer: But the flow would still be a Knowledge flow, it wouldn't be a Data flow or an Information flow?

Participant: It could be all. You can find Knowledge in ... you could get direct Knowledge as sort of verbal, but you can also receive Data but then you sort of have to compute the Data for it to become Knowledge. I think the Data itself does not become Knowledge until you do something with the Data.

Interviewer: What about Information?

Participant: It's the same, I think. You still have to do something. You have to put it in perspective.

Interviewer: Three very quick questions. First, is Data atomic, can you divide data? Is there something that's more... fundamental than Data?

Participant: Not that I can think of.

Interviewer: Is there something above Knowledge? Is there something that combines Knowledge into something bigger than itself?

Participant: That's a religious question.

Interviewer: Yes, it can be.

Participant: I wouldn't be able to describe it. I sometimes think that there's something even above Knowledge.

Interviewer: But we don't have a ...

Participant: No.

Interviewer: When someone tells you to do something: "You should do this" are they communicating Data, Information, Knowledge, or something else to you?

Participant: I think they're giving me Information, I would think. Instruction is ... that comes with a consequence. If I have to do something. I think they're sending me Information, I would think. Which I would probably transform into some kind of Knowledge. Of what the consequences would be if I don't go along with the Informational instruction.

Interviewer: If you're making predictive statements about the future, or from the past to the present, are those predictive statements, Data, Information, or Knowledge? And what are you using to create them?

Participant: It would be Knowledge. And I would be using Information and Data to produce these Knowledge.

Interviewer: How does experience turn into predictions?

Participant: Intuition.

Interviewer: You'd say experience is Knowledge?

Participant: It's not perhaps, it's part of the Knowledge.

Interviewer: and you'd say that causal models are part of Knowledge? Like if I've got these inputs, this output would happen?

Participant: Yeah, sort of, predictive models ... they would be Knowledge, yeah.

7.9 Interview 9

Interviewer: Let's model both topics.

Participant: Do we have time?

Interviewer: No, but I want to model both. Let's do the noun 91 first.

Participant: it's probably more relevant to this particular group. In that there was very little interaction from the reline in terms of work that I did.

Interviewer: I'm going to skip the entity dictionary. Normally what I do with people is: let's just go through and figure out what entities there are. What I'm going to do instead is describe to you what an entity could be, and we'll just jump in because I'd love to get both. An entity is a role that someone or something plays. Tools can be entities if they're not ready to hand. A pen isn't an entity because when I'm writing, I don't go "pen." But, SAS is probably an entity because I'll be doing lexical analysis through it. If it's providing a transformation, it's an entity. You, yourself, can have multiple entities. As a noun 91 student, I've got position 56 hat, I've got teaching assistant hat, I've got grader hat, I've got database position 31 hat. Just because I'm one person doesn't meant I'm one entity. So, starting with the noun 91.

...

Interviewer: This is looking at both your noun 91 methodology and the noun 91 noun 97. The surrounding... We'll start by modeling a trivial noun 97 in your noun 91. because it's best to start slow. Which entity would you like to characterize yourself as first?

Participant: Me?

Interviewer: You as?

Participant: I guess as a position 56? The one doing the work.

Interviewer: As a position 56. You've got flows of stuff. I use the word stuff because I'm not going to day Data, Information, or Knowledge. Or other. To someone else. What is a trivial flow of stuff to someone else and who is that other person?

Participant: I guess the trivial stuff would be reporting: talking about progress, results. So my main position 92s here are [name] and [name].
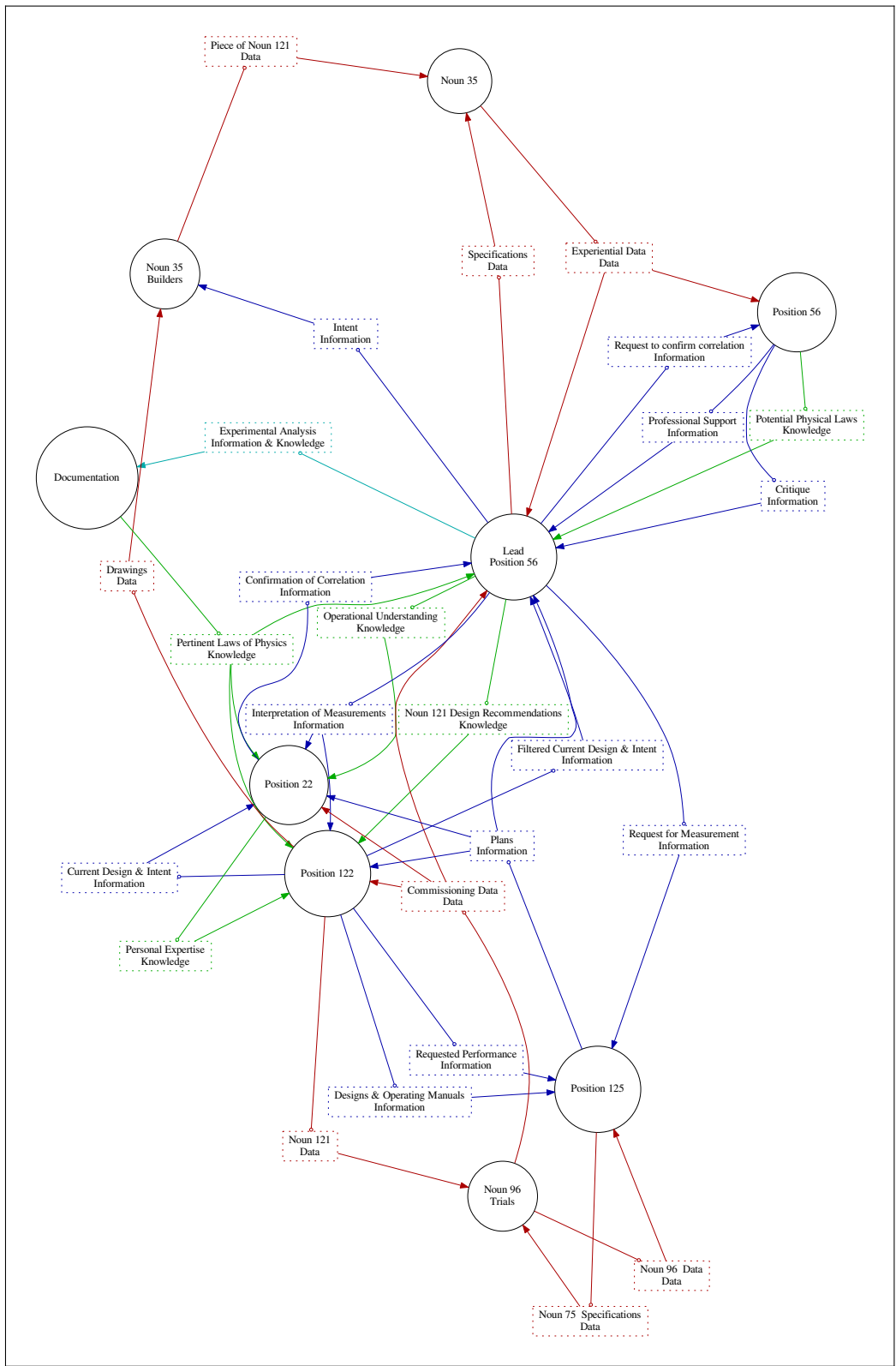
Figure 7.12: The SDFN Diagram for Interview 9

...

Interviewer: So, position 56 to position 92s. What are you sending them?

Participant: I'm sending them Information. I meet with them about once a week to talk about what I've done. It might be results that I've gotten, that I can show them. It might just be talking about how things are going.

Interviewer: Do we want to differentiate that? This diagram is atemporal. Everything is superimposed because dealing with reality is such a pain. Let's break that down into multiple flows. You're telling them what you've done.

Participant: yeah.

Interviewer: Do we have a better label than "what you've done?"

Participant: Progress?

Interviewer: Now, you send them a flow of stuff about progress. Now this flow of stuff could just be a conversation, but it's still a transfer of... stuff. The stuff you identified is Information?

Participant: As in results?

Interviewer: Your progress update. Is your progress update Data, Information, Knowledge, or O?

Participant: Information?

Interviewer: You may also send them Results. These results are?

Participant: Data.

Interviewer: Do you send them anything else in these usual meetings?

Participant: What about concerns? Can you send concerns? Emotions? Frustration?

Interviewer: Do you want to summarize them as one flow?

Participant: Concerns, I guess, is the right one.

Interviewer: These concerns that we all have are what?

Participant: Well they're not Data. They're not really Knowledge. So I'm assuming they're Information.

Interviewer: Don't use exclusions.

Participant: Can they be emotion?

Interviewer: Sure. Is the communication of emotions a unique concept? Or is covered under the other categories? Is emotion distinct in itself or is it part of some larger superset?

Participant: I would have thought that it would be distinct in itself. ... I'm going to think that emotions are separate. [someone else] probably wouldn't look at that.

Interviewer: This is a really useful interview.

Participant: Because it's true! I'm a lot more emotional than a lot of the [group] are. And so, for me, communicating emotions and getting feedback and getting my emotions right is important for me...

Interviewer: And this is distinct from Data, Information, and Knowledge?

Participant: yeah.

Interviewer: Do you send them anything else?

Participant: No, I think that covers my part.

Interviewer: What do they send you?

Participant: I'd probably break it down into feedback and guidance.

Interviewer: Feedback is?

Participant: It's Information. Yeah, I'd put that one as Information. That's responding to what I've done and telling me how they think things are going. Whereas guidance is probably more Knowledge, because that's more pushing me in certain directions or giving me the understanding that I need in order to progress in different directions.

Interviewer: What you're identifying here are normative statements: "You should do this." I'll want to unpack those later. So, your position 92 sends back feedback and guidance. Do they send anything else?

Participant: Occasionally they'll send – there's a problem where I've done something and it hasn't given the results we expected. And so [name's] done the work himself and then provided results back for me to compare to.

Interviewer: So these are comparative results?

Participant: yeah. They'd be Data.

Interviewer: Who/what else do you work with? Do you get or send flows to as a position 56?

Participant: My other position 92 who is my [location] position 92. Because he's kind of separate. Because I don't talk to him that often.

Interviewer: adjective 93 position 92? or ignored position 92?

Participant: Hey, it goes both ways.

Interviewer: You to your adjective 93 position 92 send what?

Participant: The main thing would be progress, than anything.

Interviewer: as Information?

Participant: Yes.

Interviewer: Is it the same progress?

Participant: More condensed. So it's not the same progress. I'll talk to him every few months or so and I'll talk to them every week.

Interviewer: Condensed you would say?

Participant: Maybe summary.

Interviewer: Summary of progress is?

Participant: Information.

Interviewer: Does he send anything back?

Participant: ...

Interviewer: Sorry, in the best of all possible worlds, does he send anything back?

Participant: Occasionally?

Interviewer: What does he send back?

Participant: Occasionally he sends back guidance, Knowledge.

Interviewer: Is it the same guidance?

Participant: No, different Knowledge.

Interviewer: How can we differentiate these guidances?

Participant: I don't know. Direction?

Interviewer: Tell me this guidance versus this guidance. They send you what?

Participant: They [local position 92s] send me kind of week to week guidance. At the moment I've been developing a ... model to try and describe [undesired chemical interaction]. And so I've developed something to counteract that. So, over the noun 97 of developing that, [name] has kind of given me direction in terms of: "Oh, should we consider this model instead?" and I'll try that and it won't work. And I'll say OK, that one didn't work, what about this. And point me in different directions. and then I'll do the work and come back and say none of the work serves what I've done, which is what happens. Whereas with my adjective 93 position 92 it's more of, I'll present him a summary of what I've worked on for the last 2-3 months and say "This is where I'm planning to go." More of the higher up position 92 than a day to day thing. And I'll kind of talk more about this is my plan for where I'm heading based on what I've done. And he'll provide kind of direction and, if necessary, Information on different things I'm planning on doing.

Interviewer: So this directed guidance, as Knowledge, so he provides ...

Participant: when I say Information, I mean papers, or contacts, or things like that. And that kind of stuff I would classify as Knowledge. Because it's stuff he has that I don't.

Interviewer: so his Knowledge contains Information?

Participant: Yes.

Interviewer: He sends back directed guidance as Knowledge which contains Information. But the flow is labeled Knowledge, because it's an expression of his Knowledge?

Participant: That makes sense.

Interviewer: It actually makes complete sense to me. What else?

Participant: I guess the only other thing he does, as an adjective 93 position 92, is that he can put me in touch with other people. I don't know if you have any points for networking.

Interviewer: of course I do. Let's start with a flow from him to you. He creates networking opportunities by sending you what?

Participant: Contacts?

Interviewer: Contacts. These contacts are?

Participant: people? names?

Interviewer: These people and names are what?

Participant: So they're Information. They're not really Knowledge. I guess they could be Data. It depends on your definition.

Interviewer: Of course. Are they Data? Are they a different sort of Data than your other Data?

Participant: If they are Data, then yes they are.

Interviewer: Are they Data?

Participant: I'm an position 58 by background, I can't tell you anything.

Interviewer: It's so much fun pinning position 58s down, because they squirm.

Participant: Yeah, I can feel myself squirming. I would say: "Look, you can put them in a list. There is a defined quantity. They can be grouped, they can be ordered or whatever. They probably are Data. You can put them in a spreadsheet.

Interviewer: But they're different Data. How can we characterize this Data as different Data?

Participant: They're people.

Interviewer: They're not numbers.

Participant: People don't like to be defined as numbers. They get all kind of emotional at me.

Interviewer: So, contacts are Data. Do we want to differentiate this flow of Data from your other flows of Data? Is it a difference in nature and kind or just nature? Are you comfortable with us labeling this Data in the sense that it will be perceived as the same as this other Data?

Participant: Meh.

Interviewer: Excellent. What other entities do we have?

Participant: With my position 56 hat on?

Interviewer: Yes.

Participant: There are other people who I send Information to but don't get anything from.

Interviewer: which entity is this?

Participant: [name] as my research position 41. As in the [local company] research position 41.

Interviewer: Research position 41. You've got a flow to him of?

Participant: summary of progress.

Interviewer: The same flow. And the type of flow?

Participant: Information?

Interviewer: Are there any flows back to you? Are there any flows back to other people regarding this interaction?

Participant: Not particularly. It's just one sided. It's just as a courtesy, keeping him in the loop sort of thing. So he knows what's going on but doesn't necessarily have a direct hand.

Interviewer: What else do we have?

Participant: I guess you can put other position 56s at [academic institution], but I don't do much with them.

Interviewer: If you don't do much with them, then they're not relevant. Two other areas to explore. Communications with yourself wearing other hats and communication with tools.

Participant: Does my computer count as a tool? That's really my only tool.

Interviewer: Do you use different programs on your computer in meaningful ways?

Participant: Not heaps. I do all my coding in C. Well, yeah, I've done a bit of computer programming before but nothing too extensive before this. But so far I've learned C, C++, FORTRAN. They're not too bad, when you don't know anything better, they're not too bad. So I've done most of my coding in that. I often use one visualization package that I use to look at results or I just put them in excel.

...

Interviewer: Okay, let's go with visualization software, and we can lump excel and your other vis stuff in here unless you think we should

Participant: No that's fine.

Interviewer: do you send different stuff to your visualization software than excel?

Participant: Yes and no. I get the same kind of things back from them. I don't do an awful lot of transformation stuff. I don't do a lot of formulas in excel. I just put the numbers in and get graphs. The visualization stuff I'm putting numbers in and I'm getting pictures.

Interviewer: And you've got, and what entity would you say your programming is? How do you conceptualize it in your head?

Participant: It's code.

Interviewer: You as position 56, send the entity known as code, what?

Participant: I send it Data. I created it though, so how does...

Interviewer: That's the question. You send it Data. What Data do you send it?

Participant: Well I send it the lines that it uses to create the program. So I send it... I don't know how you would describe that. I write the basis of it. I tell it what to do. But I also send it numbers to work with.

Interviewer: Let's differentiate those into two flows. The numbers it needs to work with are?

Participant: Data.

Interviewer: And we can label them?

Participant: inputs.

Interviewer: You also send it what?

Participant: I guess I'd classify it as Knowledge. No, it's Information. It's kind of all of the above.

Interviewer: we can certainly combine elements.

Participant: Not really Information.

Interviewer: so you send it what?

Participant: I would call it Knowledge. Because I'm imparting my Knowledge into the code and telling it what to do.

Interviewer: and what will we label this flow?

Participant: Maybe model theory or something?

Interviewer: What other flows are there?

Participant: Well it just sends back results.

Interviewer: Same results or are these results different from these results?

Participant: They are different. But not in nature. Just in ... obviously I'm not going to take every result I take from the code and send it on. Because that would be ridiculous.

Interviewer: So can we say selected results over here?

Participant: Yeah.

Interviewer: What else?

Participant: That's really about it.

Interviewer: when you do debugging, does it send anything?

Participant: Only when I tell it to.

Interviewer: when you tell it to send you things, does it send you a flow outside of results?

Participant: not really. I would say...

Interviewer: just different kinds of results? How do you interact with the visualization software and does your code directly interact?

Participant: no.

Interviewer: How do you interact with the visualization?

Participant: I send it the selected results.

Interviewer: the same selected results?

Participant: Different selected results. You can make it the same as the one that comes out of the code because, for all intents and purposes that's what I use. The visualization package doesn't always use all of the results that I send it.

Interviewer: but it's the same results, there's no transformation, there's just filtering.

Participant: the filtering is usually done inside the package.

Interviewer: you send it results, what does it do?

Participant: it gives me pictures or pretty graphs. like circles and balls. I've got very unexciting pictures.

Interviewer: It sends you back pictures and graphs...

Participant: The pictures come from one software and the graphs come from the other.

Interviewer: These pictures are?

Participant: Data?

Interviewer: Why are they Data?

Participant: Because they're based on particular things.

Interviewer: So it's just Data as representation of Data.

Participant: Yeah.

Interviewer: It also sends you back graphs. These graphs are?

Participant: the same thing.

Interviewer: Data. What else do we have?

Participant: That's mostly it.

Interviewer: Do any of these entities exchange stuff... about your work?

Participant: to a minor extent. I would say that there's a minor extent link between my position 92s and my adjective 93 position 92.

Interviewer: OK, what links are there?

Participant: All I can think about is discussion.

Interviewer: Who is discussing with whom?

Participant: Well, it's a double sided arrow.

Interviewer: This discussion is?

Participant: Information.

....

Interviewer: Let's see what we can do on verb 94. Which entity do we start with for verb 94?

Participant: if we're looking at it from my point of view, so I was an position 58.

Interviewer: And what flows are there, to what?

Participant: So my position 41 on the project, we'll call him the lead position 58. He was just responsible for our group.

Interviewer: flows, you to lead position 58.

Participant: I'd send him any work that I've done. It's kind of hard to quantify what I did.

Interviewer: Are there categories of what you did?

Participant: Yeah, there are. Maybe I could break it down to three main categories. position 58ing noun 95 work, so noun 95 calculations and stuff like that. Maybe noun 95 calcs?

Interviewer: And these noun 95 calcs are?

Participant: They're Data. noun 95 calcs, and adjective 96 stuff, documentation I guess I'd call it. adjective 96 stuff is kind of stuff that's done, so we do adjective 96 and we go use a Noun 18 or test a valve. Documentation is all of the documentation that we wrote up ...

Interviewer: Your ... documentation is?

Participant: I don't know what you'd classify it as. I guess it's a type of Data.

Interviewer: Can we separate it out into sub-categories?

Participant: Not really.

Interviewer: So it's not Information, it's not Knowledge.

Participant: It's a kind of a bit of Knowledge and a bit of Information and a bit of Data all rolled into one document. So what is the document?

Interviewer: Let's start by: "What is the document?"

Participant: It depends on which document.

Interviewer: What categories of documents are there?

Participant: The main ones are descriptions of the noun 97, describing what the noun 97es do. I need to explain that I worked mostly on the noun 98 stuff.

Interviewer: So we've got noun 97 description documentation. This noun 97 description documentation is?

Participant: Well it's all the above. It's all of them.

Interviewer: So Knowledge + Information + Data?

Participant: Yeah.

Interviewer: And documentation not including noun 97 description documentation has what other bits in it?

Participant: adjective 96 reports.

Interviewer: So we have adjective 96 stuff, but that's different from adjective 96 reports.

Participant: Yes.

Interviewer: These reports are?

Participant: They're Information and Data.

Interviewer: But not Knowledge?

Participant: Nope.

Interviewer: Why not?

Participant: Because the Knowledge contained in the noun 97 descriptions is a kind of background behind what was done and why. Whereas the adjective 96 report is just a description of what was done during adjective 96 and the results that were obtained. So the dumb stuff is the Information. And the results are the Data.

Interviewer: are there other categories that we can extract out from documentation?

Participant: That'll do for now or we'll be here all day.

Interviewer: In documentation, the remainder of the documentation is what?

Participant: Mostly just Information.

Interviewer: The adjective 96 stuff is what?

Participant: I would classify it as adjective 96 results and call it Data.

Interviewer: All right, what else do we have? What other interactions?

...

Interviewer: are there any flows between position 58 and lead position 58?

Participant: The lead position 58 tells me what to do?

Interviewer: What can we label this telling you what to do as?

Participant: Directions sounds a bit too nice but it will do.

Interviewer: Direction is?

Participant: Information, I guess.

Interviewer: Does the lead position 58 send anything else back to you?

Participant: I guess they send back Knowledge. I'm trying to think of what form that would come under. But say there's something in the adjective 96 results that's not quite as you'd expect it, then they can give advice as to what could be causing it, or what to do to test something

else. So you could call it noun 97 Knowledge. The point of the lead position 58 is for them to have more comprehensive understanding of how things work. You can put some other entities in if you want to. "other Position 22s." My main flow of communication with them... My background with the project was that I knew how all of the systems worked because I noun 95ed the system. ... I knew how things worked and so I would explain to them how things worked. ... I guess my main flow to them would be noun 97 understanding. Call it Knowledge.

Interviewer: Do they send anything to you?

Participant: They would me adjective 96 results.

Interviewer: The same ones?

Participant: The same Data.

Interviewer: How are they different?

Participant: Well, it could be... they do adjective 96 on?? other systems. So different results. Whereas the stuff that I send would be whatever I was involved in. If you want another arrow there, they send those adjective 96 results to the lead position 58.

Interviewer: so these results are?

Participant: They're Data.

Interviewer: What else?

Participant: The main one I had was the noun 63 position 58. I'd teach them about Position 22ing. I'd probably give them the same noun 97 understanding as the Os. They would send back. Well, they would send back understanding of how the noun 63 works. So that would be – I guess that would be Knowledge. It's very hard to break it down into data.

Interviewer: where do you get these noun 95 calculations from?

Participant: I do them. If we want an entity for them, the best thing is that I do noun 99 modeling, and I use the noun 99 package. That's a nice little name for it.

Interviewer: What do you send to the package, and what do you get back?

Participant: I'd send input, Data. And it gives me, results, Data.

Interviewer: Anything else?

Participant: Well, I don't know where all the equipment is.

Interviewer: Is the equipment important?

Participant: Probably not. I mean it is important for the job, but probably not.

Interviewer: Does the equipment provide Data, Information, Knowledge, or O?

Participant: Of course it does. It provides all of the above. Well, maybe not Information or Knowledge... It's very weird thinking of it as a Knowledge transferring. Maybe you could call it the noun 63 that's giving us Information back. The [object] through it doesn't really tell you anything.

Interviewer: OK, the noun 63. What do you send to it? What does it send to you?

Participant: We don't really send it anything. Well, I don't send it anything.

Interviewer: What does it send to you?

Participant: It sends noun 97 results. Data, because I don't have anything better.

Interviewer: now you mentioned Information at some point, with regards to this.

Participant: Well, it's the Information that comes from the results.

Interviewer: where does that Information get...

Participant: transformed? Through the position 58. Well, an position 58 looks at it and goes "Yes, that number looks right. Or no that number doesn't look right. It's broken."

Interviewer: And that's a transformation into what?

Participant: It's a transformation into whether things are working.

...

Interviewer: Define for me, Knowledge. Are there types of Knowledge?

Participant: I would define Knowledge as something that an individual or entity ... something that an entity possesses. That typically has been learned from somewhere else or gained from experience or whatever that enables them a better understanding then they would have otherwise. My understanding of how things would work is that Knowledge is the noun 97 by which Data is turned into Information. So I would say Data is things. It's numbers, it's raw Information. It is something that, on its own, doesn't mean very much. It's just stuff. Information is an interpretation of that. So it's kind of something that is understandable to someone without Knowledge, without even concept of what the raw Data is.

Interviewer: So it's encapsulated something? What is it that's encapsulated?

Participant: I mean, it can be an explanation of results. It can have nothing to do with it whatsoever.

Interviewer: So we've got Data is transformed by Knowledge into Information. What does Information do and where does it go? It doesn't just sit there.

Participant: Information is used as a means of communication.

Interviewer: What is it communicating and does it cause any changes in anything?

Participant: I don't know.

Interviewer: Where does emotion fall into this if anywhere?

Participant: Well I'd say it's fairly separate. However, it can influence how things are perceived. So it probably has an influence on the Information.

Interviewer: So emotion changes the transformation of Data into Information?

Participant: I'd say it just changes the perception of Information.

Interviewer: So Data goes into Knowledge, what's this flow?

Participant: Data doesn't really go into Knowledge.

Interviewer: Sorry...

Participant: Data doesn't turn into Knowledge. Data turns into Information.

Interviewer: So Data turns into Information. Knowledge...

Participant: creates that link. So if you don't have Knowledge, the Data won't transform very well.

Interviewer: does Knowledge go into Information?

Participant: Yes, Information can be a communication of Knowledge.

Interviewer: Does anything go into Knowledge? How do you get more?

Participant: Through Information.

Interviewer: What is Knowledge?

Participant: understanding inherent to an entity, is how I'd describe it.

Interviewer: Three questions: Is there anything to either side of Data, Information, or Knowledge? Is there anything that's lower than Data? Or is Data atomic?

Participant: I would say Data can be atomic. It stretches across that width of the spectrum. I would say that it goes from as far as atomic to as big as macroscopic. It can be anything within that. I'm saying that in any sense.

Interviewer: Data can be divided or it cannot. If Data is divided, what is it divided into?

Participant: Subdata?

Interviewer: Is subdata a thing? Or is subdata Data?

Participant: I would say it's not any different. You could have different scales of it, but... you could have a group of numbers. You could select some of that group of numbers. That's not any different. It's just still Data.

Interviewer: So you have a subset of Data which is Data.

Participant: Yeah.

Interviewer: and at some point you just can't subset any more.

Participant: You could get it right down to the last point... It's still Data

Interviewer: You could have many points, and that collection of points is still Data.

Participant: Yeah.

Interviewer: But you have a collection of points, and you apply Knowledge and you get ... Information? And then you've got this loop of Information to Knowledge and Knowledge to Information. Is there anything beyond Knowledge?

Participant: I would say no. I would say the only thing beyond Knowledge is god. I mean God is all Knowledge. That just gets confusing. <inaudible>

Interviewer: Because it is part of some people's ontologies, but it's outside the scope of this investigation ... When someone gives you a normative assertion, it's what of these, if anything? When someone orders you to do something, is that Data, Information, or Knowledge?

Participant: None of them.

Interviewer: What is it? Do we have a handy label for it?

Participant: I don't know. If anything, I'd say it falls under Information, but I don't think it does.

Interviewer: So what is Information?

Participant: It's a lot of stuff.

Interviewer: What's a component of Information?

Participant: It could be an interpretation of something. It could just be words. Understanding, words. I would say if someone asks someone to do something, provides them a direction or a directive, it's not necessarily a should, but, provides them guidance to what they should do, I would say that's Information, they're giving them Information on what to do.

Interviewer: The guidance is what to do based on what?

Participant: Based on that entity's understanding. Based on their Knowledge or their Information. Or the interpretation they've got of things.

Interviewer: Any final thoughts?

Participant: Nope.

....

Interviewer: ... your use of Emotion as a perception of manipulation engine is fascinating. So, thank you for that.

Participant: That's OK. Maybe that's where should comes into it. Maybe it that when someone tells you that you should do something, they're kind of applying their emotiveness behind it. In that, ??? said, it's a manipulative thing. It's not necessarily a "I'm guiding you to do this." It's a I'm making or telling you that you must do this. So they're exercising their manipulativeness over them. Which can be a good thing. Like a position 41 with an employee might help them to do something. And that's manipulation.

Interviewer: and so it's an emotive communication of...

Participant: status in the hierarchy. Maybe that's emotive, maybe that's not really emotive.

Interviewer: You tell me

Participant: I don't know. I'll just confuse myself.

7.10   Interview 10

Interviewer: Let's model this. A brief discussion on entities. An entity is something, someone, that takes in these flows, performs some sort of transformation on them, and outputs flows. Or produces flows, or takes in flows. I would say that, when I was doing my masters, I had three roles: I was a student, I was a teaching assistant, and I was the database position 31. For purposes of this diagram, I would be three different entities because not only do they deal with Data, Information, Knowledge, in different ways, but they can talk to each other. Even though it's me talking to me, it's me as DBA going: I just wrote this lab me, here are the questions we need to deal with. It's the DBA talking to the graduate assistant, there are different ways of dealing with stuff. We will ignore things that are ready-at-hand. This pen is not an entity because I don't think about it when I'm writing, I just write. Say I use that computer and I load up SAS, and I load up the 100,000 words of transcript that I've written so far. And it munges them and it produces something. SAS is an entity because it's taking stuff in and giving me something. It's noun 97ing it, and I'm aware that it's noun 97ing it. It's an entity to me because it's not just an invisible flow it's go there, come back. Scientific instruments can be entities, they can take stuff... What I'm trying to do is be very vague here so I don't prejudice your philosophy. This is why I'm saying if you think doing an entity dictionary is useful we'll do it. If you've got a fairly decent idea, we'll start trivial and expand out.

Participant: we'll start trivial?

Interviewer: Let's start by talking about one of the really beautiful noun 37 cases we just talked about with that semipermiable membrane. With ..

Participant: technology / business .
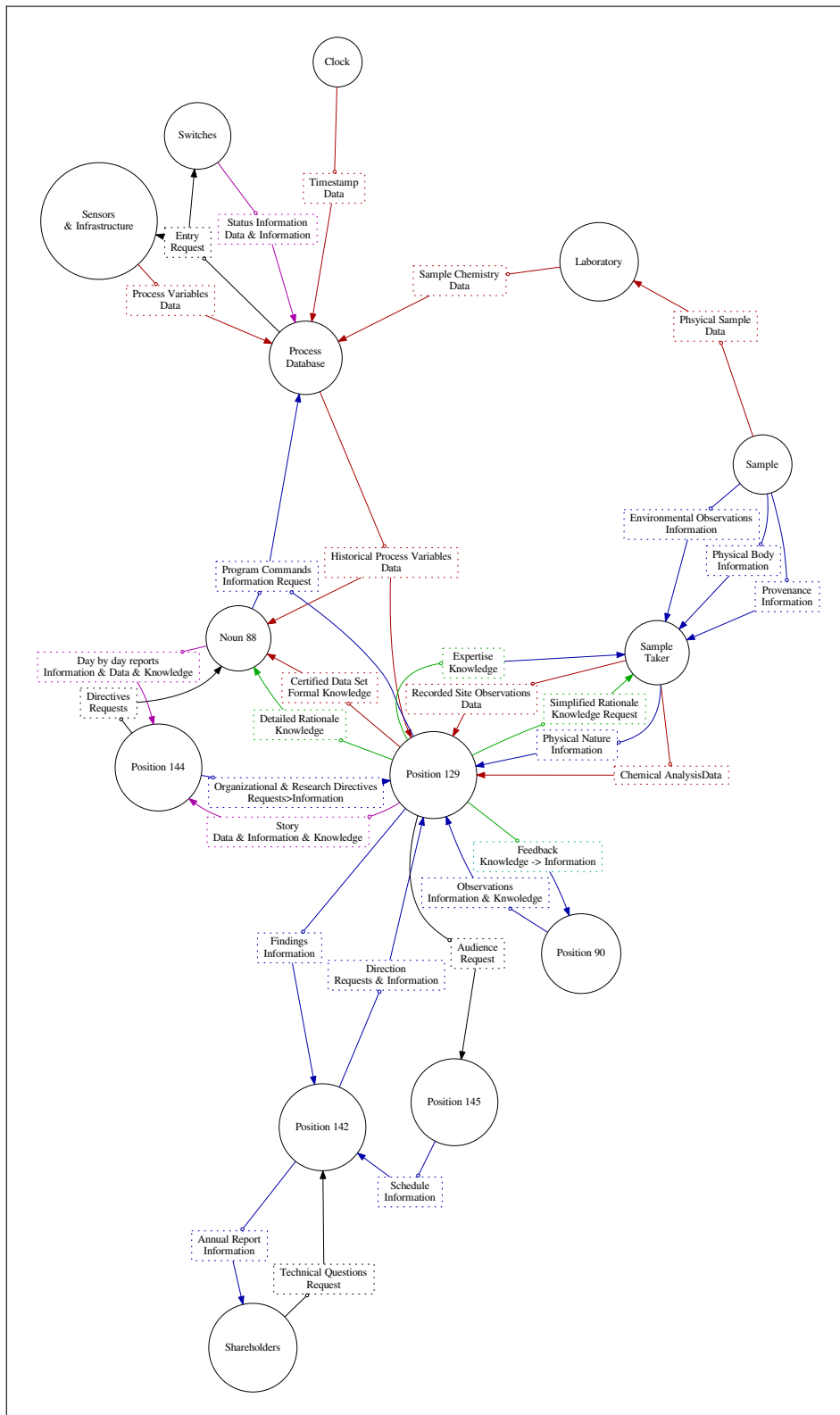
Figure 7.13:   The SDFN Diagram for Interview 10

Interviewer: we start out by going what is one of the central entities in this noun 97? probably a role that you play, but it doesn't have to be.

Participant: As the collector, just call it the noun 129 entity. noun 129.

Interviewer: multiple people can be in one entity.

Participant: noun 129. What happens in the noun 129 well: the inputs are observations....

Interviewer: What is effectively sending you observations?

Participant: position 90s. We'll probably do this around [name]. So the position 90s. These are corridor conversations. You'll be sitting there and some position 90 will say: "Do you know such and such?" ... and the position 90s will come around for a little chat. Because they know I'm there. And you pick up all sorts of things that are going on. That's the informal observations.

Interviewer: So we have position 90s sending a flow to noun 129. This flow is observations?

Participant: Just observations. They push the buttons in the noun 97. They hear things, they see things. They know how the noun 97. They know the differences between the shifts. And if you want to put a credibility factor on it, around 20

Interviewer: position 90s send you a flow of observations. Here's the fun bit. Would you categorize this flow as Data, Information, Knowledge, or something different?

Participant: Can I pick two?

Interviewer: absolutely.

Participant: Information and Knowledge. Because the Information is the things that I can reduce to something technical. The Knowledge is their understanding of the noun 97. ... what's their perception of the basis of [their choices]. Information to me would be the technical side of it. The Knowledge is the operational expertise. So you're getting some of those two things coming through.

Interviewer: So they send you observations. The observations are Knowledge, or are Information.

Participant: Pretty nebulous. Handwavy stuff. Some as good as gold.

Interviewer: Is there a flow from noun 129 back to the position 90s?

Participant: Yes. There probably are many. But again, it's limited to the ones that you talk to. And it's, for me, trying to "this is the next level down of telling the story" of why, why do these things happen? hang on fellas, that's not the way it works, this is it. Some of them? [company] ... ??? Interesting story, giving a presentation. This was a few years ago. It was noun 131s. These things are like big bowls. So we're doing a formal presentation, there's about half a dozen, a position 106 from here... I start giving this background of noun 131s. And this guy says: "[name], what you're really saying, independent of type of noun 61, they all form

for the same reason" and I say "spot on." Did group 130 get that? no. But one of the position 90s picked it up in 5 minutes of me talking. And so, that was a formal feedback to the position 90s, but when I go back, there's a lot of informal value in the corridor. noun 131s is all about this. Noun 21 noun 132 is all about this. noun 137 is about this. So, for me personally, there is an informal feedback.

Interviewer: So there's feedback back, whether it's formal or informal, it takes about the same form?

Participant: Yes. Just talking. Just start drawing lines on a whiteboard. Putting together the pieces of the jigsaw puzzle into some sort of digestible ... my job is to tell a story. The complexity of that story depends on the audience.

Interviewer: Would you say the feedback is Data, Information, Knowledge, or O?

Participant: It's my Knowledge, but I'm giving it back as Information.

Interviewer: Let's unpack that. Would you say the flow is absolutely generated from your Knowledge, and you assert that they perceive the flow as Information?

Participant: Yes.

Interviewer: Would you say the flow itself is a flow of Knowledge, or a flow of Information?

Participant: Put the two together I'd say.

Interviewer: Why?

Participant: Because the Knowledge base that I'm drawing on, is much more substantial than these.

Interviewer: So you're passing them Knowledge, but that they accept as Information?

Participant: Yes.

Interviewer: What do they do with that Information?

Participant: They probably stew about it. Would they incorporate that into their actions? No.

Interviewer: Which is why you would categorize it as Information?

Participant: Yes. Because I have no power to direct which way they want to noun 97. That would hopefully be operating through the form of noun 88. My role in that loop is to just give them the background why? They might agitate for that, "But [name] says this." But [name] doesn't have a clue. Every now and then, if it's the right sort of segment, they will push that in.

Interviewer: So you would say that it's not so much a flow of Information and Knowledge, as that it comes from you as a Knowledge flow, and then becomes an Information flow they accept. ... Are there other flows between these two entities?

Participant: in the example I'm thinking about, no. Because I'm here as the blowing consultant. You're hoping that the formal loop of noun 88 would be reinforcing that message. That would be from this entity up to the noun 88 entity.

Interviewer: What do we want to label that as?

Participant: noun 88? This is the group that sets the operational procedures and standards. This is how you shoot this wheel into that position. You will run this noun 97 at this speed. You will do this if the noun 132 in the Noun 21 exceeds this value. That's the formal bit. So, what I provide to them is the written report which is Knowledge, I would suspect.

Interviewer: So there's a flow from you to them, of Knowledge.

Participant: Formal Knowledge. This is documented.

Interviewer: So you would categorize it as a different kind of Knowledge?

Participant: Yes. This is the next level up in technical understanding.

Interviewer: This formal Knowledge flow... what's the content of the flow? Just as the content of this is observations.

Participant: This will be datasets, analysis, and some observational material that supports it. This is really about the formal Data set.

Interviewer: So you're passing to noun 88, the distilled version...

Participant: you probably do release your Data sets. There will be tables and numbers. If it's an analysis, there will be a table of numbers.

Interviewer: But this is your interpreted noun 97, refined...

Participant: This is my certified... Data set. I believe the Data set that best and most reliably represents the analyses that we've taken as part of this. Simple things like the tables add up to a sensitive number. If you look at noun 133 analyses, the table is not necessarily 100

Interviewer: But they're certified, because...

Participant: They fit more or less internal guidelines that you have for quality of Data. If it's outside, there's something wrong with it. I've gotten analyses out of there that only total 70

Interviewer: Where do you get the Data for this certified Data set?

Participant: From noun 37 and analyses.

Interviewer: In the chain of entities that may provide you with this Data set, what is the penultimate one?

Participant: For me personally, it is a physical noun 37 for which I have an analysis.

Interviewer: So we have noun 37?

Participant: noun 37. That's only because of how I take problems, and the sort of problems I take. The position 134 says I want the noun 37.

The position 135 says I want the analysis. For the people here, it will be a block of noun 97 Data. [Name]: "I've got two years noun 97 Data." I've got a relationship." His noun 37 is actually the set of noun 97 Data. And for me, occasionally, that's probably my second set. I'll look down to the noun 97 Data.

Interviewer: So we've an entity "noun 97 Data"

Participant: Yes.

Interviewer: Do we want to say that the entity is noun 97 Data, or say that it's a noun 97 DB?

Participant: I'm extracting specific variables from a DB. [name], could you get me tail Noun 23 and noun 28 chemistry? Can you get me Noun 21 noun 132s from noun 136 such and such? Can you get me the Noun 21 noun 137? can you get me the delivery noun 137 stock? That's a DB. Someone else will extract that for me.

Interviewer: Let's start with noun 37. Do you send any flows to the noun 37?

Participant: I hope not. No.

Interviewer: What flows do you get from the noun 37?

Participant: I get noun 37, observation. I see the physical nature of the noun 37.

Interviewer: Let's go with physical nature?

Participant: Yes. This looks like normal Noun 21 for [company.] This stuff's undergoing [term]. It's different. It shouldn't have done this in three days. There's something about the physical nature.

Interviewer: So you're taking physical nature. The flow of observations of physical nature is Data,Information, Knowledge, O?

Participant: Information. But I'm using my Knowledge to assign the Information, parameters, based on my previous experience.

Interviewer: Unpack that for me. You would say that it's Information...

Participant: It's Information. But that Information is actually based on an experience and technology???.

Interviewer: Here's a better question. You say that this Information is a function of your prior Knowledge. Where is the input of your prior Knowledge such that this Information is produced.

Participant: It's my long term experience with ... Or anything for that matter. If we went to: "What do I do with these noun 37?" First is observation. The second is some sort of chemical analysis. And the third thing, which i'll come back to is what it looks like at the end of a microscope. And one of the more or less concepts that I push to people is that the story should be independent of scale. In the sense that what I see down the microscope and what I see in the specimen. Or what I see from the scanning electron microscope, they should be a consistent story. When I write the summary those observations should

be independent of scale. They should be what I see. A lot of it is what I see down the microscope has to match in some way, shape, or form, what the SEM is telling me. They tell you different bits of the story, but they must be consistent. They must be the same story. Or compatible with the same story. And, I keep talking about what you see in textures and materials should in some way or form, be explainable between the two. The SEM and the optical microscope. Some people don't see it that way. "Oh! but this is different." No, there the same materials. How you see them may be slightly different, but the story should be consistent. The next level down. It may not look the same, you're just seeing different visualizations of the same thing. That's probably something about the noun 37.

Interviewer: Okay, what it looks like is that physical nature and information isn't a direct flow. That story tells me that there's a role that acts as an intermediary between noun 37 and noun 129. What is that role?

Participant: That role, that's the Knowledge role? Knowledge or expertise?

Interviewer: It's the person making the observations. What name can we put to that role?

Participant: I'm loathe to say. But it's almost like an expert, it's an expertise I have.

Interviewer: what I see is that noun 129 sends to this role Knowledge.

Participant: yes.

Interviewer: This role gets something from the noun 37, combines it with Knowledge, and sends physical nature as Information back into the noun 129.

Participant: into the development to write the story.

Interviewer: So we've got role X here ... you as noun 129 sends Knowledge to this role. This Knowledge is expertise?

Participant: Yes. Accumulated expertise which goes right back to noun 134 training in a lot of ways. Everything has a precedent in the past. An organization that forgets its past doesn't have a future. It's the same thing.

Interviewer: And you get the physical nature from this entity. You as noun 129. I'm pretty sure they're both you. Now, this entity, we can call it expert. Can we call it something else? What would you call it if you were telling someone else to take noun 37 for you.

Participant: If I had pre-knowledge of what we were verb 73, I'd say look out for.... I was just thinking that yes I had been in a situation. Usually the critical thing is: just get me a noun 37. "Oh, oh, we've got the noun 37." But that, in the industrial sense is extremely hard. Usually if you've got a bit more control, you want to refine your location. The noun 138 that you'd like to take a noun 37 from. ...

Interviewer: Not just I want a noun 37, but I want a noun 37 from that thing.

Participant: The noun 139 there. I want a noun 37 out of thatnoun 140 and I want it pretty soon... And I will go to the pile and pick it out. "Well, that's different, that's different." Because you're almost doing the analytical stage in your head. "Oh, we've got a noun 37." and the next question is: well where was it? I don't remember. That's part of it's: what is its relationship to its other entities? What was its orientation? Is this the top or the bottom? Simple things like that. Usually in giving guidelines for verb 73, it's "get me a noun 37."

Interviewer: position 141.

Participant: Yes, position 141. And usually you keep it s simple as possible. The best option for me is get me a whole noun 37 and I'll go over it, and I'll pick out the noun 37 I really need. And that's the way I usually work when I can. But if its ??? just get us a noun 37. And with pictures. That's probably the best thing. Digital cameras are absolutely terrific. Take a few pictures of where it's coming from. And then go 'I've got use for the things."

Interviewer: So that means that noun 37 is sending to the position 141 something. The position 141 is also sending to noun 129 pictures.

Participant: Observations. Recorded observations that you can make something out of. Recorded site observations. Here's the picture. It's usually got a time on the bottom of it. We usually know where it's come from. And out of the picture you can probably scale things. "Well, that's usually there too..." It's .25 mil, it's about half a mil across. So this digital image that's an observation, but it's recorded. Once you've got it recorded, you're in business.

Interviewer: This flow of recorded site observations is: Data, Information, Knowledge, O?

Participant: I'd say it would have to be Data. This is stuff you can do things... this is something you can reduce to a number. It also has observations, but it's reducable to a number.

Interviewer: The noun 37 flow to position 141 is what? The position 141 is the one that uses your expertise to produce the physical nature. So it's taker, but it's also the person who processes the noun 37. So the noun 37 to the sample taker is what?

Participant: It's embodied Information, isn't it? Shoved another qualifier on the front of it. He doesn't inherently know what it is. But the noun 37 embodies Information. Information, it might, and if you do an analysis on it, it's Data.

Interviewer: So where's the Data here?

Participant: Probably just noun 37 needs to be split into its an observation, it's a physical noun 37, and it's also what you do with that noun 37.

Interviewer: Let's look at flows here. We've got physical

Participant: A physical body that you can hold in your hand.

Interviewer: We would say that the physical body is an embodiment of Information?

Participant: Yes. And ultimately it can be reduced to hard numbers. The Data.

Interviewer: That reduction of hard numbers is performed by the position 141 and given to noun 129?

Participant: Yes.

Interviewer: So we've got another flow here. These hard numbers, what's the label of it?

Participant: A chemical analysis. This chemical analysis is... this is all hard numbers. Here is an analysis. This is real. This is hard Data. This comes down to numbers which hopefully will ??? properly. By standards, it's real Data.

Interviewer: We also have other things going from the noun 37 to the position 141, besides the physical body. Right?

Participant: Yes.

Interviewer: What are those?

Participant: Information. Simple as "this is the date I collected it, it was raining."

Interviewer: environmental?

Participant: environmental. It was picked off this noun 136. This noun 136, it had changed color. It's thats sort of observations.

Interviewer: These environmental observations are?

Participant: They're not Data. Information?

Interviewer: As a note, don't use exclusion. find some word that it fits even if we have to make up new words.

Participant: I think it's Information. it's creating a picture of the problem.

Interviewer: Do we want to say provenance instead of environmental observations?

Participant: Provenance is a slightly different sort of thing.

Interviewer: Is there a flow of provenance or is provenance a function of?

Participant: If you're looking where other people have taken noun 37, provenance becomes extremely important. If you know, the position 141, but there's another position 141 out here, it's like an expansion of this – noun 129 request and we do have examples of this. ... position 141, then there's another loop out here,... research doing things, and they're doing various things. And the provenance of any of their noun

37 becomes critical. Where it's like this is the noun 37 and I'm the position 141, provenance is not an issue because it's embodied in the observation. But if it was on the other side there, you'd like to know something of the provenance of the noun 37.

Interviewer: Therefore, where's the provenance flow? And what is it?

Participant: Provenance flow, it's probably a subset of this interaction here, but it is another point removed.

Interviewer: Point removed from what?

Participant: The immediacy of the noun 37? There's a flow there, and it's really Information.

Interviewer: Are there any other flows of Data, Information, Knowledge, or other coming from the noun 37 to anyone?

Participant: No, I don't think so. It's all being more or less channeled through this loop here. It's channeling through one entity. Sometimes this will be circumvented because they'll do it themselves, but really, in the way I'm thinking about things, it's this. This is the noun 129 role. I take in all this Information, distill it down to a story, and pass it on.

Interviewer: Is there a flow of story here?

Participant: Usually a story back this way so they know what they're doing.

Interviewer: So noun 129 sends to position 141 story?

Participant: Story. Why do I want to do this? What is the background thing? What is the rationality for doing this? This is not a stupid exercise.

Interviewer: So noun 129 sends to position 141, rationale.

Participant: Rationale. Why are we doing this?

Interviewer: This rationale is?

Participant: It's Knowledge. Because I know the people involved, I know that they like to have some idea of why are we doing this? Oh, [name] says this about it. otherwise it's shit. But if they understand why, they're likely to do it. If they don't understand ???

Interviewer: Even if the role of position 141 is in your own head, you still have to justify to yourself the rationale.

Participant: Why would I be spending money on this thing if it won't go? There's a rationale for myself and there's a rationale for other people.

Interviewer: Does the rationale go to anyone else, or is it a different rationale that goes to other people?

Participant: The rationale heading in that direction is probably much the same.

Interviewer: Is it the same rationale?

Participant: If you distill it down to its simplest form, yes.

Interviewer: So it's a simplified rationale.

Participant: It's simplified going this way, this is more detailed.

Interviewer: So noun 129 to position 141 is simplified.

Participant: I need this noun 37 because it will tell me something about this event. I need the Data to see what's going on.

Interviewer: Whereas noun 129 to noun 88...

Participant: It's detailed. "These things don't ever add up, I think we need a noun 37, and we need this about the noun 37" Same basic motivation, it's got a lot more detail.

Interviewer: Are they both Knowledge?

Participant: Is gut feeling Knowledge?

Interviewer: It's entirely up to you whether or not you want to categorize it as that.

Participant: I think it's Knowledge, because it's something your own makeup that says this is valuable. It's a form of Knowledge. It's Knowledge. Knowledge is the sum of experience, activity, that sort of things. "Why are we doing this? Well, in the past, this has worked. I've read about this activity in the paper. I've put 2 and 2 together and this is my rationale for doing this sort of thing. The Knowledge is the sum of what has been successful in the past, what has been a flop, why was it a flop, where did I go wrong? The formal learning.

Interviewer: Is learning different from experience?

Participant: That's a question for an education person. I think it is. ... The activities you've engaged in, previous experience, experimental sort of things. Learning, And what you've seen other people do.

Interviewer: What would that be?

Participant: It would be an observation. You've seen them doing this, and "jeeze, it worked. I'll try that." This is probably not out of the textbook, either.

...

Interviewer: We have the noun 97 DB over here. It looks awful lonely. Who plays with the noun 97 database?

Participant: noun 88...

Interviewer: noun 88 to noun 97 database... does what?

Participant: They, in my vision of things, they check the integrity of the Data that is actually going into the DB. In the sense that we've had a little bit of yesterday's noun 97 Information. That point's not right. Something's gone wrong there. There's some sort of quality review of the material that has actually long-term resident in the DB.

Interviewer: In that case, we have a flow... let's start with the first flow of: stuff goes into the noun 97 database. From where?

Participant: [list of redacted terms] ... the whole infrastructure. Computers, ... servers. And probably another aside to think about over my working life, there's been a very distinct change in the perception of those sorts of measurements. In the sense of – when I did my noun 91, measurements were were extremely expensive to do. A noun 74 cell was an extremely expensive piece of gear. They certainly weren't refined in the Noun 4 precision or accuracy that we have today. And the were extremely expensive. And you noun 95ed your experiments effectively to extract the maximum amount of Information. Maximum reliable Information from experimentation. That's probably what I'd go on with. That's part of the Knowledge DB. This is the experience, this ???. Nowdays, because a Noun 4, a DP setting you can get it for a thousand bucks. To wire it into the computer is another thousand. So two grand for a point. Ah, we'll go for more Data. You see noun 60s instead of 20 years ago, you might have a thousand points in a noun 60, now you might have twenty thousand. Y'know, we've got all this Data, but you think "are we doing anything with it?" One of my [old] jobs [somewhere else] was condition monitoring and melting. "When are we actually getting to the noun XX ? When are they going to verb XX?" So they put in a condition monitoring system. My argument was always: "You need to watch this data ??? times ago, it's your best Data." "Oh, no, no, we've got in a conditioning monitoring system. " That was the group 130. "And we can measure the Data." But no one looks at it. To me we have this philosophy, oh we'll just measure everything and hope that the analysis just trickles out some sort of sense. Whereas I come from an era where you noun 95 the experiment and you have a good idea of how you're going to extract it. You don't really much else. It's planning. You see that, we keep adding more and more noun 97 Noun 4. And you ask the question of [other company], "what are you going to do with this?" "Something." noun 88:"We might have a look at this thing." but in the context of the whole operating campaign or at least the last year, you don't know the reliability of the Information, you don't see the built in trends, and you're probably shooting in the dark. But you've got the Data points. Three quarters of the Data in the DB, you don't know how it was [collected]. Just a change in the way businesses perceive things now. But we've got this and we've got that. But you've got to be actively worrying the Data. Pulling out the stories, making sense of it. And it's a little bit upsetting, but that's the nature of the game. You'd probably hear much of the same story from [name]. Who: "Oh, we'll go back and have a look at the Data" Uhhh. You've got to start sorting it. They're trying to make the stories out of the Data. Whereas, me, coming here, would probably noun 95 the concept for the system and how we're going to extract the Data, what we want to do with it, and you start from times of your own. The focus is on getting the position 58ing, and once the position 58ing's there, we'll be right. No, no, it doesn't quite work that way.

...

Interviewer: The noun 60, Noun 4, and infrastructure send to the noun 97 DB...

Participant: Data. This is hard numbers.

Interviewer: This Data, these hard numbers, do we have a label?

Participant: noun 97 variables. This is a measurement of Noun 12, or noun 74, or flow, or concentrate.

Interviewer: Some variable of the noun 97. These noun 97 variables, they go into the noun 97 DB. Does anything else, in a general sense, go into the noun 97 database?

Participant: There are probably operational items which do go in. Probably from switches, more than anything else. Start a cast. A binary noun 111, it's on or off. Bang, that was the commencement of casting.

Interviewer: Switches send to the noun 97 DB what?

Participant: Status Information. Casting, not casting. Power on, power off.

Interviewer: Is this flow of status Information, Data, Information, or Knowledge?

Participant: probably Data and Information. Some of it are numbers like ladle weight for instance. That's a hard number. Start a cast, end a cast, that sort of thing is Information. It's on or it's off. It's going to be a combination of hard numerical stuff, quantities, masses. And the Information is we are in this state, we are not in this state. It's just, tick when we're not casting, not running noun 28. And probably how many bowls of noun 28. That will probably be matched up with hard Data. In the sense that there will probably be a chemical analysis.

Interviewer: noun 97 variables or something different?

Participant: Something different.

Interviewer: Where is the chemical analysis here?

Participant: It's probably another one of these noun 37.

Interviewer: From noun 37 to noun 97 DB?

Participant: Yes. Via the laboratory.

Interviewer: So we've got noun 37 to lab. And we've got lab to noun 97 database?

Participant: Yes. And that doesn't come through this loop, it's a completely different loop. It's another group of position 90s out there.

Interviewer: And they just input their things into the DB.

Participant: Yeah, they'll have "we've sent the noun 37 to the lab, this is its number, bang blam bang." All that is not observational, the noun 37 is a physical entity, the Data that comes from that noun 37 is Data. There's no observation.

Interviewer: noun 37 to lab, this flow?

Participant: is a physical noun 37.

Interviewer: Would you say the physical noun 37 embodies anything?

Participant: It embodies Knowledge. It embodies the actual chemical composition.

Interviewer: And that, you would say, is Knowledge?

Participant: That's Knowledge.

Interviewer: So there's a flow of Knowledge from the noun 37 to the lab

Participant: Yeah, in the third party. The noun 37 inherently contains Knowledge. Or is it Data?

Interviewer: You tell me.

Participant: Data. It's Data. That will come down as a hard analysis and stuff.

Interviewer: The noun 37 to the lab sends Data. The label for this Data is?

Participant: Physical noun 37. And it embodies Data, hard numbers.

Interviewer: Versus the physical body to the position 141 which embodies Information.

Participant: Information, yeah.

Interviewer: The lab takes this physical noun 37, the data flow of physical noun 37,

Participant: Is extracted by the lab, and the lab sends to the noun 97 DB, the embodied Data.

Interviewer: Is it a data flow?

Participant: I'd say it's Data flow. They're not doing anything with it. It spits out of the noun 133 or the the spark. Bang, how much of these. No one makes an interpretation, it' just a string of numbers.

Interviewer: The lab sends Data to the noun 97 DB of what label?

Participant: noun 37 chemistry.

Interviewer: The noun 97 DB takes these three things in, are there any other major components that it takes in.

Participant: Apart from the time, no.

Interviewer: Time. Does it take in time from something different?

Participant: It is probably a timing against its own timestamp. When something arrives, noun 97 Information, it arrived at this time. A clock entity. 24 hour clock starting at midnight, standard noun 97. Anything happens, it starts at midnight. Why? that's 00:00.

Interviewer: The clock sends to the noun 97 DB a timestamp.

Participant: Data.

Interviewer: noun 97 DB. Now there's interaction between noun 88 and noun 97 DB. The noun 97 DB sends to control, what?

Participant: DB sends to noun 88 Data. It is interrogated and it downloads number streams.

Interviewer: So it sends a flow of Data to noun 88 of what?

Participant: Numerical Information. noun 97 variable values.

Interviewer: What do we want to label that flow?

Participant: I would say Data.

Interviewer: So it's a Data flow of Data?

Participant: You're getting me confused. It's a data flow of hard numerical Information against predefined variables. So this is a ??? it's its value.

Interviewer: Do we have a more pithy name than hard defined Information against predefined values?

Participant: Just call them noun 97 variables.

Interviewer: We have noun 97 variables here. Are these noun 97 variables different from the noun 97 variables that noun 88 gets?

Participant: No, they will be the same throughout.

Interviewer: What about all this extra stuff?

Participant: I would call that a noun 97 variable, in the sense that if you're doing a chemical analysis, hot metal silicon, it's like a noun 97 variable. So they may not be instrumental values, but they would have a defined tag to define them. And it's a number against the tag. In the same sense as the tag.

Interviewer: Can we put an adjective before either of these things? Like stored or current?

Participant: The DB goes back goodness knows how many years. So it's effectively stored and stored for some time.

Interviewer: Do we want to say historical?

Participant: Historical.

Interviewer: Historical noun 97 variables go to noun 88.

Participant: From, from. Current time back. Then we can pull them out, and that goes back to noun 88. And it can also be interrogated by noun 129.

Interviewer: Therefore noun 88 and noun 129 need to be able to send stuff to the noun 97 DB in order to interrogate.

Participant: Yes, that's the program commands. I want to see this string for this time period of these variables. It commands...

Interviewer: These program commands are?

Participant: Information requests.

Interviewer: Information requests is the category?

Participant: Yes. You send it a command and the command is to fill an Information requirement.

Interviewer: So noun 129 and program control send an Information request to the noun 97 DB. The noun 97 DB then sends historical noun 97 variables to both, back.

Participant: whomever's requesting it. ... Just on here, you sketch plan there, you've probably got from between ten and a hundred people interactions. people, man-machine interactions. You're looking at ten people. noun 97ing, effectively on one shift, 30 people. the side loops, we're talking generalities, not time based.

Interviewer: What else do we have here?

Participant: I think we need a bigger sheet of paper. There's not much more I think you can talk about as a separate entity. This loop here, there's probably four strings to it. You can talk about noun 97 noun 37, you can talk about raw material noun 37, but I think they all fall into the same sort of concept. And there's multiple streams for these sorts of things. Probably the next step is how that comes forward and that comes forward to the business given that they're part of the business , these are one step behind the business . The barriers is, you can probably put it through there. For them, that's the barrier. For noun 129 the barrier has come back, because you've got this communication loop around here.

Interviewer: noun 88 to noun 129.

Participant: This is where it starts to get hazy. And it gets hazy. There's probably line group 130, which is the operational superintendent. His Position 22s

Interviewer: Everyone who's doing work on the line

Participant: On the line, yep. Who supervise the noun 97, who probably have a commission to see the noun 129 of the noun 97. This is the maintenance and money. This is getting up to the dollar signs of the business . The next one up here is group 142. This is ... line management would be [name]'s superintendent. Raw material position 92. The group 142 would be like general position. Very very much removed. You've got to summarize it for them. When you go here, it's a one page summary, and up to here is a one paragraph.

Interviewer: Because they've got so much...

Participant: They've got so much, and they're distilling down...

Interviewer: Any other big big entities here?

Participant: I don't think we'll go as far as the shareholders.

Interviewer: They go money...

Participant: Yes, but I think there are some shareholders that care about the businesses they invest in. I've invested in this business for

ethical reasons. So I'm interested in what they're doing. Yes, money is probably the root cause of it. People do carry interests about the business, probably some of the interest is "can I get my investment out of this?" and some of them, yes, I'll invest in this business because of ethical reasons. Some of the noun 129 work does filter through. This is from experience, some of this work does filter through into annual reports. It was greatly filtered through, I might add, but it does filter through, and it can happen.

Interviewer: Here's a question. Are the secretaries of any of these bodies a different entity? Or equivalent gatekeepers?

Participant: Yes. They are usually the access point to a lot of these people. Especially when you get up here. And I've found that... my access to [person]... the access to him is through his secretary. Because if you go straight to him, whoosh, into the aether. Gone. No response to any e-mail. But you ring up [name], I need to do this, "right [name], Ok, when do you want to see [name] sometime on this day, this day, or this day?" So there are gatekeepers for the very senior levels of group 130.

Interviewer: Are there line secretaries?

Participant: Not in the [company] organization, not in its current form. They can make your life easy or they can make it damned difficult.

Interviewer: Flows.

Participant: There's a big one here. This is the day to day reporting, direction...

Interviewer: So noun 88 sends reporting to line group 130.

Participant: Yes, they report on a day by day, problem by problem basis.

Interviewer: What shall we label that as? Are there two flows?

Participant: It's a day by day. Day reports. "And we had this problem overnight, and this abnormal report, and I don't know what the reason is. And the noun 97 is ticking over. [elements] are down. We're having problems with the Noun 21." It's all that sort of Information, Data. Some of it's numerical, some of it's observations.

Interviewer: So it's Information & Data?

Participant: Information & Data. With some interpretation.

Interviewer: And the interpretation is the Information?

Participant: Probably the interpretation is Knowledge.

Interviewer: Is that a separate flow?

Participant: They're all in the one. Because these are usually [short term hires]. "I'm in this job for a couple of years and I don't really understand what I'm doing." These blokes, these are the long term campaigners. They carry all of the sort of history, the Knowledge, what's been done in the past, sometimes we get it wrong. Most of the time it's going to be pretty right.

Interviewer: noun 88 sends the day to day reports to line group 130. These reports are Information, Data, and Knowledge.

Participant: Yes. And a whole host of abnormal event reports.

Interviewer: Separate flow?

Participant: They're all embodied in the day to day Information that's going across. What do we need to do? What are we going to get roasted on? It's all this mass of stuff.

Interviewer: It's this embolism of reports. Line group 130 to noun 88?

Participant: Are usually directives. What we would like to happen. Where do we need to push the noun 97? Why aren't we… what's the situation in the Noun 21 salvage questions. And it's directives. And in very rare cases there may be Information coming back.

Interviewer: Directives are a flow of?

Participant: Requests, Information?

Interviewer: Information?

Participant: Information. Someone says "I want this to happen in the … noun 60."

Interviewer: And you say that that statement is Information?

Participant: In your classification scheme, sure. Mine would be a request. "We want you to do this. I want you to have a look at this problem."

Interviewer: and this is a request that is neither Data, Information, nor Knowledge.

Participant: No, they're requests for Information.

Interviewer: in all of this, are there any that we need to change to match your classification scheme/

Participant: Probably only the one directing the position 141. But remember that request comes with a rationale.

Interviewer: So which one needs to be turned into a request?

Participant: Simple rationale.

Interviewer: So that's Knowledge and request?

Interviewer: That's probably a request + Knowledge. "I have a reason for asking that, but I'm asking you to get a noun 37." Yeah, it's a request. You might almost say that the noun 97 database is requesting. "I've got this hole in my DB, where's the Data? Send me this number."

Interviewer: Who is the DB requesting?

Participant: noun 97 Noun 4 or… It'll send out flags if it….

Interviewer: So Noun 4 get a request flag?

Participant: Yep. I need a number to fill this box in the DB.

Interviewer: Fill request?

Participant: Entry request. You might also get the same from switches. It probably is also interrogating switches. Switches are also effectively a request.

Interviewer: This request, does it have a different label or is it the same flow?

Participant: The same sort of flow. Say the same flow. We're getting too complicated.

Interviewer: Over here, day by day reports, Information+D+K. Line group 130 sends back directives which are requests. It also sends back questions? are they directions?

Participant: they're questions too.

Interviewer: Is that a separate flow?

Participant: I think it probably is. "I need to know all about this." as distinct from "You will make sure that this is being done." "I would like to know about this." slightly different.

Interviewer: These questions are Data, Information, Knowledge, Request, O?

Participant: I'd say it's just request. There's probably some Knowledge involved in making that request, but it's not transmitted in making that request. Because the link to here: there's something going on here that triggers some of these requests.

Interviewer: So let's model that.

Participant: This is starting to get... this is conjecture. Beyond the simplistic bits here and to the development, this, you can only surmise what is going on.

Interviewer: So you believe that group 142 has a flow to line group 130, via secretaries or direct?

Participant: Direct. Because these meet the noun 97 people, usually once a day if not more. These people will probably have a meeting once a week or they may do the walk around and actually go and see people. That's getting rarer and rarer too.

Interviewer: group 142 to line group 130, what's the flow there?

Participant: I think there's flows in both directions, reporting on condition, "this week we made X number of tons, our ... production rate was such and such," reporting. Usually about weekly is my sense. "You can't see him, he's going to the weekly meeting, or gone to a business meeting."

Interviewer: So weekly reporting is a flow of?

Participant: Probably more Information than anything else. The Data has dropped out of it. We're now starting to just give the summary. Yes it is numbers, but we've reduced... it's aggregation. I've reduced ten

261

thousand individual data points into one number. All the analyses ... and there's probably three or four hundred analyses. "The average fill rate was such and such. Our target was such and such. We're below target, we're above target. Why? Why?" It is now Information. The Data, the hard numerics, is gone. That number embodies the Information. Slightly probably different sense of...

Interviewer: but it's an important difference. group 142 to line group 130 sends what?

Participant: Probably the requests again. And the high level business Information. Request: I want noun 143 to be instituted in the noun. noun 143 is a new managmenent in [company]. Basically it's "no paper on your desk at the end of the day." You have a clean office, everything in its place. Have you ever seen a technical person that isn't mounded up with paper? Me, I want to walk out the end of the day: "That's where I got to this day."

Interviewer: It's an external memory.

Participant: Bingo. I sit down. Alright, neurons start firing, we're back into it. This probably comes from the phone center approach, where you don't have the same desk every day. Wherever the phone is, that's where your job is. You have a clean desk. I think that's where it comes from. "You will institute noun 143 at the mine site" It's a directive, it comes down, it has to be done.

Interviewer: So basically what you do is you have a desk for management then you have an informal paper center?

Participant: we have another place where we actually go and work.

Interviewer: High level business directives? What would you use?

Participant: Directives that are coming down from further up the pyramid. There is also the Information. That would be the business Information. Kind of the last month, we've broken even. We've made a profit, we've made a loss, this has been the lossmaking area, so it's that ethereal business Information. And the people down, once you get beyond the line position 41. "What's that all mean? What does it mean on a day to day basis?" Very very hard to see.

Interviewer: This high level business Information is...

Participant: I think it's Information. Is it Knowledge? No, I don't think it is. It's just got an immediacy to it. And people who just basically do what they're told. The board has said this, we will do this. I don't care how you implement it. I want want this done. It's Information and requests, the Knowledge is probably very very subsidiary. ...

Interviewer: other flows?

Participant: noun 129 to, we talked about this one.

Interviewer: Are there ones besides that?

Participant: noun 129 which is where I am, yes, I do have a line to there in the normal course of business . And in my current sort of job, I have

links to here to. Bing, bing, usually, because to see him, I've got to go through [name].

Interviewer: noun 129 to position 144 sends what?

Participant: I send, it's a distilled version of Knowledge, Information, and Data.

Interviewer: So you send a distillation...

Participant: I send a distillation embodying all of those things.

Interviewer: Distillation of what though?

Participant: The Data. I'm telling a story about...

Interviewer: We're going to say story.

Participant: A story, a specific problem, event, a series...

Interviewer: This story is Data + Information + Knowledge?

Participant: It's Data, Information, and Knowledge. It's built up from my experience. It extracts Data, or it has a Data component, and I'm reducing it to some sort of Information. But they're not independent.

Interviewer: Does line group 130 send anything to you?

Participant: Yes, it also sends directives as to what it... "I like this piece of Information, can you do more?"

Interviewer: Research directives?

Participant: Yeah. We have this problem, can you look at it? That sort of thing. They don't know what the solution is or what you can do about it. "That report, can you do a photocopy of that too, and can you send it to someone else? Has this position 41 been included in the circulation of this?" They are, what you say, organizational directives and directions in which way to produce investigations. "Yes, we like that idea, can you do any more." That's the one you want to hear and the one you hear the least of.

Interviewer: These organizational and research directives are...

Participant: Usually requests. And sometimes there's Information.

Interviewer: Is it Requests + Information, or just requests?

Participant: Requests + Information but requests dominates. Requests > Information. I think that's a fair summary.

Interviewer: noun 129 to position 145..

Participant: is a request for an audience. "I need to talk to [name] about such and such" or "[name] wants to see you."

Interviewer: So they are sending back to you a request for audience as well?

Participant: Yes.

Interviewer: So it's a bidirectional flow.

Participant: Bidirectional, yes. But it's only to get to see them. There's no Knowledge, Information, or Data, in that statement.

Interviewer: So it's a request flow for...

Participant: a request flow for audience.

Interviewer: So it's audience and type of flow is request. Then position 145 to group 142 send what?

Participant: They usually look at their Microsoft diary and see where you've got to get.

Interviewer: So the position 145 send the position 130 their schedule.

Participant: "You've got a gap there, [name]" "Yeah, I suppose so." It's that sort of...

Interviewer: this schedule is...

Participant: Information.

Interviewer: group 142 to noun 129 sends?

Participant: I think it's like the position 144. It's usually requests and Information. Usually the Information flow is large to them. So we're communicating what we're finding...

Interviewer: Findings?

Participant: Yep, yep. Where we are in the project. It's a distillation, it's the story.

Interviewer: These findings are?

Participant: Primarily Information.

Interviewer: Then group 142 to you sends?

Participant: Probably requests and Information. Requests: "Look, this looks alright, we don't want you to do any more on that, we want you to this." Or, alternatively, "Did you know that someone else is doing this and, you really should talk to them." So it's a blurring of the two, but usually, if they're happy with which way you're going, whether they support it or keep going. It's the Do/Don't Do type...

Interviewer: So this flow is labelled... direction?

Participant: Direction and Information. "That's very interesting but."

Interviewer: Direction?

Participant: Yes, Direction. And the Information is: "Have you considered this?"

Interviewer: but that's a function of direction?

Participant: Yes "We would like you to do this." Effectively it's Information, but really it's probably another form of request to do something.

Interviewer: Shareholders.

Participant: There're so far into the aether. But really it's feeding through this line here. I think the Information is that way, as far as the technical is... If you have a look at the report... you know ...

Interviewer: Report?

Participant: Annual reports. Have a look at the [company] annual reports and you'll see there's a little spiel about their technology side of things. ... It's part of our corporate profile and it gets a couple of pictures. In the report it's usually half a page. It's saying we're doing these sorts of things, we're looking at these problems. We're trying to be proactive in terms of the environment. All those sorts of things.

Interviewer: Do the shareholders send anything back?

Participant: Only if the business goes bad.

Interviewer: So it's just a one way flow.

Participant: Question, has a shareholder said anything of a technical nature that has registered with the board? I don't think so. Only if a noun 97 is going really bad. ... But it's probably only a thousandth of a percent of what the shareholders are communicating about. Institutional shareholders, it's about the state of the business . But sometimes things will pop up. But it's only a small dotted line.

Interviewer: Technical question, concern?

Participant: Someone's see something, someone's asked the question. Some of those things might be via investor relations, and you really don't know. But every organization has an investor relations.

Interviewer: It's?

Participant: They're probably trading on Knowledge,

Interviewer: So they're sending Knowledge?

Participant: No, they're using their Knowledge to generate a question/request.

Interviewer: So therefore, it's a request?

Participant: More a request.

Interviewer: Do we have any flows that we're missing?

Participant: At this stage, I don't know.

Interviewer: The sum of rationale, activities, experiences, learning, and observation create Knowledge?

Participant: Yes.

Interviewer: Let's have Information, Data, Knowledge, Request. What goes into Information?

Participant: Information is, It's the technical story. It's usually is... business Information, what's the state of the business ? In a technical sense, Information is a distillation of these entities here.

Interviewer: Therefore it distills Knowledge?

Participant: It takes Knowledge, Data, observations...

Interviewer: Is that separate from Knowledge observations?

Participant: Observations build into Knowledge.

Interviewer: So Knowledge + Data... is there a function here that turns it into Information?

Participant: usually my brain.

Interviewer: Your brain doing?

Participant: Generating the Information is like a jigsaw puzzle.

Interviewer: You take what and combine it with what?

Participant: You take Data and Knowledge and merging the two to generate Information. And it's an incomplete set. These are incomplete sets.

Interviewer: So you have Knowledge and Data create Information. Is it equality or directionality. Can information be unsausaged?

Participant: Sometimes yes. You want to deconstruct the Information to get back tot he original concepts. So I take this Information and plot it. There is equality.

Interviewer: Whereas with Knowledge you take this and go.

Participant: Yep. Given that these are adding to it all the time.

Interviewer: This sigma is adding to Knowledge all the time?

Participant: Yes. I've found out this, and I'm going.

Interviewer: Data is?

Participant: Usually one off.

Interviewer: This, but no sigma?

Participant: Data is the hard numerical material to my way of thinking. It's the experimental measurements, it's the noun 97 measurements. It's the chemical analyses. It's the things that I can get which are hard. A number which has a physical association, or a variable association. It represents some property.

Interviewer: Does Knowledge, Information or Request have any function in Data? Or is it just an experimental measurement?

Participant: It's just the experimental measurement.

Interviewer: We have experimental measurements. We combine experimental measurements with Knowledge to create Information. We have these perceptive things which are aggregated to create Knowledge.

Participant: Yes. And the request is "We want you to look at these things." The request brings into play all those factors.

Interviewer: So a request is a one way function of Information, Data, and Knowledge?

Participant: Yes. You are asked, or you ask yourself, "I would like to know about this." And then the higher factors come into play. "How do I get an answer out of this problem? "and it's the jigsaw puzzle. Consider a 500 piece jigsaw puzzle. For most of our games, we're making the story on about 25 pieces of Information. In a really really good situation we might have 100 pieces to actually put the story together. But usually, because it's in a business nature, we've only got 5

Interviewer: Would you say that this accurately corresponds to what you consider the ontology of Data, Information, and Knowledge?

Participant: I think it's pretty close.

8 Personal Reflections

This chapter presents an analysis of the evidence gathered by my interviews and surveys. I will explore whether or not my methodologies have demonstrated different conceptions of data and if they are capable of exploring unique realities of data in the first place.

The initial discussion is an examination of my research paradigm and approaches, looking at the philosophical basis of my conclusions. I employ an abductive[*] approach to iteratively define universes of interest, rather than applying inductive or deductive approaches.

I am testing two questions of interest, which guide the course of my analysis: whether people have different realities of data and whether my methodologies can discover someone's personal construction of data. By guiding the abductive research process with these two questions of interest, I may be able to limit my conclusions to the most simple and credible story possible from the evidence.

Using the questions of interest as a guide, I will iterate over the results of a recursive analysis of every interview and survey, looking for evidence of different conceptions of data by looking at their statements about data and the relationships they suggest between data and information and data and knowledge. The evidence suggests that I have successfully answered both questions of interest by demonstrating the existence of three realities of data that emerge from both the surveys and the interviews. The success of this research suggests that my dissertation offers a foundation for future studies and

---

[*] An abductive research strategy features a process of rapid hypothesis formation based on inconclusive evidence. While this shares similarities to the inductive strategy, the approach looks more to suggest boundaries to a region of possible answers and less towards finding (inducing) the governing rule of all the data collected. It is the approach used when there is insufficient evidence to even start guessing [56 and 57].

can be deemed successful.

8.1   Summary of Reflections

My personal reflections found the following constructions of data. While these summaries should not be viewed as "authoritative" over the "raw" data, they indicate my personal beliefs as to the interpretation of that data and the final result of the recursive analysis.

- Pilot interview:

  - Data is a menial-interpersonal communication.

  - Data is a menial-technical communication.

  - No hierarchy specified.

- Interview II

  - Data is an electronic container for human-produced information and knowledge.

  - Data does not transform.

- Interview III

  - Data is a subjective observation.

  - Experimental data is not data, an exception that allows for quality and less subjectivity.

  - Sometimes data is a communication.

  - Data can create information and knowledge and information is structured data.

- Interview IV

  - Data is a subjective, discrete observation with specific provenance and reliability: "the ability to specify 'what, where, when' would define a piece of that, what the measure of it was, and its location at the time."

  - Representations of data are not data.

- Data participates in a cyclic hierarchy, being filtered by knowledge and being contextualized into information.

- Interview V

  - Data are the fundamental relationships of matter, which are themselves measurable and observable.

  - Data can become information and knowledge by adding context and abstraction.

- Interview VI

  - Data are facts created from objective sensors and experiments.

  - They can be input to mathematical models, contextualized with information, and generalized with knowledge.

  - Standard hierarchy: data produce information, which produces knowledge.

- Interview VII

  - Data are numbers: "Data is a multi-dimensional collection of data points. A data point is a number." Data points are created from experiments.

  - Standard hierarchy.

- Interview VIII

  - Information as container for data and knowledge.

  - Data are factual representations of measurements.

  - No hierarchy.

- Interview IX

  - Data are numbers, no requirement of objectivity: "Look, you can put them in a list. There is a defined quantity. They can be grouped, they can be ordered or whatever. They probably are data. You can put them in a spreadsheet."

  - Cyclic hierarchy: Data, interpreted with knowledge creates information. Information analyzed creates knowledge.

- Interview X

- Data are contextualized "hard" numbers as observed representations of reality.

- Data as apex of ontological hierarchy.

- Tuomi-reversed hierarchy.

8.2   Summary of Survey Analysis

- Survey I

  - Working database labs of university.

  - Data is a symbol without meaning.

  - Standard hierarchy.

- Survey II

  - Participant works for a defense department processing information.

  - Data is an unanalyzed sign.

  - Standard hierarchy.

- Survey III

  - Participant is a manager of a consumer electronics repair workshop.

  - Data is a statement.

  - Hard to classify.

  - Standard hierarchy.

- Survey IV

  - Participant is a social entrepreneur.

  - Data is an observation.

  - No hierarchy.

- Survey V

- – Research scientist, incomplete.

- – Data are observations of the world.

- – Standard hierarchy.

- Survey VI

  - – Participant is a service desk employee for telecoms.

  - – Data are structured records.

  - – No hierarchy.

- Survey VII

  - – Participant is corporate strategy manager.

  - – Data are observations without interpretation.

  - – Standard hierarchy.

- Survey VIII

  - – Participant is counterterrorism analyst.

  - – Data are objective records of activity.

  - – Standard hierarchy.

- Survey IX

  - – Participant is a housekeeper.

  - – Data are factual scientific observations.

  - – No hierarchy.

- Survey X

  - – Participant is a senior software architect.

  - – Data are electronically stored observations.

  - – No hierarchy.

- Survey XI

- – Participant ambiguous.

- – Data are objective, precise facts.

- – No hierarchy.

- Survey XII

  - – Participant is a research engineer.

  - – Data are specific observations of phenomena as well as stored bits on a computer.

  - – Standard hierarchy.

- Survey XIII

  - – Participant is researcher and modeler.

  - – Data are numbers without context.

  - – Possible hierarchy, survey unclear.

- Survey XIV

  - – Participant is a developer of an in-house optimization application.

  - – Data are facts without context or intrinsic meaning.

  - – Standard hierarchy.

- Survey XV

  - – Participant is a SIGINT analyst.

  - – Data is a small, measurable, description of the world.

  - – Hierarchy of precision.

8.3 Summary of Interview Analysis

- Pilot interview:

  - – Data is a menial-interpersonal communication.

  - – Data is a menial-technical communication.

- No hierarchy specified.

- Interview II

  - Data is an electronic container for human-produced information and knowledge.

  - Data does not transform.

- Interview III

  - Data is a subjective observation.

  - Experimental data is not data, an exception that allows for quality and less subjectivity.

  - Sometimes data is a communication.

  - Data can create information and knowledge and information is structured data.

- Interview IV

  - Data is a subjective, discrete observation with specific provenance and reliability: "the ability to specify 'what, where, when' would define a piece of that, what the measure of it was, and its location at the time."

  - Representations of data are not data.

  - Data participates in a cyclic hierarchy, being filtered by knowledge and being contextualized into information.

- Interview V

  - Data are the fundamental relationships of matter, which are themselves measurable and observable.

  - Data can become information and knowledge by adding context and abstraction.

- Interview VI

  - Data are facts created from objective sensors and experiments.

  - They can be input to mathematical models, contextualized with information, and generalized with knowledge.

  - Standard hierarchy: data produce information, which produces knowledge.

- Interview VII

- Data are numbers: "Data is a multi-dimensional collection of data points. A data point is a number." Data points are created from experiments.

- Standard hierarchy.

- Interview VIII

    - Information as container for data and knowledge.

    - Data are factual representations of measurements.

    - No hierarchy.

- Interview IX

    - Data are numbers, no requirement of objectivity: "Look, you can put them in a list. There is a defined quantity. They can be grouped, they can be ordered or whatever. They probably are data. You can put them in a spreadsheet."

    - Cyclic hierarchy: Data, interpreted with knowledge creates information. Information analyzed creates knowledge.

- Interview X

    - Data are contextualized "hard" numbers as observed representations of reality.

    - Data as apex of ontological hierarchy.

    - Tuomi-reversed hierarchy.

8.4   Summary of Survey Analysis

- Survey I

    - Working database labs of university.

    - Data is a symbol without meaning.

    - Standard hierarchy.

- Survey II

    - Participant works for a defense department processing information.

    - Data is an unanalyzed sign.

- Standard hierarchy.
- Survey III
  - Participant is a manager of a consumer electronics repair workshop.
  - Data is a statement.
  - Hard to classify.
  - Standard hierarchy.
- Survey IV
  - Participant is a social entrepreneur.
  - Data is an observation.
  - No hierarchy.
- Survey V
  - Research scientist, incomplete.
  - Data are observations of the world.
  - Standard hierarchy.
- Survey VI
  - Participant is a service desk employee for telecoms.
  - Data are structured records.
  - No hierarchy.
- Survey VII
  - Participant is corporate strategy manager.
  - Data are observations without interpretation.
  - Standard hierarchy.
- Survey VIII

- Participant is counter-terrorism analyst.

- Data are objective records of activity.

- Standard hierarchy.

- Survey IX

  - Participant is a housekeeper.

  - Data are factual scientific observations.

  - No hierarchy.

- Survey X

  - Participant is a senior software architect.

  - Data are electronically stored observations.

  - No hierarchy.

- Survey XI

  - Participant ambiguous.

  - Data are objective, precise facts.

  - No hierarchy.

- Survey XII

  - Participant is a research engineer.

  - Data are specific observations of phenomena as well as stored bits on a computer.

  - Standard hierarchy.

- Survey XIII

  - Participant is researcher and modeler.

  - Data are numbers without context.

  - Possible hierarchy, survey unclear.

- Survey XIV

- Participant is a developer of an in-house optimization application.

- Data are facts without context or intrinsic meaning.

- Standard hierarchy.

- Survey XV

  - Participant is a SIGINT analyst.

  - Data is a small, measurable, description of the world.

  - Hierarchy of precision.

8.5   Interviews

To avoid weasel words such as "participant seems," every interview description will be written in strong language with categorical statements about the nature of data, information, and knowledge. These are my observations and interpretations from what the participants stated as refined through recursive analysis, not my own philosophies. Further testing is required before these statements can be codified into any kind of statement covering many people's constructions of data. As one of the end products of my methodologies is a person-specific understanding for the interviewer to the participant, I feel that these analyses will capture that specific understanding and already provide a useful tool for any designers of databases or interfaces who want to use my methodologies.

8.5.1   Data as Communications

One of the three constructions of data is that data are encoded signs. Signs that exist on paper or electronically are passed between humans who may then interpret those signs as data for information or knowledge encoded into them. In this construction, data is

a technological artifact rather than something generated and manipulated by humans. Information and knowledge are passed using data, but are not of the same fundamental nature. The relationship is a container metaphor rather than hierarchical.

The few interviews expounding upon this construction suggest that significant research is necessary to explore its tenets. Beyond that, there is some evidence that this construction may be held as a secondary way of understanding data when dealing with technological systems[*] rather than as a fully-fledged construction of data of its own.

8.5.1.1 Interview I

Technical pilot interview, exploring a Facebook game. Participant holds two realities of data: Data as a menial interpersonal communication and data as a menial technical communication.

The first interview I conducted was a technical pilot, assessing the methodology alone without exploring business specific concepts. In this interview, the participant and I explored the leadership dynamics of a Facebook Game. More details about the SDFN diagram and its analysis can be found in chapter 3, in the case study. This first interview will offer the raw quotes, as well as the recursive analysis appropriate to the section in which the quote appears, thanks to explicit permission by participant.

Data, information, and knowledge relate to each other. Data is a menial category between these three. Data exists only as itself, without reliance on anything else. "Anything else" here means external signs which frame, or contextualize, data as well as indications that the data requires judgment or discernment on the part of the recipient. I believe that the participant considers the application of context to be something required of ambiguity or importance, and therefore of increased importance.

---

[*] This pattern was also apparent in some of the surveys like Survey I (p. 308) and VI (p. 316).

Data is not a function of expertise or knowledge. A critical component of data is its lack of necessary arbitration: data is intuitively obvious, objective, and unimportant.

> Participant: It's almost like discrete packets.
>
> Interviewer: So, packets of?
>
> Participant: Data?
>
> Interviewer: Commands?
>
> Participant: Like a command. Saying: "I want to do something. I want rez the fort now."
>
> Interviewer: So is it OK to label this as commands, or does it encompass something other than commands?
>
> Participant: It does, sometimes...
>
> Interviewer: So commands and..
>
> Participant: No, that's true. It's pretty much just commands. Like I would send it I could do many things but they're all related to telling the game to do something. I don't call it knowledge, because it's like it's just, maybe they're just talking about context thing. It's something fairly simple, discrete. It's not open for arbitration or anything like that. It doesn't require arbitration, just "do it."

Two iterations of analysis produce:

> Data is discrete. Data does not need context nor arbitration. Commands to software are data. Data is simple, it can be executed on its own merits.

For example, instructions for action are data, both those that are explicit instructions to software agents and implicit orders as "status updates." In the game sense, a status update is a message sent out from administrators to the players "advising them" on the status of the structure players are defending. This "advice," however, is intended primarily as a mobilization for action, as positive activity is needed to create the structure.

Communications are information. Communications are movements of subjective statements about the world encoded in the transmitted signs. Communications must have context. This categorization of data is used to delineate between communication/information and data. Beyond a generalized context, communications must also contain a temporal context: they have to be about somewhen, whatever other elements they

contain.

> Participant: So I'm telling them, I'm sending them information. That they can now war. That they can start.

> Interviewer: And so this is a what? Is this a status, is this a command, is this something else?

> Participant: It's information because it does require context. But it's like a status it's saying: "You can now war." We can start fighting.

Two iterations produce:

> Information can be returned by software. Status is a form of information. There is a temporal aspect to changes in status.

> Information requires context. Information provides options.

Expertise is knowledge, being the perception of options for action in information.

Knowledge confers the ability to be persuaded by communications as expertise is used to frame incoming information and thereby "understand" it.

> Participant: I don't know how to class their opinions. Not necessarily knowledge. It's almost like data. It's like, they don't really send me commands, not as in the concept that we talked about, commands in the past that do something. It's almost like data. They just tell me stuff. I'll say, "We're going to do this." but they'll say, "I think we should do this." I don't think I have to reply to it. It's like they're offering their opinion. But it's where we're distinguishing the different status. And then when we go to active clan masters. Now there's a number of different inputs: "Clan status, Info, FB." Same with all of these. They're in a group, they're getting the same as everyone else. But we'll have an extra line which will definitely be MSN as well ... which will be another input to them.

> Interviewer: So you're sending them MSN? And what are you sending over MSN?

> Participant: I don't know? I can call it anything, can't I? It's like knowledge? I guess? It really is knowledge that I'm sending back to them that way. Similarly, they'll send me back data and opinions via FB messaging. But they'll also have that subsequent line via MSN. It's that knowledge sharing. Even hilariously in the past, I've actually called some of these guys on the phone. ...

> Interviewer: Now what are you doing over this MSN & Phone?

> Participant: Well, that's the genuine sharing of knowledge. Knowledge of: "I think this; I want to do this; no, no, no, we can't do that." It's a collaborative sort of working together. It's a sort of consensus-building. That's where it's knowledge, because by sharing that knowledge you can build a common consensus which we can work around. Because if we don't share each O's opinions or experiences then we never build

the consensus to actually operate, otherwise you'd just be individuals because it, generally as a clan, we have to come up with a consensual view of everything..

Interviewer: Do we want to have consensus as another line? Or is the sharing consensus-forming?

Participant: Yeah, the sharing is consensus-forming. We exchange knowledge and out of that comes consensus. Even that knowledge is: "Do you consent to this?" It's not really a command but it's a knowledge.

Two iterations produce:

Opinions, lacking privileged status, are not knowledge. Activity provides context. Responses from active members are classified as knowledge. Responses from semi-active and inactive leaders are data through opinions.

Knowledge can be shared. Collaboratively sharing knowledge builds consensus.

There is no explicit differentiation between different kinds of data. There are examples in which data is either technical in nature (for example, a command to cause activity) or a communicative activity (wherein the communication was not produced as a function of expertise). Data is interesting only in the sense that it represents the mental activity of the communicator, not as any kind of thing on which to take action. The differentiation only extends to the comment that something is "sort of data" potentially representing an internal cognitive dissonance between the two ideas.

Participant: It's when you're talking to another person and they're passing their knowledge of what happened to them and what happened in the past to you. So you can either learn from your own experience or O's experience. It's really a mixture of that. In relation to knowledge gathering.

Interviewer: Would you say that there's any relationship between knowledge and data in this sense? Or is it just...

Participant: No, I don't... I think. I think the expression of data in this context, it's interesting actually. Now you've got me thinking. No, but yes. Yes, but no. Well, it is sort of, yes. I agree. I don't often express some things like certain relations to [the game] as data. but there definitely is if you stop and think about it, because. But yes.

Two iterations produce:

Knowledge can be gained from talking to other people, learning things through the game. Experiential Knowledge is what you get from doing

things. You learn from criticism, experimentation, failure. A different kind of knowledge is someone else's experience. Knowledge is privileged over data. Represents "knowing" over "observation.

There was no discussion on transformations of data. Moreover, no conclusions can be drawn from inference, as it is equally likely that no transformative framework exists such that the standard hierarchy is present. Each term seemed to be discussed on its own merits without reference to the Os. There are two constructions of data found here: data as a menial-interpersonal communication and data as a menial-technical communication. In both senses, data contains no privilege over information or knowledge. As menial-technical, data seems to be considered to be bits or human instructions rendered as bits: self-encapsulating orders for action. As menial-interpersonal, data seems to be expressed observations of the world without a knowledge basis: empty words to be corrected and brought in line with the consensus formed by more experienced people.

8.5.1.2 Interview II

The participant is an information worker. Data is a container for human-produced information. Data is the only thing manipulable by machines. Information and Knowledge cannot be transformed into Data, only contained.

The participant works with information systems. Data is a container for human-produced information. Information and knowledge relate to each other as a function of human interpretation, integration, and analysis. The main function of data is as a communicative medium of technology.

There is no hierarchy or relationship between data and information. Data is completely different from information or knowledge.

Data does not belong in the hierarchy of Information or Knowledge.

The results of a search are Knowledge, but are formatted and contained within Data.

These quotes indicate a strong differentiation between information/knowledge and a strongly physicalized data. Although information and knowledge are related, the relationship expressed seems to be a function of analysis and encoding.

Communications between people, as discussed in the interview, are clearly information.

> Information is non-physical. It can be a communication, a conversation, a story.

Knowledge, or personal experience, is *encoded* in information for transmission to other people. This ontology requires this encoding because knowledge is not directly communicable between people. Knowledge guides interpretation of information and, in the case of knowledge encoded in information, the recipient's knowledge must be used to analyze the incoming information-communication to produce new knowledge:

> Knowledge is associated with an experience. Information doesn't have to be read or analyzed. Knowledge must be analyzed and used.

Data is a container for human-produced information stored in physicalized entities such as books or computers:

> Data is a record that can record and contain Information. Records, Data, are merely physical entities.

As examples of data, the participant offers: "Physical containers are pieces of paper or e-mail. Digital containers are still physical." With no evidence for other interpretations. In a simple sense, the participant believes that: "Data is a bucket for Information and Knowledge." To restate: data is a semiotic representation of and container for content.

Data as semiotic container is a strong ontological stake. It expands beyond the simple data-as-bits construction that may accompany other realities of "real" or "raw" data. The universal application of data-as-container and information as root-of-meaning suggests that this construction extends beyond a simple communicative reality that can be used in conjunction with other constructions. Rather, it strictly defines a role for

data in the storage of signs and relegates it to a strictly supportive role with respect to meaning. Nor does this construction leave any room for data as observation or numerical representation.

8.5.2 Data as Subjective Observations

The next interviews explore the nature of data as a subjective observation. These constructions differ strikingly from the data-as-communications realities as they accord data a place in the knowledge/information/data hierarchy. There is evidence of Ackoff's hierarchy in interview III that contrasts with a more cyclic hierarchy in interview IV.

Present in both interviews is the acknowledgment that data is inherently subjective, that it is produced from observations made by people, and that it must be contextualized with information to provide useful guidance. Data's lack of inherent value due to its subjectivity requires that it be "filtered" so that only the most usable aspects remain.

8.5.2.1 Interview III

> The participant explored business dynamics of a meeting. Data is the basis of information and knowledge, and formed from observations.

This interview, exploring the interactions of a meeting, was one of my more theoretical interviews. While most of the other interviews focused on more technical topics, the communicative/business focus of this discussion provided an interesting counterpoint to the other interviews featured in the rest of this analysis. Observation is the fundamental component of data, allowing interpreted data to be useful information. Information becomes knowledge, with the hierarchy continuing beyond knowledge:

> Knowledge is probably not the top of the hierarchy. Universal truths,
> e.g. things called "Laws" in science, could be called wisdom (W).

In the interview, the participant began with a discussion of their opinion of the relationships between data, information, and knowledge. Later in the interview, however, they admitted to a change of their preconceptions due to the interview: "I consider [data and information] to be a hierarchy, but it's not very clear in my mind" is a comment made by participant at the start of the interview. Later in the interview, the participant observed that participation in the SDFN produced a change of understanding of data:

> Coming into this interview, I thought that the hierarchy was Data, Information, Knowledge with maybe Wisdom or something above that. Or something crazy like that. Way out there. During the meeting, I think I... with your... non-directional coaching, you helped me to see where things were probably Data, where those Data were used by Os to form Information. And then, pieces of Information, there's something higher above information.

Structured and organized data, as delivered for a persuasive argument, becomes information. Data can be transformed into knowledge directly, such as a scientist predicting movement with Newton's first law of motion. In a prediction of the future as information, the predictor has to take data (observations of something happening), and input them into knowledge (a scientific model). Interestingly, data also encodes information and knowledge as well as forming the basis of those two concepts.

Knowledge is the theoretical basis of understanding-of-the-world. Two people discussing a simple mechanical operation are communicating information, but the theory driving that discussion is knowledge. Furthermore, the jargon used in that discussion also represents knowledge, whereas a generalized description is "somewhere between data and information." Thus, an understanding of the state of the art and the status quo is knowledge of what is possible or effective to do. Operation of effective management practices is also knowledge, despite the management theories themselves not being transmitted. Because the knowledge of effective management inspired the communications to modify behavior, the communications themselves are knowledge, representing

287

the practice of those theories.

Information is data structured for a purpose. Structured, organized data as delivered for a persuasive argument becomes information just as an historical-relationship context is also information structured for purpose. Data becomes information by a person or machine collecting data and then "drawing something out of it." Just as information is a structuring of data, the participant suggests that knowledge may be a structuring of information.

Data is an observation. A listing of daily events is a data flow just as a progress update is data. Nonverbal cues between people across a meeting table, such as eye contact or arm movement, also constitute data. When one of the meeting participants aggregates the nonverbal cues to supplement persuasive mechanisms, the participant considers that aggregation to be data.

Another example of data is an intuitive insight, a "moment of clarity," into the fundamental nature of a problem. "Data gathering" is used in the context of an experiment differently than the term "data" alone, with few relationships in common between the two. "Experimental data", unlike normal "data", can have different levels of quality: "we can classify the data better if we can see it clearer." There seems to be little support for data as encoded information or knowledge.

The resolution of cognitive dissonance in this interview suggests that the SDFN can be used as a way to reconcile a participant's theoretical understandings of categories with their everyday use. This consequence, while useful in the exploration of philosophy, can also be useful as a practical way to resolve "silos" as it is a way to change a participant's

understanding of the items under discussion.

8.5.2.2 Interview IV

This participant engaged in multiple interviews, producing quite a lot of evidence and strong elaboration of their ideas. There is a cyclic hierarchy of data as observations, localized

> The participant is an engineer. Data are subjective observations filtered by knowledge and localized into information.

by information, and filtered by knowledge, becoming information and knowledge. Knowledge then directs the gathering of the next cycle of information. There are no obvious demographic characteristics to explain the strong subjective nature of this participant's views, though they seemed to engage more strongly with databases than did those of some of the other participants.

Data is used both in singular and plural senses, data can have qualities, and measurement generates a data point. Data relates to information and knowledge in a cyclic hierarchy[*], such that some assertions about the world (knowledge) form a context within which to view data. "Knowledge filters data, which produces information, which produces knowledge. The acceptance of new knowledge is predicated on extant filters."

One example of context modulating the nature of data is the understanding of a sensor's purpose changing the observer's interpretation of the data produced by that sensor. In this understanding, background education is a primary contributor to a piece of data's context.

Information's localization of data from the "study of the issue at hand" helps the recipient to understand what the data means. Information emerges from a contextualization of data: "In order to have context [within] which to understand data, background

---

[*] Reminiscent of Tuomi's ontology (page 28).

education (knowledge) and the study of the problem at hand (information) must be combined." Knowledge and information represent different levels of understanding, with knowledge formed by integrated conclusions versus the ordered and contextualized data of information.

Understanding in this case is an awareness of applicability. All data, information, and knowledge interact with mental models of the worlds. Locally true items are information, whereas knowledge informs models that are more global. Because the generation of knowledge is a function of incoming filtered data, we sometimes get conflicts in explanations:

> Explanation is contextualizing [models] for specific circumstances. Explanations should expand knowledge. Sometimes they contradict understanding, potentially due to cultural factors or mistranslation.

Knowledge filters for meaningful data, and data is contextualized from information into information. Contextualizing knowledge frameworks allow for an understanding of incoming data by situating that data within knowledge. The framework does so by providing a relevancy filter for incoming data:

> Information is more concrete – in the hierarchy. Data is concrete. I measured it via a defined process at a particular time. We might argue what it means, which is knowledge, but it is what it is. But the information ... I've reviewed a whole pile of things. Those are information sources, it's clear how I obtained them. They may be wrong, their provenance may not be certain. Data is specific, discrete. Information can be a collection of data.

It is the process of collecting and localizing data which turns data into information. This set of localized data is *contextualized*, which increases relevance at the same time that it increases potential error or subjectivity. Despite claims of concrete data, data in this construction must be subjective: it is filtered from human observations. However, the self-reflection of that subjectivity creates meta-data about reliability and thereby filters for useful data.

Data is a discrete observation. Discrete data means the ability to specify "what, where, when: would define a piece of that, what the measure of it was, and its location at the time." Data has a specific provenance and reliability. Data is treated as if it is atomic, despite the ability to perform different, closer observations of the same thing. The presence of the term "filter," one of the most important keywords that I found, suggests that data as observation is produced from any observation of the world and is thus a subjective construction, tied to the observer, rather than evidence of the world.

Meaning is produced through the negative feedback of knowledge. Data can be formed into data sets, which may be evidence of a somewhat local scientific language translating between purely subjective data and the "reliable" data that some scientists demand. Relationships between sets are information.

There is no link to technical data in this construction. Representations of data (bits) are not the thing. Categories of representations are meta-data and bit streams are merely representations of data. "When people can attach meaning to bit-streams, they can become information. Normally glyphs become information, but when they're just representations of data (like the 'A' signifying a part of the base pair in DNA) they're still data."

8.5.3 Data as Facts

Conceptions of data in this category tend to be more scientifically oriented, representing data as *only* the results of careful consideration and measurement. Here, data as facts orient the observer to a recorder-of-measurements, rather than as someone who must filter observations or interpret semiotic encodings.

This category of interviews is the largest, as can be expected while interviewing

a scientific research establishment; because of the scope of the interviews, I probably missed some significant nuances. Exploring the different constructions of data and fact from a philosophy of science standpoint could be extremely productive.

8.5.3.1 Interview V

The participant is a scientist. Data are measured observations, objective reflections of the world. Information is a product of human thought to contextualize that data.

In strong similarity to Interview I, the participant differentiates between what can be understood by consciousness and what is in the realm of machines. The participant is a highly theoretical researcher, in contrast with some of the more pragmatically trained participants. Data are specific and measured observations.

Data, information, and knowledge are ordered categorically in terms of precision. Information is an unconfirmed interpretation of data and knowledge represents understanding of the real world. The flow from specific to general is interesting, especially when considering the implications of the participant's conception of data.

Data are measured observations, with observations or raw measurements as "raw data," and interpretations of data to be data-as-"interpreted data." Experiments return data as numbers collected by various instruments, but data is also embodied in the world. Geometry and design of equipment and the operational set points are data. "Equipment geometry and operational set points are embodied data." The idea of data embodied in the real is a strong feature of the data-as-facts construction.

There must be an observer to manipulate things outside of data because, "[The] experiment itself is not a conscious entity; you cannot send anything but data to it." While it is interesting to consider information and knowledge as functions of consciousness, the distinction between data as embodied or observed seems to belie my assessment

of "data as measured observations." If data is allowed to be embodied in the physical universe, then to retrieve the data and thereby to understand the object is to perform some sort of structured and measured observation in the context of an experiment. The participant notes that data are "raw numbers and no context." This construction requires that data must be understood as the fundamental relationships of matter that can be observed and quantized.

The category of information trades precision for context. Information is an unconfirmed interpretation of data. There is a clear distinction between the objectivity of data and the interpretation/necessary subjectivity of information.

> It's not just raw numbers ... but it's not knowledge because it's the experiment and a very contrived environment.... It is interpreted data about the contrived environment.

This interpreted data-as-information is not knowledge because it does not represent a phenomenon of the object under consideration. Data produced as part of a simple experiment, does not form the basis of action in the real metalworking area.

Experimental information (or interpreted data) as the product of simple experiments may not conform to the real world. However, information may form the basis of understanding of the real world. Information, combined with a statement of "I think you should do this" (classified as knowledge) may allow design engineers to create new designs. Flows of information wholly contain "raw data"[*] *and* its explanation, whereas an independent flow of data would be subject to external interpretation.

Knowledge is an understanding of the real world. It may be formed from personal experience, but may not be about "unreal things." An unreal thing is "Information, because it's not a real thing yet. It's not knowledge about a real device. It's still a

---

[*] The use of "raw" here is a specific example of a conceptual trading zone discussed on page 32.

hypothetical."

8.5.3.2 Interview VI

> The participant is an Information Architect. Data are facts produced from instrumentation and sensors. Information is multidimensional data.

The participant is one of the information architects that I interviewed. They explore the implications of computerized data. However, in contrast with other interviewed information workers, they clearly articulate data to be facts exported from measuring instruments. Although there are elements of contextualization here, the view of data as records of measurements indicates a different construction of data when compared to the other information specialists[*]. More research is necessary to explore how interactions with technical devices shape conceptions of data.

There is a strong relationship between information and knowledge:

> An example of information is a time series trend. An example of knowledge is advice (expert correlates information, presents a suggested course of action). Information can be provided to support decision-making. Examples of information: Train timetable, an index of crime levels.

There are suggestions of a relationship between knowledge and data: "Sensors provide data, something you can work with." Then, information plots data versus data. Knowledge contextualizes data (and the plot of data).

Knowledge is a generalized statement about the world. Information is multidimensional data such as data versus time. The act of contextualizing data by forming data sets and comparing it against other parameters like time produces reports as "information." These reports are locally true, and may be explored for insights into trends and patterns (knowledge).

---

[*] Contrast these results with Interview I (p. 280) and Interview IV (p. 289).

Data produces information, which produces knowledge. Mathematical models are information, which complement experiential knowledge. However, models are useless without data as input. Information-as-model requires contextualization (knowledge) by a person to be useful: "People's experience, as knowledge, tells them when to transform data into information into knowledge." Thus, data is treated as if it is atomic, despite being decomposable in other contexts.

Looking at the higher levels of the hierarchy, knowledge must be explicit and encodable (such as into source code), but it has fewer restrictions than information because information is less discrete. Data are components of information. People treat "derived data" as "real" despite knowledge that it has been processed. In the local jargon of the research environment, real seems to mean "authentic and an atomic representation of a measurement."

Data are measured facts. They can be input to mathematical models, contextualized with information, and generalized with knowledge. They are created from sensors and bits. This construction clearly articulates Ackoff's hierarchy of data, information, and knowledge, viewing data as unimportant building blocks in the path towards generalizable statements.

8.5.3.3 Interview VII

This participant is less computer-focused than those of the prior interviews, instead investigating physical phenomena. Although data are also facts for participants, there exists an

> The participant is a scientist. Data are numbers formed from data points. Clearly articulates Ackoff's hierarchy.

entity below data: the "data point" or specific measurement. Sets of data points become data, restricted purely to the numerical results of measurements.

Data are numbers, and only numbers. "Data is a multi-dimensional collection of data points. A data point is a number." As well as "Data plus analysis creates information. Analysis is a function of a mental causal model. Understanding of the model is knowledge." The ultimate goal of data collection is the creation of knowledge as the basis of science.

Information is the basis of action and a container for data: "Information are instructions to perform work of a specific nature, guidance into what it is they're going to test." Information is what allows humans to judge between different outcomes: to order one product over another because of scientific results *contained* within the information. Those results, interpreted and packaged, are data. Although the participant uses the container metaphor here, information is not acting as a passive sign-repository. Rather, the contextualization of information is explicitly differentiated from the pure numbers of data, rather than data being transformed into information.

Data, analyzed into information, may create novel information. Novel information can generate knowledge. Knowledge is a predictive model of the universe, "based on generated meaning through interpreted results (information)." Knowledge as predictive model is far superior in this hierarchy because of its generalizable quality, something that individual decisions to act do not contain. A meta-analysis of action, the results of acting on information received, can create meaning or knowledge:

> Knowledge is a model derived from analysis of information. You analyze data and generate new information. New information can generate new models. Models are knowledge. A shift in [a] model (physical or whatever) is due to new knowledge from discussion.

This construction demonstrates Ackoff's hierarchy [*], with the scope increasing as each level of abstraction is gained. Models, as the ultimate discussed step, represent

---

[*] See page 27.

knowledge, because they represent that which interprets incoming information. At the base of the hierarchy, however, rests data as facts measured into numbers.

8.5.3.4 Interview VIII

The participant's background is strongly different from that of most of the other interviews, as they are not directly affiliated with the company. The interview was taken spontaneously while they were visiting the site on one of the days

> The participant is a researcher from a different company. Data are objective observations with no restriction on representation. Data can be generalized in Knowledge and contained within Information.

I conducted interviews. Data are rendered measurements. Unlike Interview VII, the measurements are not constrained to purely numerical entities, but require the same level of objective rigidity.

In the interview, the participant did not discuss any hierarchy of data, information, or knowledge. While there seem to be strong verbal differences between data and the Os, there are no transformations from data into information or knowledge. Information is understood quite differently from data or knowledge: "Information is a container for data and knowledge."

Knowledge is transmittable experience: "Knowledge is perception of something, what you think you know." Experience manifests as intuitions and requests for specific data are an expression of encoded knowledge. In one example, the participant transmits a request for an image to the database: "I would be sending that knowledge, even if the DB doesn't acknowledge it." Knowledge is transmitted in the act of using knowledge, despite the receiving components possibly not being able to understand that what they are receiving is indeed knowledge.

Specific data retrieval is a function of knowledge. Because knowledge is present

in the act of using knowledge, data requests to computing devices are formed from knowledge and therefore represent the sender's knowledge, rather than the recipient's received data. This represents an interesting meta-component of this conception of data, representing an explicit trading zone formed between entities sharing what *one* of them considers to be data.

Information is a container. Bits are information and form the substance of the container. Beyond the container metaphor is also a question of quality:

> Information is a quality. It mixes with data and knowledge. Something that you relate, it goes from one source to another. Any kind of data or knowledge is some kind of information.

Thus, information is more than a simple container, and may seem to be an intrinsic building block of data-knowledge interaction with the world. Information seems to equate to semiotic ideas, as it can be a representation or transmission agent. In strong contrast to the constructions of the data-as-communication camp, *information* is communication. Information, being formed of signs, requires interpretation for the extraction of the data and knowledge components of use to computers and humans, respectively.

Data are factual representations of measurements. They usually form numbers, but, "Information is text and data is numbers, but even text could be data. Text as data are descriptions of measurements." The fundamental characteristic of data is that it is the factual product of experiment. At the same time, the participant also seems to use data-as-technical, because they describe punching in a few numbers to get a drawing as data as well.

This interview represents a fascinating departure from the normal affordances expressed as it provides a strong definition of data and knowledge and a very weak, subsidiary, and technical view of information. One possible explanation is that, coming from another country's cultural understandings, participant linked the term information to

the computerized bits and had less experience with other uses of information. This interview, alone, suggests a strong need for an international investigation into the different cultural conceptions of data.

8.5.3.5 Interview IX

The participant engaged in an interesting interview: they produced two diagrams exploring both an academic project and an engineering project. In both frames of thought, participant describes data as something inherently numeric, with some basis in fact.

> The participant is an engineer. Data are numbers that are orderable within a list. Source code is also considered to be data as it is comprised of numbers and instructions for those numbers.

Data, information, and knowledge relate in a cyclic hierarchy. Data interpreted via existing knowledge creates information. Information analyzed creates knowledge. There is a container metaphor in this construction of data as, "Information can appear alongside knowledge or as a component of knowledge." As a practical example of this understanding: equipment returns data that is processed into information via an engineer. Analysis is then performed as to "whether things are working." Information is produced by an analysis of data.

Knowledge is fundamentally experiential: "I would define knowledge as something that an individual or entity... something that an entity possesses. That typically has been learned from somewhere else or gained from experience or whatever that enables them a better understanding then they would have otherwise." Knowledge as understanding represents knowledge as model, articulating the nature of the universe by making predictions about future events.

Knowledge acts as the agent that transforms data into information:

My understanding of how things would work is that knowledge is the process by which data is turned into information. So I would say data is things. It is numbers, it is raw information. It is something that, on its own, doesn't mean very much. It's just stuff. Information is an interpretation of that. So it's kind of something that is understandable to someone without knowledge, without even concept of what the raw data is.

Knowledge is understanding: advice about the causes of unexpected events. Knowledge can be communicated and the creation of understanding of the process under investigation is knowledge. Documentation contains knowledge because it presents a history of actions, and can provide methods for analyzing data. The analyzed data, in some ways the "understood data," is collected in things like reports. The report, however, just contains facts about what was done and the results. "[Less meaningful] stuff is the information and the results are data."

Data are numbers: "You can put them in a list. There is a defined quantity. They can be grouped, they can be ordered or whatever. They probably are data. You can put them in a spreadsheet." However, the emphasis seems to be on numbers rather than on scientific results: not on the objectivity but on the coding mechanism. The participant believes that code is data: source code is instructions and numbers to work with.

Information is connected with the communicative act. Facts are transmitted via information, but there does not seem to be any evidence to indicate that facts are information, merely that information is the semiotic encoding [58] process by which they are communicated. Knowledge can be communicated through information. The recipient, in possession of this encoded knowledge, must then analyze the information to retrieve its knowledge component.

In some ways, information described in this construction echoes the information as container construction of Interview VIII. While there is evidence of the participant's

cyclic hierarchy, there are interesting ambiguities relating to the relationship of information, "raw information", and facts. While data are encoded pre-facts expressed as numbers, this interview presents difficulties in classification and represents an edge case worthy of future research.

8.5.3.6 Interview X

This interview also involved someone external to the company. Their conception of data represents something with a strong sense of embodiment; data is physically present in the world, in relations between matter. Measurements and

> The participant is a researcher external to the company. Data is embodied in reality and can be discovered through measurement. Data represents the fundamental relationship of things to one another.

observations can *discover* data, but not create it out of subjective whole cloth.

One of the more interesting concepts of the interview was that of formality. The idea of formality is raising a McLuhan-like "medium is the message" level of meta-analysis. It notes that the context of a communication and its source within an organization contribute significantly to the actions one takes upon it, despite the same communication being passed from different directions: "Information passed from knowledge won't be turned into knowledge by the recipient due to a lack of formal power within the organization. The information is just 'the background why' that doesn't change actions."

Formality, the coding of the "authority" encoded in the message's medium, is something that modulates knowledge rather than a unique category to itself. Formality represents *an actionable instance* of whatever the category is. Written reports, delivered to responsible people in the organization, are an example of formal knowledge. They are incorporated into the organization's stock of knowledge and acted upon. Those same reports delivered to interested parties, informally, do not have the same actionable

qualities.

Formal knowledge is an expression and transference of technical understanding. Components of the flow described by the participant include distilled and certified data sets, the analysis of those sets, and other supporting observational material. The certification of the data sets is to insure that the analysis is internally consistent. Certification is checking the fit of "internal guidelines of quality." To designate something as certified is to express an opinion of its reliability and repeatability. Because of the additional process of certification, formal x can be acted upon. The participant believes that normal flows, lacking confirmation, are not processed in the same way.

There is clear evidence for a reverse hierarchy in the interview. The Tuomi hierarchy suggests that we use knowledge to build information that allows us to extract data from the world; this data then allows us to update our models of the world. In this way, as information frames our requests for data (subjective or objective as they may be) it is possible to check the integrity of the data and contextualize it by inspecting the originating information. Experiments can be designed to extract information from sensors and computing devices operating on machinery. When sensors are rare, experiments are designed to maximize "reliable information from an experiment."

Knowledge is expertise. Expertise, however, is expressed as an understanding of the world that *shapes* action. The participant observes that "corridor conversations" discussed in an informal setting can contain information and knowledge. Information, in this case, "is the things that I can reduce to something technical." Knowledge "is their understanding of the process, how and why they turn levers and their perception of the basis of that." Thus, as expertise accumulates from the past shared precedents of actions, events, and training, it acts as a guide to action. It informs information-gathering methodologies, which then inform data-collection strategies. The use of "reduce" in the

quote above is better understood as "abstraction" rather than implying that information is "under" knowledge.

Information has a temporal element or has a different level of reliability. Whereas more sensors, both cheaper and more connected, produce more data, the output of multiple sensors over time produces information. "In the context of the whole operating campaign, you don't know the reliability of the information, you don't see the built in trends... but you've got the data points. Three quarters of the data in the database, you don't know how it was acquired." Although information may appear to be used synonymously with data, there is a suggestion that a thing can be classified as data or information based on some kind of underlying or meta-characteristic.

For example, a physical sample can embody both data and information, depending to whom it is sent. This suggests that observations and numbers are latent in the world. In this way, observation *transforms* latent information or data in the atomic or molecular structure of the thing being experimented upon into more easily recognizable numbers with defined error limits. This transformation is not an external imposition of measurement, but a recognition of something that is already there.

Information is an observation of the world. Observations of physical nature are information assigned from parameters informed by knowledge as previous experience. The participant strongly believes in the objectivity of information. They require that different types of measurements of the same object correspond. In this instance, a sample represents embodied information. There is no inherent knowledge of what it is in the object, but it is information. If the information is analyzed, it then can become data. In this vein, environmental observations surrounding the sample are information, creating a "picture of the problem." Binary signals from switches can be information or data, depending on whether they represent state status (information) or measurements of

quantities and masses.

The apex of the hierarchy is data. Data are contextualized hard numbers. Data can be generated from "something you can reduce to a number. It has observations, but it's reducible to a number." Optimally, samples will be accompanied by recorded site observations, situating the sample in some sort of context.

Unlike the other interviews, this one holds no indication that contextualization serves to transform the data into information. Instead, contextualization creates data that are more valuable. All things contain within them latent information that, through the correct experiments, may be observable. Only through analysis can the observed and recorded results be promoted to data. One method of analysis is a chemical analysis because the results of chemical analysis are deemed "real" and "hard data" numbers according to standards.

At the same time, data has different natures according to the roles that need the data. "The geologist says I want the sample. The chemist says I want the analysis." The participant notes that for other roles, data's substance can be process data or the discovery of relations within the set of process data.

Data exist at various levels of accuracy. "Industrial quality" data has accuracy considerations different from those of academic or theoretical quality. Furthermore, dud data-data that is erroneous-is possible and is a likely result of cost-saving measures.

8.5.4 Reflection Conclusion

My findings fundamentally agree with and extend both Chaim Zins' work and the definitions of the *Oxford English Dictionary*. While more research is necessary, there is little doubt in my mind that I have answered the two questions of interest:

- I have discovered constructions of data that profoundly differ from each other and that agree with 0, related research.

- Some evidence points to the existence of trading zones between different constructions of data. The use of the terms "raw data" and "derived data" in the interviews strongly suggests this. However, more research is necessary before there is any useful evidence for how trading zones form with constructions of data and what impact evaluative accents have on data and misunderstandings.

- There is evidence for both Ackoff's and Tuomi's hierarchies of data, information, and knowledge. Ackoff's work has penetrated modern thinking, especially management thinking, quite deeply. At the same time, the more "modern" relativistic philosophies of science seem to correspond quite well with Tuomi's ontology. Far more evidence is needed to create suggested "trading zones" that can bridge these and other discovered conceptions of data.

- Some indications hint that the developer's understanding of data is a factor in database design, though the current evidence cannot create useful rules for database changes based on the recipient's current conception of data. More implementation research is needed here, as well.

- Both the interviews and the surveys discovered different realities of data, but not without failures. More research is necessary to explore the utility of my methodology and to refine it for different purposes.

8.6 Surveys

To demonstrate that the results of the research were not a fluke of my methodology, I then decided to run a set of surveys to explore another perspective on the reality of data. This section documents the results of my second survey attempt and is a check on the conclusions drawn from my interviews. Many different participants participated in this survey, being drawn from academics, IST professionals, other workers at BlueScope, and intelligence workers from the United States civilian and military intelligence establishments.

Survey analysis proceeds in a roughly straightforward fashion. I discuss interesting aspects of each incoming survey, and try to derive their conception of data by looking

at how they classify data, information, and knowledge. Their classifications and the relationships they articulate between them may give some clue about each participant's conceptions of data. Less time will be spent on each individual survey due to their less comprehensive nature. The terms "participant" and the singular "they" will continue to be used to indicate the survey-taker.

These surveys were the first deployment of my survey methodology that resulted in any kind of success. As such, the instructions were not close to perfect, with a few survey-takers misunderstanding the instructions. I state this with certainty as one participant (name withheld, but permission given) had the following to say in an e-mail:

> Also, I had a hard time with the questions. About midway through the survey, I realized that some of the questions were looking for a semantic understanding of the meaning of the highlighted words. I began the survey looking at it through an analyst's eyes, with the understanding that data coming in only becomes information or knowledge if it reveals something about the intended target – usually a person or group of people (network). As I went along, I realized that you were differentiating between data, information, and knowledge as semantic definitions that all described different types of incoming "media" or "communications". I got the impression that you viewed data as something measurable, information as something that feeds understanding, and knowledge as something that is either factual or knowable. Since these are pretty disparate understandings of the terminology, it was difficult to reconcile my answers between the two methods of differentiating the data types.
>
> I offer this as perhaps an anecdotal story about the problems with this kind of question. Like so many issues surrounding understanding within a community (e.g., religious doctrine), the real question here is perhaps not so much a matter of different perceptions within a subculture, but whether or not the subculture has taken the time to clearly define their terminology to enable effective communication. For more "tangible" elements of conversation, like computer design, engineering, or medicine, the subcultures tend to build a sophisticated lexicon of industry jargon very quickly, because the objects described (whether processes, methods, tools, concepts, or physical things) are easy to understand. With subcultures that focus on more ambiguous topics (freedom, truth, spirituality, goodness, guilt, culpability), the clarity of their jargon becomes diluted by variation in understanding of the meanings of the subculture's terminology. You (purposely?) do not define what information, data, and knowledge mean in your survey. Therefore, these terms will carry a wide variety of denotations and con-

> notations with them. The implication is that effective communication within the intelligence subculture cannot begin until the subculture's definition of these terms is made explicit.

Therefore, the products of this analysis should be taken with a grain of salt, and used to suggest that far more work can be done in this area, as opposed to presenting any kind of definitive or reproducible look at the reality of data.

As stated in the methodology[*], the surveys allowed participants to self-identify by role and then classify a series of scenarios[†] as to being data, information, knowledge, or other. The self-identification of role will be communicated in summarized form in each survey below. The summary will be generalized to preserve confidentiality. From each survey, I will abstract the most interesting explanations given of categorization. Included in each summary will also be a description of the general proportions of the answers. Each participant not only has their own conception of data, but their own rationale for identifying things as Data, Information, Knowledge, or other. The analysis will also briefly touch on the nature of each participant's categorizations of the O, and how they may influence their construction of data.

Each introductory paragraph will contain a count, as noted, of X data, Y information, and Z knowledge. Although it is possible that interesting statistical observations may be made from the relative proportions of classification, the main items of interest in my research were the participants' explanations of their decisions, rather than the quantization of the decisions themselves. In addition, the number of surveys is not sufficiently high for any statistically useful generalizations. As ever, this research is just laying the foundation for more intensive research, and I feel that this survey structure

---

[*] See page 71.
[†] See Appendix A.

could form the basis for an excellent long-term study of many people's conceptions of data and how they change over time.

8.6.1 Survey I

The participant is working in the database labs of a university. They classified most scenarios as information and knowledge, with the counts of the various classifications being: 3 Data, 11 Information, 10 Knowledge, and 2 other. Participant was careful to classify other scenarios as a "lack of data" and "ambiguous" that increases the reliability of prior categorizations as it demonstrates that participant is not afraid of novel categorization.

Data is a symbol without meaning. Some quotes:

- "This doesn't necessarily contain information and can be simply random."

- "A design may or may not convey information."

- "[secret] Code, in this case, is data since it does not have information. However, if it is decrypted, it becomes information."

The participant presents a novel interpretation of data. One explanation of the unusual sense of symbol without meaning is that the participant's role is sharply different when compared to the other interviews. Furthermore, when considered in light of databases, a purely technical understanding of data could support the symbol-without-meaning interpretation if the participant does not assign computers the ability to manipulate meaning.

Information is classified as "anything with meaning":

- "Information, since data doesn't necessarily have a meaning and knowledge is factual/practical information. A letter, therefore, can be knowledge, but is normally just information."

- "'Locations of parts' would [be] basic information. The data has meaning, however it does not detail practical or factual steps of information."

- "The data is known to be a planning program, so it has meaning."

The notation of meaning in information has links with the information-as-communication theme found in some interviews above, but suggests a purely semiotic boundary: symbols-with-meaning versus symbols-without-meaning, without any kind of communicative subtext.

Knowledge is classified as practical information. Some quotes:

- "This is something that contains factual/practical information; i.e., knowledge."

- "Toughie; the story pertains to steps detailing a story, therefore it has chronological sequence (i.e. - steps, procedure), so it is knowledge."

- "Step-based, procedural information is knowledge."

This position is an interesting one to take, inasmuch as knowledge is a subset of information that relates to action or time. Knowledge is anything with practical meaning, or in other words, anything that is actionable.

The participant belongs in the data-as-communications camp, assigning computers the role of data processors while granting a privileged human-centric view of information and knowledge as symbols with meaning (and utility). Although the relationships between data, information, and knowledge are unusual, they do not present any unexpected new conceptions of data.

8.6.2 Survey II

The participant works for a defense department as an information processor. They present a strong bias towards classifying scenarios as data, with 11 as Data, 6 as Information, and 9 as Knowledge. They did not take advantage of the other category. One hypothesis is that as an information processor they have a very strictly operationalized

definition of information. Still, the lack of any categorization of other suggests that Data

or Knowledge were used as catchalls.

Data is an unanalyzed sign, set of signs, or communication. Some quotes:

- "No analysis or recommendations for action are involved, leaving this as simply data."

- "The code has not been sorted, so it is still data. The fact that the data is encrypted offers some insight for analysis, so that may be used to create information from this situation."

- "Lacking detail about what the letter contains, I would default to the lowest category of analysis, and assume it to be data for the time being. I can be sure that it is data, but cannot determine from the description given if it is information."

While "data as a fundamental building block of information" is unambiguous, the

attached semiotic definition results in an uncertain hypothesis. The interesting phrase,

"Lowest category of analysis," suggests a strict and articulated set of relationships among

data, information, and knowledge. The very strictness of the hierarchy, of course, means

that accessing it is quite difficult: much of the participant's knowledge about data is tacit

and assumed, making it difficult to extract from their explanations of reasoning.

Information is analyzed data:

- "The data of how the user interface worked is changed to information by the analysis of the users, categorizing things into 'good,' 'bad,' 'useful,' etc. However, this does not include recommendations for action, and so doesn't quite go as far as knowledge."

- "Quotes are chosen to reflect a given point, and so are offered with inherent analysis, by the questioner, the respondent, or both. "

- "Poetry, while perhaps not designed as traditional analysis, does offer insight, and information-at least about the poet. The poet has analyzed words, scansion, rhythm, etc., changing raw data (perhaps Dadaist poetry?) into traditional poetry."

Analysis of data creates information. However, if we explore the meaning beyond

analyzed data, defining the term becomes quite tricky. It is easy to exclude knowledge

as actionable information, but the very specificity of the term "analyzed data" precludes

most understanding. In some ways, the meaning of information could be as broad as information as statements about nouns in the world.

Knowledge is information applied to decision making. In many ways, this classification echoes the knowledge of Survey I. Some choice quotes:

- "Bob has used information (prior experience with traffic noise) to make a decision on how to behave, transforming his experience into knowledge-knowledge that he uses to judge what sounds to ignore."

- "Charlotte has not only analyzed information on behavior (making it information), but has taken that a step farther, into knowledge, based on its applicability."

- "Dave is providing not only analyzed data (information), but is doing so to empower action (a change in answers for the next time Eve encounters the situation, on a test or in her outside life), changing it to knowledge."

As knowledge is applied information (similar to Survey I), knowledge must be actionable information. This relationship is a reasonable derivation of Ackoff's hierarchy, put to different ends. While Ackoff's hierarchy generally has knowledge-as-model, another understanding of model could be "information with practical consequences." This view necessitates filtered information, as information contains elements both with and without practical consequences. By assigning practicality to knowledge, one can relegate the elements of pure-theory without application to information until they are filtered and contextualized by use.

8.6.3 Survey III

The participant is a manager of a consumer electronics repair workshop. They gave a normal sorting of responses, with 7 Data, 11 Information, 6 Knowledge and 2 other. The Os were identified as structure:

- "If it is a program only with no facts or data, perhaps this is just something that transforms data into information."

- "I think the empty document is something that assists in organizing data into information, but in itself is just a structure."

The wording on these strongly suggests a standard hierarchy. Furthermore, the relationships allow transformation: data becomes information becomes knowledge.

The participant's definition of data is slightly more problematic. Although data can nominally be defined as a statement without structure, there is little evidence to indicate what data is a statement *of*:

- "I'd say this is data, as the facts have no relevance on their own. I'm not entirely clear on whether it is data or information, though."

- "This one is interesting, also. I'd see each individual mp3 file as information, but the group of them together as data, as it feels like more of a random sample than a structured grouping."

- "The randomness and lack of interpretation leads me to think this is data."

- "The quotes themselves are data until they are processed, at which time that would become information."

The first quote suggests data are facts, but does not indicate the participant's understanding of fact. The second quote belies the traditional hierarchy by placing a set of data above information, though the trend of randomness or "lack of structure equals data" is still present. Thus, in my analysis, I have to default to the description "statement" without further elaboration due to the participant's answers.

Information is analyzed data:

- "Now having the structure to put the information into, I think it is now information."

- "Data with structure."

- "This is a statement with clear intent and purpose, so I'd say the demand is information that the mugger is conveying."

In keeping with the traditional hierarchy, information is data with structure or data post-analysis. The last quote also implies some hints to information-as-communication as well, though that meaning is overwhelmed by the data-with-analysis usage.

Knowledge is understanding:

* "I'd say this is knowledge, as she has taken the information and used it to form an understanding of the situation."

* "Someone's impressions of something would be the conclusions they've drawn from the information, so I'd say knowledge."

* "That selection is a conclusion based on understanding of what he is doing."

From these quotes, the participant suggests that analysis of information becomes knowledge. Knowledge as understanding or conclusions does not suggest the use to which that knowledge is put, however.

This survey presented difficulties due to the fuzziness of the participant's explanations. While there is some evidence of a standard hierarchy, the relationships between the terms are poorly articulated. Here the survey did not provide enough guidance to extract from the participant their ideas on how data, information, and knowledge relate to one another. At the very least, this survey presents evidence of how deeply Ackoff's theories have penetrated the "white knowledge"* of society.

8.6.4 Survey IV

Participant is a social entrepreneur. Their self-description of their role emphasized management and planning skills with a correspondingly high information consumption. Participant overwhelmingly classified things as information, with 5 Data, 13 Information,

---

* White knowledge is the everyday knowledge that "everyone knows" but that doesn't need sourcing or attribution due to its sheer ubiquity and acceptance. I believe Sir Terry Pratchett extended the term "White Noise" for white-knowledge, capturing the same sense of randomness and lack of coherence. [59]

5 Knowledge, and 3 other. The use of other suggests a willingness to differentiate, although it seems that from descriptions of what they wrote, some of the categorizations are in error: a few times participant categorized something as information but explained it as data. Participant differentiates emotion from the traditional hierarchy, as they categorize both mugging and poetry as flows of emotion. There is insufficient evidence to articulate how emotion relates to data. Despite these problems there seems to be some evidence for the traditional hierarchy.

Data is an observation:

- "Random noise can be data."

- "Data only-no context or filters to apply."

- "Encrypted information is structured data without a context for Alice to extract relevant knowledge from."

Data is filtered and structured by knowledge, thereby producing information. The participant understands data as subjective observations, as indicated by the keyword "filter." These observations are filtered and contextualized into information, in keeping with Tuomi's cyclic hierarchy. This set of responses matches interviews III and IV quite well, suggesting some commonality of background. Most importantly, it suggests that the realities of data explored in those interviews were not flukes, as the SDFN also produced these results. This minimum demonstration of reproducibility is heartening as it validates both methodologies.

Information is comprised of filtered, structured, and relevant communications:

- "Instructions are information, organized data that conveys...."

- "A short story is data filtered through the author's perspective."

- "An empty word document is information because it has structure and that information may be the presence of structure only."

These quotes suggest that structure, while necessary, is a component of the analysis performed upon data to create information.

Knowledge has explanatory force. It entails conclusions and explanations of the world. The participant is clearly using knowledge in the traditional understanding-the-world context.

- "The key is knowledge that once learned creates a new set of filters for Alice to perceive the data through."

- "Dave is passing on the knowledge of why one procedure is right or wrong."

- "Design is choice, choice involves knowledge."

However, knowledge is sharable and a way of creating context. The context-creation component of knowledge in many ways suggests Tuomi's cyclic hierarchy rather than Ackoff's hierarchy in the way that knowledge plus data create information: "the key is knowledge that once learned creates a new set of filters for Alice to perceive the data through." Knowledge is also a way of generating choice.

8.6.5 Survey V

This survey, produced by a research scientist, was incomplete and very terse. The participant categorized six entities as Data, six as Information, one as Knowledge, and one as other. The instance of other is "probably both data and information," which tells us little about whether the participant was willing to explore the non-data categories.

Data are observations of the world:

- "At this stage the numbers and times are data. Study and interpretation of these will lead to information or knowledge about the thing that is unknown."

- "The noise is a meaningless stream not constructed into anything more significant."

- "Just listening to a live music stream is receiving data. Gaining pleasure from the story being told by the data stream is a construct inside one's head and is producing information or indeed knowledge of a larger portion of music, e.g. the symphony."

There is no evidence for any required activity with respect to the observations. The identification of noise as data suggests that there is no intentionality associated with the data. Unlike some of the interviews, wherein data is the product of sensors and experiments, noise as data suggests that data is inherent in the world, and that the observer's knowledge is what allows for filtering and contextualization into information.

Information is contextualized data:

- "A collation of data within a context, not just random."

- "At one level the planning program might be considered to be knowledge in itself. However, it is information at the level before the unknown thing is revealed."

In this instance, the differentiation between data and information is context. The participant applies the term information to the planning program without exploring the term. They give no explanation for how information in the planning program is both unknown and contextualized.

Knowledge is understanding:

- "This requires an understanding of the machine as a whole, i.e. a map, so is beyond data and information."

The use of the term "beyond" suggests the traditional hierarchy, but tells us little beyond the simple use of the term understanding. The single categorization of knowledge and the paucity of responses in this survey make intuiting their conceptions of data problematic.

8.6.6 Survey VI

The participant is a service desk employee for telecoms. They place an unusual emphasis on the term information, with 5 Data, 16 Information, 3 Knowledge, and 2 Os. They note

that an experience is not knowledge, nor can art be classified as data, information, or knowledge.

Data are structured records:

- "The quizzes are on relational algebra, which is more data-driven than say a quiz about philosophy."

- "Data because it is a record of data. A record of few parameters. Day + temperature outside her apartment."

- "Lists of things are perfect data, to me."

Data seems to be a subset of information: structured records containing statements about the world. However, the classification of quizzes as data is strange, although the participant may be responding more to the topic of quizzes than to the entities themselves. Thus, information with structure is data. There are no suggestions as how to utilize either with knowledge.

Information is a statement about the world:

- "What did Alice receive? A letter. The fact that it was a letter is a piece of information."

- "The files are information."

- "Bob ignores the noise. This is a piece of information."

- "What did the mugger do? Demand his wallet and watch. I don't see this as data, it is not a few parameters like a temperature chart."

- "I would say this is information. Yes, a lot of knowledge is behind the information, but such is the cases of a lot of information-there is knowledge behind it."

In the last two quotes, the participant defines the boundaries of information using the example of the temperature chart for data and "background knowledge" for knowledge. In many ways, all the examples are simple facts without relationships and the facts are statements about the world.

Unfortunately, the few answers the participant gave for knowledge do not provide a useful definition:

- "The parts are at different locations. Where they are is a piece of knowledge."

- "A lecture is knowledge-based. It is not merely a list of items. It is not a list of information."

While there is a suggestion that knowledge consists of factual or true statements about the world, there is no way to reach that generalization from these responses. Furthermore, as the participant uses the term "fact" when describing information, knowledge has to be something beyond simple fact.

8.6.7 Survey VII

The participant self-identifies as a corporate strategy manager. They gave balanced categorizations to data, information, and knowledge, with an extraordinary 9 other versus 5 Data, 6 Information, and 6 Knowledge. They used the other category to delineate art, wisdom, and lists. Of all of the surveys so far, I am the happiest with this one because the high incidence of other categories suggests strict definitions for the three categories of interest.

Wisdom structures knowledge, creating behaviors:
- "A learned behavior from a past experience."

Some of the scenarios were not even data, one categorized as a list or sequential phrases:
- "Not even data because there is no metadata to give context to the list."

- "I found this a tough one. At first I thought information however because the context in unknown it doesn't fit with my view of what information is. I also don't think this is data because it is not observational and again lack context. The best thing I could think of was sequential phrases. You know each phrase is in a specific order but you do not know the intent."

Data are observations without interpretation:

- "I think it's either information or data because the code is just a transformation of the content to maintain security. Once transformed the content may be contextual in which case it's information or observations which suggests data."

- "It's just a list with no interpretation attached."

The first quote indicates observations are data; the second indicates a lack of interpretation. Interpretation-as-context suggests a normal hierarchy.

The participant indicates an understanding of information as context:

- "Without any descriptors as to the type of letter the word letter has an ambivalent connotation about it. There is a contextual element to a letter that renders the content more than data but not necessarily knowledge which to me suggests value and worth retaining."

- "I think it's either information or data because the code is just a transformation of the content to maintain security. Once transformed the content may be contextual in which case it's information or observations which suggests data."

The indication of context *as* information is slightly unusual because the usual formulation is data *with* context. However, information as the contextualizing element is presented well enough here to leave little ambiguity as to how they think.

Actionable responses are knowledge:

- "Because the location is critical for the part to perform its role in the larger machine. The location and the part are linked and relative to another part and location. If you just had the concept of the machine and location, you would only have data.... If you understood how each part worked you would have information."

- "The secret key provides an actionable response to accessing the coded content therefore it is knowledge."

The precision of the answers is fantastic, especially with the examples that the participant gives. They clearly indicate that actionable responses are knowledge and give examples of how the relationship between categories progresses to said action.

8.6.8 Survey VIII

The participant is a counter-terrorism analyst who reads many messages daily, exploring trends for counter terror issues. They identify 6 Data, 10 Information, 6 Knowledge, and 4 other. The participant used the other category when unsure, rather than inventing new categories. They believe that live music and other entertainment do not provide data in themselves, and that entertainment is a minor and irrelevant category: "one cannot have true knowledge about something when the results are deliberately trivial." They articulate Ackoff's hierarchy, but seem to employ a cyclic hierarchy in practice.

Data are objective records of activity:

- "Dave has provided himself a reminder, which is a type of data; or he has uploaded a format with appropriate font, pitch, etc., that he needs to use (perhaps he works for a bureaucracy that is very particular about these sorts of things). Because it is empty it cannot be more than data."

- "Impressions of an interface are simply impressions; the aggregate will be information. It is not "data" because impressions are not data, data does not allow of emotive responses. But the impressions are not knowledge or wisdom, they are merely impressions."

- "Receiving an email is merely a fact; the email could be anything from a joke to a sales proposal to a social invitation to collaboration on a project. the fact of receiving the materiel is unable to be evaluated without content and context."

Data are factual records that cannot contain subjective elements such as impressions or thoughts. Entertainment, as stated before, does not produce data, but the record of someone consuming entertainment is data, because it is an objective fact about their person. Furthermore, all facts must be about activity of some sort. The participant is not interested in scientific facts per se, but in records of actions that can be contextualized into information and analyzed into knowledge about the target.

Information is data with context and assessment:

- "As stated before, the context provides so much about m.o., etc., that this is more than mere data, but because there are critical parts missing or not understood, this is only information."

- "This is raw data regarding what was watched and when. There is no assessment of the data nor is there a greater purpose that is being described. There is no wisdom, knowledge, or information from this, only data."

- "There is no historical context as to why this methodology was chosen, so this is not 'wisdom,' and because it is a set of instructions, this is more than data."

The assessment of the data may be performed by knowledge, but is never explicitly stated as such. If knowledge contextualizes data, then participant is using a cyclic hierarchy, wherein knowledge contextualizes data, which in turn produces information, which can be analyzed to knowledge.

Knowledge allows someone to make predictions about the world:
- "Prediction moves beyond data and information and – presuming the analysis is correct – gets to real knowledge."

From an intelligence standpoint, this makes sense: predictions about the world are one of the ultimate intelligence products. Thus, the outcome of knowledge is a prediction. The nature of a prediction differs from the usual knowledge models of data-as-facts constructions. Although the participant indeed considers knowledge to be models, the nature of a prediction is such that it is an exported and actionable model upon which 0 people can base their activity or lack of activity. Of interest here is the traditional view that enemy actions can communicate volumes about their intelligence sources and models of worlds, as unlikely actions may indicate the presence of a leak in one's organization. The ability to make predictions that are given to other people, therefore, is tempered with far more severe complications than physical models of the world.

8.6.9 Survey IX

The participant is a housekeeper. In this job, they have a low exposure to workplace-enforced understandings of data. The participant has a very unusual categorization: 7 Data, 5 Information, 4 Knowledge, and 10 other. The high rate of other represents a lack of broad categories. In a sense, this represents a lack of formalization from learned patterns.

The participant makes a brief mention of data as technical-communication. In other words, they consider data-as-bits a separate use of the term data:

- "Computer data-a different use of the same word.... Information, in this case, music, stored on computer hardware."

In this, they also conflate the term with information, which is a common usage. However, this separate meaning should not be confused with a conception of data as data-as-communications. While using data as a term for computer records is certainly possible among all of the three realities of data, those who use it as their primary construction elaborate upon it more; the idea of it being a side definition never really occurs in practice, there *usually* is a seamless transition in usage depending on context. However, this area is another in which far more research will be needed before it is possible to make any definitive statements, especially the exploration of the use of the two terms, data and information, to mean the same thing.

Data is a factual scientific observation:

- "My father used to talk to me about the difference between facts and opinions. Short quotes have the potential to contain either, but will more likely be composed entirely of opinions than of facts. Why would anyone collect quotes from students reciting facts? Chances are, the information would already be known, and be of little interest to anyone watching the campus TV station."

- "This is data of a specific kind collected over a week by observation-albeit by the subject of the observation himself."

- "The answers given on the quizzes are collected data about the knowledge and correctness of the people who did them."

These factual observations are required to have a strong correlation with reality. With the prior term of "computer data" specifically excepted in their comments, they consider data to be inherently factual and representative of the world.

Information is a statement carrying meaning about the world:

- "A completely empty document contains no information-it is an absence of information except for the fact that the document itself exists. Now, this may in some circumstances be useful information, but the document itself is devoid of meaning, since it's completely empty."

- "Poetry is an expression of emotion rather than knowledge, confirmed information or collected data most of the time. While one's emotional response to winter wind may be important, and interesting information, poetry about it is in a different class than other forms of information about wind. Its purpose is emotional rather than informative."

Information appears to be a fundamental component of data and knowledge, serving as a general communicative vehicle. It also has some elements of meta-data, as the presence or absence of the document is termed "Information." More research is required to see how significantly different professions interact with data, information, and knowledge.

Knowledge is formalized, reliable instructions for activity:

- "Reasons behind an answer could be knowledge or information - it doesn't really seem certain enough for me to call knowledge."

- "Lectures about database design... or other academic lectures are to confer built-up knowledge from one person to 0s. It's also information."

Knowledge is transferable, certain, and constructed. The phrase "built-up knowledge" suggests a subjective, constructivist view of knowledge. Knowledge is viewed as a subset of information because something can be both knowledge and information at the same time.

8.6.10 Survey X

The participant is a senior software architect and acts as a programmer and database modeler. They classified 12 Data, 10 Information, 4 Knowledge, and no other. They make a strong differentiation between data and information, but isolate out knowledge as a privileged case.

Data are electronically stored observations:

- "A letter is akin to an email which is electronic data. A letter in this case ins hand-written or typewritten data."

- "This would be digital data. Data that is store on a given medium, in this case, a flash drive."

- "Live music is data. It is sound waves traveling from the instruments to the user's ear."

- "This contains specific data points, meaning the temperature outside."

While there are strong elements of data as bits, there are also elements of data as observations, which suggests that the participant uses both ideas simultaneously inasmuch as data are stored observations. It seems that collecting data is a database-centric act, i.e., data is something which is stored in a database. There are no semiotic overtones or indications of stored meaning.

Information is a semiotically significant transfer:

- "A short story is information. Generally, I would think of this as a fun trivial kind of information that is meant to be temporary and enjoyed at that given moment."

- "This is information, signaling the mugger's intent."

- "A program is a set of code that will run on a computer. It provides information."

There is a component of intentionality to the transfer, but it is any kind of communicatively significant act.

The participant identifies knowledge as reasons behind interaction with the world:

- "The instructions are knowledge. They are specific pieces of knowledge that Bob is giving to Alice."

- "This is a knowledge transfer from the professor to the students."

- "This is an educated decision on what to pick and why."

Knowledge as the basis behind the ability to make decisions can then be extended as the ability to make predictions. Knowledge as decision-making is exploring possible future outcomes and making choices that enable a desired visualized outcome, based on the constraints of the search.

8.6.11 Survey XI

The participant self-described as someone who makes an action plan for a business problem. They identify 10 Data, 9 Information, 5 Knowledge, and 2 other. 0s are categorized as "don't know." The participant makes a strong differentiation between information and knowledge.

Data are objective, precise facts:

- "Definitely data. This time chart can be analyzed and information derived from it. A time chart requires no interpretation."

- "Definitely data. A list. An objective measure such as temperature. You can use it to make information."

- "Not sure why but I feel this is data. Guess it being files with objective stuff in it (like a digital 1/0 recording) it is data. But I could be convinced in some part it was information depending on the content of the files."

- "Again, not sure but I feel this is data. Hard to put this question in a work context and "live music" is sort of an emotive term. Music requires no interpretation (maybe lyrics do....) and is objective (i.e.. you can follow music notes on a sheet)."

Data are objective and listable components of information. All data are factual and obvious to all observers without need for interpretation.

Information is the communication of unambiguous elements such as procedures:

- "Imagining a situation where I receive instructions on how to take a sample for the task force. Being a procedure, I will call it information. It will be detailed but designed for me to follow, not to interpret. When it doesn't need interpretation, then it is information."

Although the participant differentiates between information and knowledge, there is the strong suggestion that they are similar. The participant articulates an objectivity-subjectivity spectrum: data being the most objective to knowledge being the most internalized and subjective.

Knowledge is a framework of learning designed to situate problems:

- "Imagining a scenario where I receive a short technical report to read on the task force. ... This is knowledge. Detailed, contextual and requires interpretation. I can learn something from a report and use it in my own framework to view similar problems."

The act of interpretation indicates that something is knowledge, whereas information is far less subject to interpretation, though there are some suggestions that information is used to build context.

8.6.12 Survey XII

The participant is a research engineer who assists with troubleshooting and does basic research. They provided a standard distribution with 7 Data, 13 Information, 3 Knowledge, and 3 O, with the Os noting categories of music and filtering. Unusually, the participant noted that experimental design was a process.

Data are specific observations of phenomena as well as stored bits on a computer:

- "Specific observations of events - data. No abstraction or generalization."

- "Data - the individual bits that refer to the specific makeup of the songs are data - Its file structure and formatting is information based on the mp3 standard - which is information."

- "The computer sees it as stored data - it is also David's preparation so far saved as one Datum."

There is no need to filter scientific data because they cannot occur without intentional observation and measurements. The duality of data as observations and data is very hard to reconcile. While it seems that data-as-bits is a side definition, there is no evidence to blithely dismiss it. Participant also does not indicate that data are objective, merely that they are specific, intentional observations. This answer illustrates one of the difficulties of the survey: the inability to follow up on interesting leads that are ambiguous.

Information is a set of analyzed data that can create predictions:

- "Information made up of pieces of data" and that it is in some way a description of the world: "The demand is information passed to Bob describing the state of the muggers mind and his potential, this info is based on the muggers generalized knowledge that Bob probably has a wallet - the specific demand is too abstract to be data, but too specific to be knowledge. "

- "Contains some data i.e. measured temperatures etc., but usually has information as well - i.e. predictions"

- "Can't be knowledge because it is not general or transferable, assuming it is primarily fiction-will not contain data. You could possibly stretch the definition a little and say it contains data about a fictional world. But the story does give some information about the fictional content and contained in that Bob's underlying assumptions and experience, etc."

Information exists between the abstract generalizations of knowledge and the factual observations of data. This set-based conception resonates with some of the ideas of data sets from the interviews and falls into the same philosophical framework: collections of facts can be analyzed for patterns.

Knowledge produces highly generalizable concepts about the world:

- "Knowledge because it involves abstract, generalized concepts that presumably have been produced based on previous experience and analysis (data and info)."

Knowledge is predicated on both information and data, creating a gradient of specificity. However, as the other descriptions of knowledge were poorly explained, it is difficult to fully generalize knowledge's relationships with data. However, working from the data-as-facts construction, knowledge must be patterns that provide *correct* predictions of the world.

8.6.13 Survey XIII

The participant is a researcher modeling industrial processes from an engineering perspective. They classified 4 Data, 13 Information, 4 Knowledge and 5 Os. Unfortunately, the majority of other responses were not well discussed, but they indicated a special kind of data, a process, and curiously a null set:

- "Null set. If file contains the correct number of pages and formatted, then 'data'".

The participant indicates that meta-data is a form of data, and technical data is a special case.

Data are numbers without context:

- "Only says what he has watched. Does not specify contents of programs. More information however than just a series of time durations (data)."

- "Time series (no attached informations, e.g., weather conditions, etc.)."

- Sound is "special data."

All of these data are numerical and contextless. Sound is intrinsically numeric and thus is considered by the participant to be data. They show no evidence, however, for data

being anything other than numbers. The use of music/sound rules out measurements as data and there are no suggestions for objectivity.

Information may be contextualized data:

- "I would consider instructions to be more than 'blind measurements' but less than a discussion on how the machine works."

Information, as catchall, has a non-standard relationship with knowledge and data. Although there are hints to a hierarchy, there is no discussion of what actually comprises information. From inference, information is contextualized data.

Knowledge is explanation, chief of a hierarchy of primacy:

- "Depends on quotes: 'Day is Hot,' then information. Explanation to the meaning of life: knowledge."

- "As it stands, then information, if giving reason for her to take her stuff back then knowledge."

Knowledge provides reason and explanation. Only new knowledge is classified as knowledge, and information seems to be used as a catchall.

8.6.14 Survey XIV

The participant is a developer of an in-house optimization application. The most interesting aspect is the very low categorization of data: 3 Data, 14 Information, 9 Knowledge, and 0 other suggest a very interesting conception of data.

Data are facts without context or intrinsic meaning:

- "A letter usually contains a commentary alongside any facts, giving the reader the possibility to interpret those facts."

- "If the subject is unknown, the content are data. However, the scenario may be useful to somebody in determining what the data concern."

The contextless aspect is predicted by the information phrasing, but from the second quote, Data is meaningless without said context. Data, without the ability to interpret them or without a use, cannot transmit meaning. The participant may have a constructivist view of knowledge, which suggests that only data contextualized as information have utility. Data as proto-information, therefore, lack utility.

Information is data with context:

- "There is a sufficient context to determine that it is a 'planning' program (whatever that is)"

- "There is a clear context surrounding the basic demand."

- "The music itself is just data (unless it contains textual messages, which could be considered information)."

Context provides meaning to data that support judgments. The process of contextualization is what allows data to become information.

Knowledge is a judgment:

- "Bob has had to make a judgment on whether this noise is of any interest."

In some ways, the term "judgment" is disjointed from the other philosophies of knowledge inasmuch as it combines imagining and predicting the future with then assessing the best outcome. Despite the novel term, however, the understanding is fairly consistent.

8.6.15 Survey XV

The participant identifies as a SIGINT [signals intelligence] analyst and a subject-matter expert in that field. They identify 6 Data, 1 Information, 6 Knowledge and 2 other. The Os identified are: having an experience and art.

Participant draws a strong distinction between practical matters and art:

- "To me, art is something different than data, information or knowledge. Art can't be easily classified as data or information, for instance. "

- "The music is live, and not electronically stored as information. As such, the music is 'O' in this case. It isn't data, it isn't information, and it surely isn't knowledge either."

    Data is a small, measurable, description of the world:

- "The word 'quizzes' does not reveal anything particularly novel about Dave or his situation. Dave grades work from a relational algebra course, which is good knowledge about Dave and his roles and personal knowledge. Working within an educational facility also implies something about Dave. But a quiz is not a revelatory piece of information. It would be more revealing if Dave were grading Theses or Dissertations or reviewing Essays or White Papers."

- "This is data because it does not reveal anything about Dave or his network or his intentions. A Word document that contains no information does nothing to further analysis into Dave and his network. By the statement above, we already know that he is working on a conference paper."

- "The weather report is data about current weather conditions, since it is a measurable description of the weather."

    Data are the traditional building blocks of information, though the quotes above suggest that data stems from non-novel revelations. Data are measurable, small, recorded descriptions of the world or target.

    Information is an imprecise statement about the world:

- "This is more than mere data, because we know something about the sender and recipient of the letter. It is less than knowledge, because we do not yet know more about the sender or recipient, beyond their implicit connection to one another."

- "We know now that Alice is interested in building a device to cook her breakfast, but we do not explicitly know anything about the device itself. This means that we do not know how feasible her breakfast-cooking endeavors are. Because of this, the sentence is merely information and not yet knowledge."

- "We know that someone is planning something, but we do not know explicitly what they are planning. The details of the planned activity would make this into knowledge."

The term "more than" suggests a traditional hierarchy. However, unlike most hierarchies, information is used to identify statements that do not reveal useful patterns due to their level of detail or other imprecision.

Knowledge is relationships or patterns:

- "This is knowledge about the two targets within our tracked network, because it defines loosely the roles that the two targets share between each other. Bob is a keeper of how-to information, and he passes it to Alice. Alice may be a leader who in turn passes this information to the rest of the network. She may be a worker-level associate who executes construction of the machine based on Bob's information. She may be a connector who acts as a conduit to spread information on machine construction throughout the network."

- "This is knowledge because it implies a relationship between the two targets – the sender and receiver. A short story may imply a personal relationship between the two targets, or it may reveal that they fulfill roles in publishing or media."

- "If we are seeking to understand Bob's character, we now know that he enjoys symphony music. Beyond this, we can infer little from the sentence above."

- "A statistical profile may help reveal patterns of behavior in a target. This would make it knowledge about the target's behavior. There are a lot of unanswered questions here though: how big is her sample? How many different activity types by the target are included in the sampling? How many external environmental variables (associates, external events, etc.) have been considered in the analysis? Finally, the danger of statistical profiling is the Black Swan. Ultimately, this is knowledge that leads to error."

Knowledge is a pattern or relationship among entities. It allows for understanding and prediction through application of patterns to future events.

The most important aspect of this survey is that it is completely analyzable according to the very simple analysis methods that this section employs and that it produces no products that would have it be viewed as an outlier. If I were not the one doing the analysis, it would be difficult to write a rubric that would capture the different understanding of instructions sufficiently for the coders to catch it accurately. Future iterations of this research must have more questions and better instructions that can detect this level of wrong model-of-survey in potential participants.

8.7 Survey Analysis

The results from the surveys successfully triangulate my results. While limited by the same potential bias as the interview analysis, the high variation in conceptions of data from the survey suggests that I may very well be on to something. On the other hand, the phrasing of the survey instructions suggests that I may have indicated to my participants what I was looking for, thereby eliminating scientific rigor from my experiment. For this reason, the survey results are presented as a curiosity, rather than as proof of different conceptions of data in the wider world.

Different roles have different uses for data, and there are differences of opinion as to the methods of knowing, intrinsic nature, and purpose of data. In the surveys, people referred to records, observations, and numbers, roughly mapping to the terms found in the interviews.

Records, as technologically stored signs, can be data if they are stored properly. Another term for record is row or *tuple.* These tuples can be about any topic and can suffer from anomalies: authoritative conflicts in which the database may not represent the thing in the universe of discourse. Despite this, to some people, any records in the database are still data.

Observations are the act of applying consciousness to a reality. They can be subjective and inconsistent. At the same time, researchers used the term to indicate experimental observations that are objective and provide their own context. The term observation in the surveys suffers from the same traps that it does in the interviews and is worth far more research.

Numbers as data are used by both the database designers and the objective researchers. Numbers are either a more objective refinement of an observation or merely a

semiotic rendering that requires context and analysis to render into meaning. Numbers are the *sine qua non*[*] of measurement, as they are the result and the goal of factual inquiry. Those who believe data are facts can use the idea of numbers to represent this.

However, numbers are also the *without which not* of databases, being a different order of thinking about a record. A number in a database sense can be either the content of a record or the sign by which a record is stored, because a series of bits are simply numbers. Both uses are evident and present evidence of a fascinating linguistic trap for future research.

From the point of view of question of interest 1, the surveys were a complete success: people demonstrated that they had realities of data that departed in ways far beyond mere semantic differences over subjectivity and objectivity. Furthermore, different roles tended to cluster around different understandings of the term, suggesting that educational background combined with the current uses of the term strongly inform meaning. However, from a methodological perspective, the tainting of the surveys with the item of interest leads to results that are strongly suspect. While these results are encouraging, no real conclusions may be drawn from them.

---

[*]   Without which not.

9 Different Realities of Data and the Database

How can different realities of data can be applied to the philosophical conceptions of the database as corporate memory? Different realities of data provide a novel and interesting context in which to think about corporate databases and database design. To a database, data is an instantiation of a sign. Just as words represent both signs and reality, tuples[*] in a database represent signs and the reality behind those signs. When we define data as an instantiated sign, the nature of data in a database becomes very important. Using the communicative definition of data, a database contains a number of signs and the explicitly encoded relationships between the tuples.

In relational databases, there exist two categories of relationships between tuples. The first, and simplest, is a grouping of tuples into tables. A table exists as a set of identically formatted tuples representing different data of similar type. The second type of relationship between tuples is relationships between tables: while tables group tuples together, tuples can relate to each other based on internal codes or information.

The combination of these relationships allows for the computer to encode signs as data in the form of tuples and to relate signs to each O, allowing for manipulation of the signs to create new signs, and for information processing to occur outside the brain. In database design, the term "business rule"[†] defines a statement that shapes the nature of the database or the way the database manipulates information.

From a semiotic perspective, we can interpret these business rules as mental maps: they exist to shape, form, and interpret the data through our understanding of the local

---

[*] A tuple is a more formal name for "row" in a database, or a set of data of different domains stored together.
[†] A business rule is a constraint or action imposed upon the database by the business reality within which the database is functioning.

reality of business practices and norms. Through business rules interpreting data, the computer can act on the data, pass new data to other business rules, or change the business rules themselves: the exact same relationship that a mental map has with signs in a conscious system, albeit on a much more simplified level.

The act of designing a database is strongly informed by the designer's constructions of data. The mapping of relationships and the nature of each table is a reflection on both how the designer understands data and how the designer understands how the ultimate users understand their data.

9.1   Case Study

Consider a case of a fictitious winery trying to create a data warehouse. I will use this as a case study in which the ramifications of realities of data are illustrated in this chapter. This case study was developed for a data warehousing course at my old university and is of sufficient complexity to illustrate some of the issues raised in this chapter.

> TigerWine Distributors was founded in 1985 by two college friends, Miss G. and Mr. R. The company was originally based on a small family winery, founded in 1878 and passed down through Miss G's family, and has since grown into a very successful wine distribution company. Today, the original winery is only a small portion of their business. The California based company now trades as an importer and distributor of wines throughout the United States. Over the years, TigerWine Distributors has built up an extensive portfolio of premium products and represents a large number of national and international wine wholesalers as well as many small vintners.

> One of the greatest strengths of TigerWine Distributors is that it has contacts with the complete spectrum of the wine-sales trade from small, but prestigious, restaurants to the largest liquor stores and supermarkets. Their customers also include private individuals, regional vintners, traditional wholesalers, individual cash-and-carries, specialty stores, and other independent outlets.

This setup is an excellent foil, as the hereditary views of Miss G's understanding of data as tempered by her university experience could quite plausibly differ from Mr. R's

understanding, especially if Miss G studied a more scientific viticulture major versus's R's business degree. If these managers commission an analytic data warehouse for their distribution business, it is reasonable to suppose that they will understand data in different ways. To these two individuals, we will add a third, a Mrs. B, who is their Chief Technical Officer and the person developing their data warehouse.

For purposes of this case study, Miss G considers data to be reflections of objective reality, most comfortable manipulating data as observations from chemical studies of grape biology. While she admits that there is business data to be collected and analyzed, the data exist only in sales records: the movement of real things and the income derived from their movement.

Mr. R considers the most important "data" in their records to be the intangibles of customer satisfaction and error rates. He has heard that the data warehouse could explore these questions but is not entirely sure what black magic needs to be applied to the "big data" for that exploration to occur. In many ways, Mr. R considers data to be encoded human communications and is interested in how to decode those communications.

Mrs. B is the CTO, and therefore considers all data to be subjective recorded impressions. While she considers this in her role, she also needs to design the database so it can fulfill the requirements of the other two founders of the company. For purposes of this case study, Mrs. B will be making a data warehouse that only covers the distribution side of the company, and will be trying to understand the realities of data held by the founders.

Designers can encode business rules in a database in two ways. First, designers can build them into the design of the system, informing the grouping of columns, the design of tables, and the relationship of one table to another. Relationships are a way of encoding the constraints of reality into a constructed simulacrum, a database. A constraint is a

337

database's way of enforcing a rule. A relationship between tuples is defined as: "this tuple may not exist without a reference to another extant tuple." This both encodes the specific tuple the entity being created relates to, and enforces articulation of that relationship.

In the context of TigerWine Distribution, the table business rules look like:

CASE

A collection of items offered for sale; this is the fundamental unit of sale for the distribution side of the business. It can be: - a case of wine that is not on a special promotion (case-wine) - a case of wine that is on a special promotion (promotion-case) - a special promotion package that consists of one or more bottles of wine and one or more other items. (promotion-case)

This is a straightforward and unambiguous presentation of what the table contains that would be accepted by any of the realities of data. The practice of the different realities does not rely on the different formulation of specific business rules, but the very choice of specification of the rules.

Therefore, for purposes of the case study, the question is to illustrate a table or relationship that each of the three understandings would find particularly compelling or distasteful. For data as objective measurements, Miss G would find the presence of a table of written employee evaluations without any real "data" to back it up most objectionable. It would illustrate the worst of subjectivity, without even offering the benefits that computerized semantic analysis of a mass of product reviews. The tables she would most approve of would be the automated sensor readings from the distribution warehouses of the stock that was on each shelf. Miss G would also object to normalization because it splits one "observation" across multiple tables; as the meta-data for the observation belongs in a different table than the data.

Mr. R as someone who holds data as subjective observations, is nominally "okay" with "anything" in the database, so long as it accurately reflects reality. Mr. R is not

interested in the merits or lack thereof of normalization, so long as the database accurately reflects the subjective impressions of reality that he wants it to encode. In a way, this means that Mr. R treats all entities within the database as "equal data" with no one tuple privileged against any other due to its intrinsic truth or relationship to the universe. Mr R especially dislikes tables that are not internally consistent within the rules of the database: just because reality gives people multiple phone numbers does not mean that there should be multiple phone number fields in the "person" table, even after the fact. For Mr R, if reality changes, the database should be reengineered to reflect that reality accurately without consideration for the costs or impacts on established software. Without a consistent and useful database structure, the queries to manipulate and transform the presentation of data in the database will fail.

Mrs. B considers data to be encoded human communications or encoded information. Therefore, the database tables are a repository of human interactions with the world, curated and made searchable by other humans. Any design that interferes with that curation is discouraged, as it makes the task of the maintainer that much more difficult. Because the tables in a database have on ontological value to Mrs B, she is comfortable with whatever arrangement of tables that will work best with the programs that support human communication. Any table arrangement that exists for the data's sake rather than to support the programs that support human communication is pointless and wasteful effort. For example, normalized relations that introduce inefficiencies for the purpose of avoiding rare or even purely theoretical data anomalies is to be fiercely resisted: it is inefficient and does not serve human communication. Instead, it gets in the way of programmers and therefore makes the encoded data less likely to be presented to the users.

The O, more significant way, is with queries[*]: exposing and relating signs in a pre-determined format. The real power of queries, however, is that the data they output can be used as the inputs for subsequent queries: just like a sign or group of signs, one mental model can output data to another mental model for further interpretation. A query draws upon the encoded business rules in the database and applies some criteria to find only interesting entries of interest. These entries can be composed from multiple tables, even positing characteristics not originally planned for in the database but derivable from available evidence.

The importance of this similarity cannot be understated. When I discuss databases, I am talking about the seat of external memory and cognition for distributed cognition. Whereas other entities shape and perform input/output on the database, as well as on their own interpretation, most of the processing in a database happens through the use of business rules, either as constraints on database design or the implicit contextualization of data via queries.

Queries with regards to different realities of data, are simply structures to produce the desiderata as a function of expectations of the querent. The actual structure of the query to produce the same information varies little based on the coder's conception of data, as there are "right" and "wrong" ways to access information in a database dictated solely by the efficiency of a particular query.

While there are techno-philosophical approaches on more sophisticated queries with their representation within the database and the methods used to construct them: mainly varying the complexity of individual queries and whether or not they the output is collated within the "host" programming language. These choices of programming and

---

[*] A query is a way of retrieving sets of tuples from the database.

efficiency are subtly influenced by the realities of data of the programmer.

Specifically, the choice of whether or not to perform the desired operations within a purely database context or to perform intermediate stages within a programming language can certainly reflect the programmer's conceptions of the "atomicity"[*] of data. A programmer who views data as non-atomic will be more "comfortable" making giant queries as the idea of splitting apart tuples and recombining them in an extended chain does not produce a cognitive dissonance in the programmer's mind. However, for realities of data that are less about semiotic manipulation and who consider a tuple an observed or recorded fact, intermediate translation stages within the context of a more comfortable functional programming language would necessarily appeal more: there is no violation of objectivity in operating upon retrieved data as contrasted with reformulating the presentation of that retrieved data within the data retrieval system itself.

Assuming that all three founders are sufficiently comfortable with programming languages and database systems to write queries, we can approach this from their perspective. Miss G, with her reality of data as objective representation, would prefer to retrieve numerical data gathered by automated systems and sensors. She would consider much of the ancillary data in the system to be "metadata" supporting the certainty of various conclusions that could possibly be made by manipulating the "true" data within the system.

While she recognizes the existence of text in the database, the presence of the text is considered to be merely a way to inform a website, rather than "data" to be collected and analyzed like the data she gets from sensors. Her queries would focus on data retrieval

---

[*] Atomicity: whether a given piece of data can be divided into associated attributes

instead of manipulation, and the statistical summation and estimation necessary to make conclusions from the collected data would be performed in a sufficiently powerful mathematical system of her choice. "Tampering" with the data through sophisticated SQL queries, especially ones that filter or remove data on anything more than simple "boundary" filters (between this time period, from X sensors), would be tantamount to heresy. Removing the atomicity of the tuples in many ways invalidates the objectivity of their storage choices and recording methodologies because it is changing the context after the "experiment" (in her eyes) started.

Mr. R, as a business type, is most interested in tracking consumer sentiment over time. He also sees data as the most ephemeral, human, and subjective. Data are observations, to be sure, but the are human observations. SQL is powerful because it can produce unusual and post de facto relationships through the mechanism of queries. Mr. R is the most comfortable of the three producing "new" information through the transformative properties of queries, and sees nothing wrong with encoding a long query that touches many tables to produce a novel and interesting output. In a sense, Mr. R. would like to be able to "persuade" the database to produce the results he wants.

Mr. R is most interested in manipulating and analyzing the text as input in customer and performance reviews. Because nothing in the system is objectively true, the things which are more rich in human meaning, e.g. words, are more useful than the things which require significant statistical analysis to extract meaning from, such as numbers produced via sensor observations. The presence and analysis of the text in the database is an interesting challenge, rather than an irrelevancy.

Mrs B. is not particularly interested in the analysis of stored data. As all of this is a stored semiotic representation of human information, the primary purpose of the database is to efficiently retrieve wanted records for humans to manipulate without

any incorrect or wasted retrievals. As efficiency of communications is everything and humans are necessary for transforming the data into information, Mrs. B is less interested in automated analysis and complex queries than R or G. However, she recognizes that these tasks are desirable from R and G's standpoint and will try to design the database accordingly, creating specialized "views" (permanent queries which can be treated as tables) to suit both of their requirements. Her queries and manipulations of the database will be designed around efficiency and speed of retrieval and access rather than the semantic content of the data itself.

9.2   System Models and Mental Maps

The semiotic concept of the system model for databases [60 and 61] is the mental map of the designer instantiated in business rules and relationships[*]. The mental mapping is the functional product of the user's internal representations of data. This map is exploring the database designer's environment, both in the explicit requirements of the system given to him by the consumer and through formal and informal observations of the old system. The designer's understanding of requirements informs the design and implementation of the intended system model.

A trivial example of a system model mismatch can be seen in the finite state machines of a typical Sydney mass-transit ride. The correct system model of a typical subway service is that the user starts by purchasing a magnetically encoded ticket, pushes it through the machine at the front, and at the destination. If there are no more uses left on the ticket, the machine captures the ticket.

However, some users may have an "insert ticket in starting machine, then discard"

---

[*]   Relationships are specific business rules that delimit how different tuples can or cannot relate.

model in their heads, leading to problems when the inevitable ticket checks occur on a train. This confusion may be compounded by those stations that are open to the train platforms, because they are too small to justify the staff that ticket-taking stations would require. Thus, the user's actions with his or her ticket may not correspond to the *expected* actions by the designers of the ticket system, and from this basis, mistakes occur.

From the perspectives of the three founders of TigerWine distributors, they have been working together long enough for a shared pidgin to spring up around the database. Just as in the interviews the term "raw data" was used, the three founders would evolve similar jargon to map specific concepts within the database to their own internal realities of data. However, at the creation of the database these pieces of jargon would not have yet evolved.

One of the initial debates the three would have at the creation of the various datamarts supporting the distribution business. As the framing document states:

> TigerWine distribution business has now grown to the point that the partners would like to begin anticipating business trends. They want to be able to predict patterns in wine availability along with trends in wine drinking so that they can dynamically adjust their business to maximize sales. It would be helpful for them to have a data system (warehouse or data mart) that allows them to track changes in the products available from different suppliers, changes in the availability of various categories of wine, and changes in the buying habits of their customers. They would also like to analyze the efficiency of their current business processes, especially their warehouse inventory management.

Miss G would be most interested in the objective and discrete changes in product availability. These are based on objective business measurements that do not require interpretation. As such, the analysis of buying habits is something of an imponderable that is a useful pipe-dream but not discoverable from the "database" they have. Data, to her, is a purchase or a price.

Mr. R is certainly interested in changes in product availability, but inasmuch as that

reflects the *intentions* of the businesses that they are interacting with. The data warehouse being developed is a way to get inside the minds of people the business interacts with, and therefore Data is a discrete decision recorded by a computer. He would prefer more effort be spent in capturing entered text, as that provides more and better clues as to the decision making process than simple sales numbers.

Mrs. B is interested in the technical challenges of interpreting these disparate forms of human communication. As the author of the "extract, transform, and load" process, Mrs. B. will be responsible for the manipulation and transformation of the various external data sources that the data warehouse will be connected to. For her own use, she is most interested in data that is well documented, as that gives a better understanding of what the human communication is about. This documentation then allows the programs to provide better contextualization of the data for human consumption.

9.3   The Consequences of Error

The system model interprets reality into an abstract model of relationships and norms or business rules. A good system model corresponds well to reality: relationships in the map correlate with reality. The proof of a pudding is in the eating and the proof of a system map is in the using. Systems generate system maps in their user through exposed affordances. If the affordances designed into the user interface correspond with how users think they should behave, the users have a good system map.

A system model that is poorly envisioned or implemented, either because of design errors or misunderstanding of the user's realities of data, causes multiple types of potential error [62]: data corruption, data deviation, and misinterpretation. As the system map deviates from reality, the incidence of the three types of error named above becomes ever more likely.

Data corruption stems from where a system model does not allow or require sufficient updates from systems that observe reality. As parts of the system correspond to earlier realities, they increasingly fail to reflect the current reality. Errors breed. In a system where an interaction produces the right result for the wrong reason, the user does not become aware of their mistake. In fact, a mistake only occurs if there are observed secondary effects that were not desired by the user. However, further interactions may lead these second (and third) order mistakes to create their own changes to the system's state, turning what was an unnoticeable mistake into something that can bring the entire system to a screeching collapse. An example of such a problem is seen in preventative car maintenance. If a car is improperly maintained, nothing will happen initially. As systems fall out of timing, lose their lubrication, and suffer other failure-of-maintenance effects, these failures cause *other* systems to break down. These secondary systems *may* light the "idiot lights" in the dashboard, but for a bad driver, it is not until a furious boil of smoke escapes from the engine that they are confronted with their mistake. If those incidental problems are fixed without due thought, the underlying causes will simply cause *other* systems to fail until the car is written off as a total loss. This unfortunate outcome is purely due to the user's system model diverging from the car's reality.

The second type of error is data deviation. Whereas data corruption stems from data falling out of synchronization with reality, data deviation starts with the data being out of synchronization due to a database design that cannot accurately reflect reality. Incorrect assumptions, norms, or ignorance can lead to data deviation. I can draw a trivial example from a special case of postal mail: "general delivery" that may instruct mail to go to a centralized post office to be held for pickup. Corporate database systems may not be designed to recognize general delivery, nor may they have business rules suitable for letters addressed under general delivery notation. Geographic Information

Systems have even had problems mapping other post office special cases (Postal Boxes) to geographic location for purposes of illness tracking [63]. These problems are data deviations, in which reality does not conform to the database and errors occur when there is an irreconcilable mismatch.

Misinterpretation, on the other hand, stems from problems in Os' interpretation of tuples in the database. As the interfaces access the database, the system models of the interface designers may map the meanings of the relationships and the tables in the database incorrectly. This leads to interfaces which present the context of data incorrectly and thereby a cascading failure of meaning which eventually leads to mistake and error.

Corruption stemming from different realities of data requires more research at this time. My research did not provide sufficient evidence to make useful specific examples of the secondary and tertiary effects described above.

The data warehouse envisioned by G, R, and B, may experience some deviation. Data deviation would stem from a mis-implemented ETL stage by Mrs. B. In this deviation, if the transformation misunderstands the reliability or meaning of the column being translated, the new column could appear correct in its new context but have its meaning shifted significantly. One probable cause of this error is misunderstood labels for financial transactions, altering the "start and stop" time periods of some transaction aggregations.

Misinterpretation in Mrs B's data warehouse would stem from Miss G and Mr. R misunderstanding the documentation of the tables. While this looks like data deviation, the error of misunderstanding happens during data consumption rather than data recording and is a function of in-company documentation and communication practices. Because miscommunications occur when the designer fails to express her intentions correctly, it is a discrete category with different failure modes than deviation.

9.4   Database as Corporate Mind

It may even be possible to extend the new understandings of the reality of data to an exploration of the philosophy of the corporate mind, for a corporation's memory is in its databases. Donald Cunningham explains John Deely's theory of human experience when he states [64]:

> *Umwelt* model and argues that our cognitive experience is mediated through a labyrinth of signs, a personal world, created by the organism via species-specific sensory characteristics and particular experiences in the physical world. This *umwelt* comes to define those things that we pay attention to and consider as "real." Humans, uniquely, can create signs that go beyond immediate experience: Words, pictures, bodily movements and the like can become part of our *umwelt* even though they may mark objects that have no literal basis in the physical world.

This theory of human cognitive experience can be extended to memories. Only that information which is part of the corporate databases can be considered part of a corporation's *umwelt*. Although individual agents of a corporation may be aware of reality not otherwise informed by the databases, the corporate entity can only perceive reality through the inputs and outputs of its databases. This important link between the database and reality requires that the database designer's system model corresponds well with reality to reduce errors in database design.

A tragic example of the problems of corporate memory occurs in the story of the American MERS Corporation, a tiny corporation that maintains a database of most mortgages held and transferred for securitization [65].

> The implication is that MERS is superior to the local courthouse system. The evidence is the reverse. Chris Peterson has described the utterly unorthodox corporate governance system of MERS, where employees of other firms put on a MERS hat for a short period as a "MERS certifying officer" and execute documents. MERS does not supervise these individuals. Indeed, it specifically disavows any responsibility for the accuracy of MERS's records: MERS makes no representations or warranties regarding the accuracy or reliability of the information provided. MERS disclaims responsibility or liability for errors, omissions, and the accuracy of any information provided. MERS does not input

any of the information found on the MERS(r) System, but rather the MERS Members have that responsibility regarding mortgage loans in which they hold an interest. ... And there are widespread indications of member non-compliance with MERS processes, such as making assignments without having proper corporate authorization (including out of bankrupt entities). Moreover, the updating of records by MERS members is strictly voluntary.

Thus, banks are using MERS as their memory; they cause it to be a model that reflects their reality. However, they have no obligation to keep the model accurate, and have no recourse or backups when problem arise because MERS is so much more efficient than the traditional paper audit trails. This conflict of priorities produces predictable results: an error-filled foreclosure process with many parties fighting over the reality of possession before courts and the U.S. congress.

To a corporation, errors in database design allow signs to be encoded into the database that do not correspond to reality. Therefore, the corporation's actions will deviate from the most optimal, because the corporation is endeavoring to solve a problem that only has a tangential relationship to reality. Poor system models beget worse system models. It is inevitable that database design passes from the initial designer to subsequent designers employed to make changes to the database.

With a poor system model, only the most glaring of errors advertise that they are errors. Other errors simply alter the corporation's responses to reality without causing agents to realize that the responses do not perfectly correspond. In the latter case, as secondary designers attempt to change the database to adapt to changing business conditions, the initial system model residing within the designer's imagination influences all subsequent system models. Without careful consideration and reflection at each stage, the errors induced by the database structure cause the database to diverge further and further from reality.

In context of the example of TigerWine Distribution, they suffer from a glaring error in their system model: wine is identified by the case, and only with difficulty as individual bottles. Because they associate wine with cases of wine, the most obvious error is that they won't identify other opportunities to sell bottles or unusual containers of wine. The "with difficulty" above indicates that their marketing department forced through a change: promotions could contain bottles of wine as well as unusual items like corkscrews or cheese. This presents a logical disconnect within the database itself: half the company understands wine as something which comes in a case, and the other half understands wine as something that appears inside a pre-packaged promotion. While the humans involved are certainly smart enough to understand the em nature of a case of wine, the fact that these ontological descriptors of containers are hardcoded into the database informs the corporate identification of objects and its behavior towards them. As the technology ages and new systems are built upon the old technology, it is quite possible for these artificial divisions to be encoded in the new systems like unto laws of nature. This subsequent generation is when the "schizophrenia" sets into the corporate mind, as the databases state a thing that is simply not true in "reality" but fully accepted as true within the corporate memory.

Fixing errors in a production system is difficult as "the cost of fixing errors grows exponentially as a function of elapsed time to discovery." [66] Since these errors exist in a production system that has run for some time, the cost of fixing the errors is prohibitive at the stage when the errors are found. The only way to insure that the system model of the database designers keeps the system in line with reality is to inspect the database and insure that it corresponds with reality within the required specifications.

By understanding the impact the various users' realities of data has on a database and on corporate memory, this philosophy of corporate mind combined with an under-

standing of how we conceive of data can provide tools for a data modeler to create better databases. Normal database modeling methods are quite formalized, exploring the relationship of tuples to each other and whether or not to split them into new tables. Other database design methods simply instantiate the tables of a "default" database and allow the designer to connect new tables and data structures to the pre-extant core[*]. In many ways, most database-creation methodologies already expect the designer to understand the participants' realities.

When teaching database design, I had students create data-flow diagrams mapping the current reality. They then updated the DFD's so that they could reflect some sort of more coherent "future reality" and created the database design from that. The process of updating designs from current reality to future reality was never explicitly spelled out, merely left as a "well, you should solve this yourself" method for the students. An understanding of the various realities of data, and especially of how different people understand data, may provide for better database design.

An understanding of data will provide for better design neither by increasing normalization[†] nor by providing better requirements-collection methods. Instead, the discussion of realities of data will provide for better design by giving the practitioner more tools to understand the clients' requirements. Although this will not help with those clients who insist on a turnkey design, the understanding of the different realities will help render what the clients think is data into the database, including elements of data that are not currently present in the artifacts[‡] inspected by the database designer.

---

[*] Django and Ruby on Rails are excellent examples of this modeling architecture.
[†] A well-accepted set of methods for removing potential errors in a database at the cost of efficiency.
[‡] Papers, old databases, old software systems, and the traditions and rituals of employees as manifested in things.

10 Conclusion

This research aimed to understand how users perceive the nature of data. In a more limited sense, I wanted to demonstrate that users have different conceptions of data and a method for discovering hints as to how they use and understand data. In that light, I have been successful. This research should serve as the foundation to more specific and focused studies into the different social constructions of data present in our societies and what epistemological and ontological basis data may have.

With the use of the term "raw data" there is some evidence for trading zones being established as people with different usages of data must discuss it to function in the same way. In a more limited sense, this evidence may just be an artifact of the narrowness of the study. Either way, more indication into the establishment of trading zones with different technological concepts is strongly indicated.

10.1 Results

People have different realities of data. My methods are capable of generally distinguishing between different realities. I have found three distinct realities: data as communications, data as subjective observations, and data as objective facts.

In the data-as-communications category, users generally consider data to be electronically encoded semiotic representations, or bits stored on a computer. Although some people in this category believe that data are any communications (i.e., data are signs), most believe that data are re-encoded and computerized signs. This construction tends to share some space with the other understandings in the sense that people can use the term "data" in an electronic sense when referring to "computer data" and in

the "proper" sense of facts or observations. However, even in that case, there is more a blending of the two meanings rather than two distinct meanings.

There are people who classify data as observations. Observations, in this case, are any acts in which a sensor (such as the brain and eye) takes note of the world. These observations are then filtered according to some knowledge and situated with context as information. All observations are data, but not all data is important. Only data situated properly in the correct context can be assessed for utility.

Considering data-as-facts, people equate facts with objective, true, statements about the world, and comprehend data as small facts about specific parts of the world. Facts are a superset of data, information, or knowledge, being objective and truthful statements about the world regardless of scope. People in this category tend to favor strongly objectivist views of science and tend to be scientists and some types of engineers. Data as facts does not require external context, and is packed with its own meta-data: its precision, provenance, reliability, and reproducibility, among other things. Data as facts strongly tend to be numbers.

It is clear from the, admittedly biased, results of the surveys that many other variations exist on these three simple categories. These findings are in no way exhaustive, merely suggestive of what I found in the majority of my interviews and surveys. It is obvious to me that far more research is necessary before any definitive statements can be made about the different realities of data present and, especially, their possible interactions and failure modes with each other.

The Social Data Flow Network has great utility as an reality-discovery methodology. The SDFN, inspired by the data flow diagram methodology of systems design, is quite able to cause participants to differentiate between practical definitions and theoretical definitions when discussing a participant's construction of data. And is therefore better

than more focused requirements gathering methodologies at this task.

The SDFN produced many useful discussions of data, causing participants to reflect on their own working definitions and on how they contrast with formal or intuitive definitions. The propensity for generating self-reflection is a useful product of the iterative categorization of the SDFN. This process of iterative categorization provides mechanisms for inspecting peoples' functioning definitions, rather than the theoretical definitions under discussion.

The internal consistency of the results between the analysis of the transcript and the analysis of the diagrams suggests that the generation of the SDFN diagrams provides a useful triangulation method for both reflection during the interview and subsequent analysis. However, the SDFN requires significant refinement: it must be made less intimidating; steps must be made to make analysis easier; it needs to be codified such that other people can perform both the collection and analysis stages of the SDFN..

10.2   Methodological conclusions

The interviews via SDFN and surveys were successful. The SDFN diagrams were useful in producing directed discussions which uncovered personal realities of data. My chosen research group, the fine engineers, scientists, and researchers at the company I conducted interviews at, were a fantastic initial target. The group presented a number of different trading zones as scientists and engineers interacted in a highly competitive area. The research was both practical and theoretical and had to show progress on both fronts, creating many different goals and many different research languages.

The informal atmosphere also presented significant advantages. My interactions with the team allowed for a sense of reassuring informality in the interviews while the non-interview interactions (mainly sharing meals) allowed for a sense of rapport that

helped the difficult sections of the interviews slide past. This informality allowed people to communicate their insights without worries that what they were going to say would backfire.

The survey in practical matters was a partial success. Although the second survey was generally able to probe the thoughts of the participants, a number of people thought the survey had a different intent, and thus their answers presumed different questions than those I asked. As such, the wording and the scenarios could be tightened. All things considered, however, the survey's results are acceptable for use, but could stand quite a lot of work to improve their reproducibility and accuracy. Nonetheless, the survey presented some useful insights as to how wildly different people understand the nature of data.

10.3 Further Research

Further research will be pursuing two primary goals: methodological improvement and philosophical exploration of the philosophy of data. This research is about applied philosophy: using philosophy as an engineer would use science. This research is designed to go some distance towards combating the "ivory tower" [67] perception of philosophy, especially philosophies of technology*. The ultimate targets for this research are the practitioners who use data every day: engineers, businesspeople, scientists, systems designers. With this research, I hope to encourage philosophical thinking in those areas and to demonstrate that many worthwhile philosophical activities must be rooted in a firm knowledge of their respective areas.

---

* Our philosophers of technology and science, in the main, are not practitioners. While their outsiderness lends some objectivity to their philosophies, their lack of capability in the fields of which they speak (especially high-tech) means that many actual practitioners in those fields feel no obligation to listen to what these outsiders are saying.

If this research can inspire another IST practitioner to explore aspects of the philosophy of technology in which they have a personal investment, then this research will have been useful. All the further research directions *must* be focused in integrating philosophy with practice. Research that involves practitioners and enhances their capabilities for meta-reflection is good research. Research that isolates practitioners and returns nothing is bad research.

This section cannot articulate all the possible research directions, for there are too many unknown unknowns. While most of the known research directions are designed around targeting or identifying flaws in the current philosophy or methodology, those are only short-term goals. In the longer term, as the rate of change increases in the world, our understanding of data will change with it, as data is the reified basis of software, in one way or another.

10.4   Final reflection

There is science to do and research to be done in understanding different conceptions of data. This dissertation merely paints the broadest outlines of possible questions to ask. The idea of different constructions of data offers many ideas to many disciplines, and borrows just as many. In the discussion of "what is data?", there is no consensus on an answer, and investigating that question of the nature of data and the trading zones constructed around it will pose the majority of the work for the next few research projects.

One major research direction for the philosophy of data is to define a set of open questions of the philosophy, questions that will guide researchers into more specific areas of inquiry, just as Floridi's open questions have guided the philosophy of information.

Creating a framework of research as a research project is something that will allow cooperation and mutual work without much unnecessary duplication of effort.

The results must be accessible. It is far too easy to scream, "This was a triumph," when the actual conclusions are hidden behind a wall of polysyllabic vocabulary and inscrutable jargon. The worst danger for this field is articulated by Eliezer Yudkowsky: "Well, sounding wise wasn't difficult. It was a lot easier than being intelligent, actually, since you did not have to say anything surprising or come up with any new insights. You just let your brain's pattern-matching software complete the cliché, using whatever Deep Wisdom you'd stored previously." [68] If research on the nature of data is reduced to pattern-matching Deep Wisdom without any pragmatic export or falsification of results, I will have failed.

People have different realities of data. They can think of data as stored signs of communications, data as subjective observations of the world, or data as objective and numeric facts. They can think about data in ways not articulated here, and I intend to find out what those ways are.

11 Appendix A: Survey Text

This file was generated from the "Print Survey" function of Limesurvey. It represents the questions asked to all survey-takers, but without any of the HTML interface improvements.

Exploring the Philosophy of Data

In this survey, we will explore the philosophy of Data. The first question will ask you to vividly imagine yourself in one of your past job or life experience roles, so to get you in the right mindset for the rest of the questions.

The 25 questions of the survey, expected to take around half an hour, will ask you to categorize very short scenarios as Data, Information, or Knowledge and then to quickly explain your categorization in a few sentences.

The intent of this survey is to probe your philosophy of data through indirect categorization, so as to differentiate your applied understanding from any theoretical definitions you have.

Welcome, and thank you for choosing to participate in this survey on the Philosophy of Data. This survey is being presented as part of data gathering for Brian Ballsun-Stanton's study, HREC ethics approval 09298. If you have any concerns, questions, or complaints, please contact the ethics secretariat. If you have any questions or comments about the questions below, please contact brianb@student.unsw.edu.au

This survey is completely anonymous and we ask that participants not take it multiple times.

The survey software will allow you to resume later by creating a login and password. Please do not use an identifiable user name for this. While it is impossible to associate your username with the survey, it is also useful to not be able to identify the username. At any time you may choose to exit and submit or exit and clear the survey. No record will be kept of people who exit and clear the survey.

There are 53 questions in this survey

Thinking about your role

This survey is exploring what you think about Data. To do that, the survey will present a list of short "scenarios". We will ask you to categorize the scenario as involving Data, Information, Knowledge, or something else, depending on your own understanding of the terms.

We believe that people can have different philosophies, depending on what job they're doing. For this survey we ask that you think about the scenarios from the perspective of one of your jobs.

---

1 This survey has requested that you answer it from the perspective of one of the jobs that you do. Please describe the duties of that job (in general).

Please write your answer here:

For example, I could answer "I'll be answering this survey as a Database Developer. A database developer designs databases, talks with clients, and tries to figure out what they do."

Scenario 1

I am trying to understand what you think of as Data, and why.

The questions below ask you to categorize the scenario, and then explain the categorization.

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Alice receives a **letter** from Bob.

The questions below ask you to categorize the scenario, and then explain the categorization.
An example category that "letters from Bob" may fall into could be "Data" If it is simply that, then indicate the category it does belong to.

Then please, in a sentence or two, explain why you categorized the scenario as you did.

---

2 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

3 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 2

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Alice receives a letter from Bob containing **instructions** on how to build a machine.

4 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

5 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 3

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Alice receives a letter from Bob containing a **short story** he has written.

---

6 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

7 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 4

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Alice determines the **locations for parts** of a Rube Goldberg style machine to cook her breakfast.

---

8 Please categorize the above scenario. *

Please choose **only one** of the following:

- ◯ Data
- ◯ Information
- ◯ Knowledge
- ◯ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

9 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 5

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Alice receives a letter from Bob. The letter is a **time chart** of what shows he has watched on TV for the last week.

---

10 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

11 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

```


```

Scenario 6

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Bob receives an e-mail from Alice, it is a **record of the daily temperatures** outside her apartment for the last week.

---

12 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

13 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 7

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Bob receives a flash drive from Alice. It contains **mp3 music files**.

14 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

15 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 8

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Bob attends a symphony with Alice and enjoys the **live music**.

16 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

17 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 9

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Bob **ignores** the traffic noise outside the symphony.

18 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [＿＿＿＿＿＿＿＿＿＿＿]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

19 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 10

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- As Bob is mugged walking home, the mugger **demands** his wallet and watch.

---

20 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

21 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 11

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Charlotte finds a microfilm in a hollow coin, it contains a **list of numbers and times** about something unknown.

22 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

23 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 12

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Charlotte finds a microfilm in a hollow coin, but cannot decypher the **code**.

24 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

25 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 13

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Charlotte finds **the secret key** to the code, and realizes it's a letter for technical support to the spy's handlers.

26 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

27 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 14

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Charlotte finds a microSD card in a hollow coin, it contains **a planning program** for something unknown.

28 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

29 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 15

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Charlotte creates **a statistical profile** of a spy, to predict their actions.

---

30 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

31 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 16

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Dave **lectures** to a classroom about database design.

---

32 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

33 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 17

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Dave grades **quizzes** from a relational alegabra course.

34 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

35 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 18

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Dave discusses the **reasons** behind one of Eve's incorrect answers.

36 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

37 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 19

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Dave writes a survey asking people to **describe their impressions** of a user interface.

38 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

39 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 20

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Dave saves an **empty word document** in preparation for his later work on a conference paper.

---

40 Please categorize the above scenario. *

Please choose **only one** of the following:

◯ Data

◯ Information

◯ Knowledge

◯ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

41 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

[   ]

Scenario 21

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Eve writes **poetry** describing the winter wind.

---

42 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

43 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 22

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Eve interviews students for the campus TV station and gets **short quotes** for her topic.

44 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

45 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 23

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Eve looks at **the weather report** and decides to bring an umbrella.

---

46 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

---

47 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 24

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Eve receives a letter from an ex-boyfriend, **telling** her to take her stuff back.

48 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

49 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 25

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Frank **selects** which instrument readings to include in his experiment.

50 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [   ]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else"

51 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Scenario 26

Please read the following one sentence scenario. Categorize the highlighted word or phrase in context of the scenario.

- Frank **designs** an experiment

---

52 Please categorize the above scenario. *

Please choose **only one** of the following:

○ Data

○ Information

○ Knowledge

○ Other [_____]

Please be as precise as you desire. If something is not easily categorized as "Data", "Information", or "Knowledge" You are welcome to use the "Other" choice to describe it more exactly. Perhaps it's completely different like "Emotion" or "Wisdom". Perhaps it's different from the other things you categorized, "Data as something" or "Information as something else."

---

53 Please explain in one or two sentences why you categorized the scenario that way.

Please write your answer here:

Thank you for taking this survey. If you have any questions or comments about the survey please e-mail me.

If you would like to discuss your philosophical intuitions about data, you are welcome to e-mail me. This conversation will be covered under a strong confidentiality agreement and will not be linked to your survey responses.
21.08.2010 – 00:00

Submit your survey.
Thank you for completing this survey.

12   Appendix B: Graphviz

12.1   pretty.gv

```
//BEG_G { graph_t g = graph("merge", "S") }
N [label!=""] {
        label = gsub(label, " ", "\\n");
}
E [label!=""] {
                node_t l = node($G, label);
                l.shape = "box";
                l.fontsize = "7";
                l.width=0;
                l.height=0;
                l.style="dotted";


                edge_t e1 = edge($.tail,l,"");
                //e1.dir="none";
                e1.arrowhead="odot";
                e1.arrowsize=0.5;
                edge_t e2 = edge(l,$.head,"");

                if (match(label,"Data") != -1){
                        l.color = "#aa0000";
```

```
                        e1.color = "#aa0000";

                        e2.color = "#aa0000";

            }

            if (match(label,"Information") != -1){

                        l.color = "#0000aa";

                        e1.color = "#0000aa";

                        e2.color = "#0000aa";

            }

            if (match(label,"Knowledge") != -1){

                        l.color = "#00aa00";

                        e1.color = "#00aa00";

                        e2.color = "#00aa00";

            }

            delete($G, $);
//      $.head.color = "red";

//      $.tail.color = "blue";

//      edge_t e = edge(t,h,label);

//      printf(2,"%s\n",e.head);


}
END_G {

      center=1;

//      nodesep=4;

//      ranksep=4;

      //colorscheme="dark23";
```

```
        startType="regular";

        overlap="false";

        splines="spline";

}


    12.2  general.dot

Digraph G {

        //Interview [X] [Date]


        Dev [label="noun 1 Developer", shape=circle];

        DSS [label="noun 1", shape="circle"];

        PAM [label="noun 5", shape="circle"];

        BF  [label="noun 2", shape="circle"];

        CRPC [label="position 8", shape="circle"];

        DPDS [label="noun 17", shape="circle"];

        CRSC [label="position 9", shape="circle"];

        BFPE [label="position 22", shape="circle"];

        BFPM [label="position 26", shape="circle"];


        Dev -> DSS [label="Code\nKnowledge"];

        DSS -> Dev [label="Textual Advice\nKnowledge"];

        DSS -> CRPC [label="Textual Advice\nKnowledge"];

        DSS -> CRSC [label="Textual Advice\nKnowledge"];

        PAM -> DSS [label="{noun 5} Real Values\nData"];
```

```
PAM -> DSS [label="{noun 5} Calculated values\nData"];

CRPC -> PAM [label="Derived Values\nData"];

PAM -> CRPC [label="{noun 5} Calculated values\nData"];
PAM -> CRPC [label="{noun 5} Real Values\nData"];

PAM -> CRSC [label="{noun 5} Calculated values\nData"];
PAM -> CRSC [label="{noun 5} Real Values\nData"];

CRPC -> BF [label="Setpoints\nData"];
BF -> PAM [label="Real Values\nData"];
CRPC -> DPDS [label="Derived Values\nData"];
CRPC -> DPDS [label="Real Values\nData"];
CRSC -> DPDS [label="Derived Values\nData"];
CRSC -> DPDS [label="Real Values\nData"];
BFPE -> CRPC [label="SOP\nKnowledge"];
BFPE -> CRPC [label="How to use\nKnowledge"];
BFPE -> CRPC [label="Daily Targets\nInformation"];
BFPE -> Dev [label="SOP\nKnowledge"];
BFPM -> BFPE [label="Monthly Target\nKnowledge"];
BFPM -> Dev [label="Production Philosophy\nKnowledge"];
BFPE -> BFPM [label="{noun 2}\nInformation"];
CRPC -> CRSC [label="Experience\nKnowledge"];
CRPC -> CRSC [label="Talk\nInformation"];
```

```
    BFPM -> BFPE [label="Production Conditions\nKnowledge"];
}
```

12.2.1 Terminal Command

Terminal command to run the above files:

```
gvpr -c -f pretty.gv diagram.dot | \
neato -Tpng > general.png
```
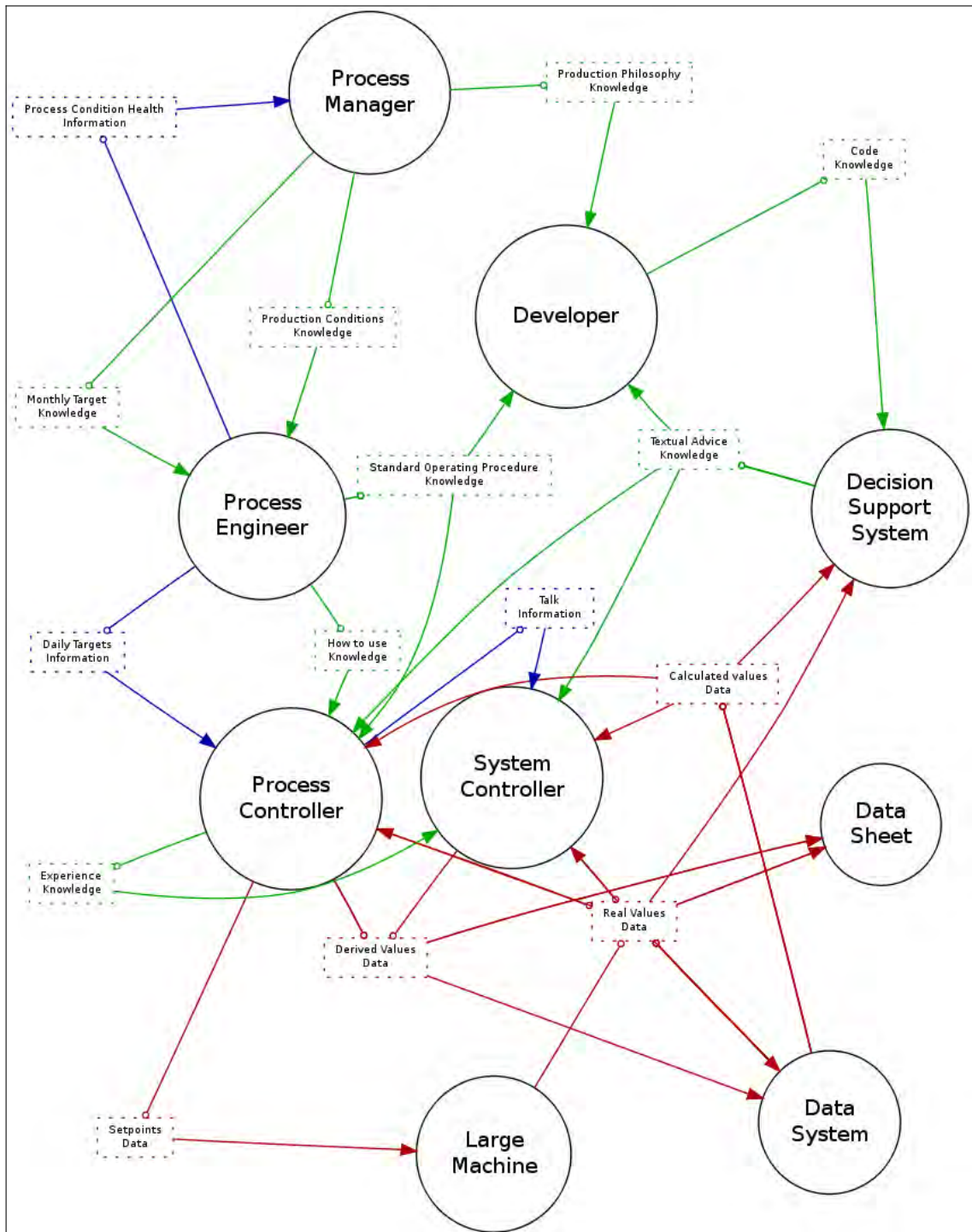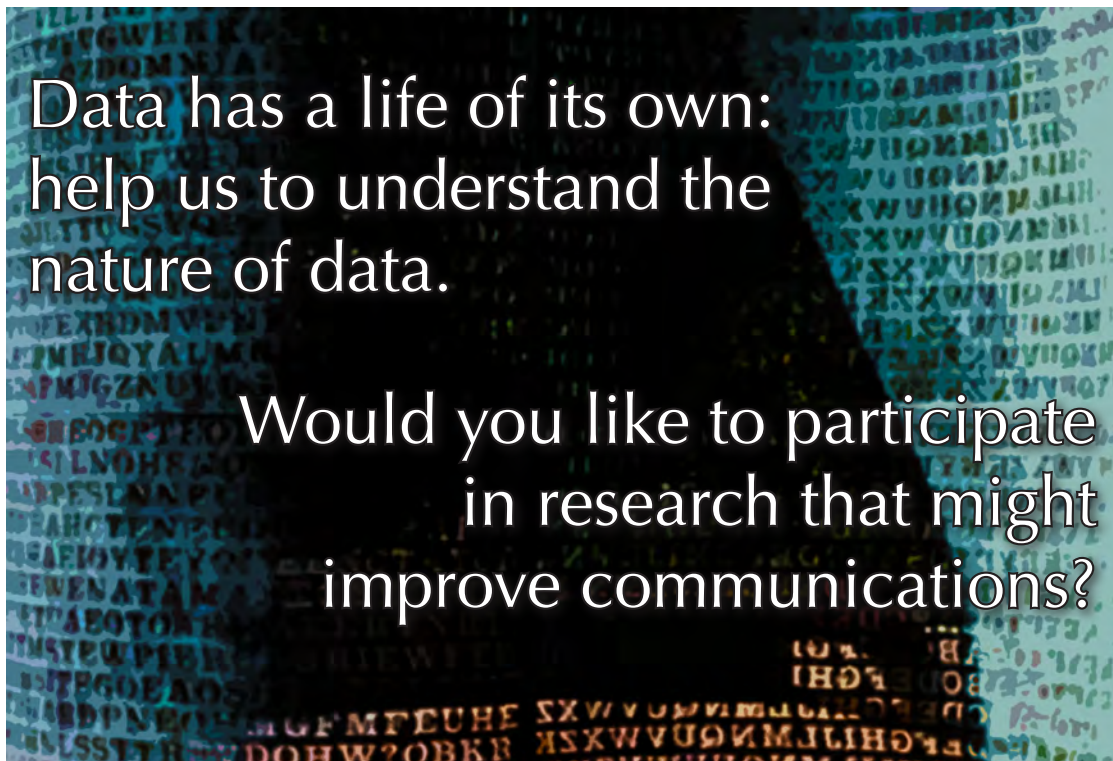
12.2.2 Output

Figure 12.1: This is a sample graphviz output produced from the preceeding elements.

13   Appendix C: Flyer

Data has a life of its own: help us to understand the nature of data.

Would you like to participate in research that might improve communications?

E-mail brianb@student.unsw.edu.au for information.

INVESTIGATING THE PHILOSOPHY OF DATA

We will be offering a survey in order to explore your understanding of the nature of data. We want to create a way to model how people think about data.

This study could lead to smoother interaction between allied groups, systems which better reflect their users' needs, and reduced error rates.

Our research uses absolutely confidential surveys to explore how you categorize data.

The survey will take less than an hour of your time. E-mail us to participate or for more information. All e-mails and participation records will be kept strictly confidential.

14 Appendix D: Asking About Data

This peer-reviewed paper was presented at the IEEE 5th International Conference on Computer Sciences and Convergence Information Technology in Seoul, Korea during the process of writing the thesis. It serves as a useful summary of my research and is presented here.

Asking about Data:

Experimental Philosophy of Information Technology

Brian Ballsun-Stanton
School of History and Philosophy
The University of New South Wales
Sydney, Australia
brian@ballsun.com

Abstract—**This paper explores recent research done into the philosophy of data. The research utilized experimental philosophy ideas combined with Information Technology methodologies to assess participants' philosophies of data. Reusing the concept of the data flow diagram, I suggest a methodology of experimental philosophy that allows participants to categorize flows into data, information, and knowledge. This allows me to explore their practical understanding instead of their theoretical understanding. My research has found three philosophies: "data as bits", "data as hard numbers," and "data as recorded observations."**

Keywords-Philosophy of Data; Experimental Philosophy; Philosophy of Information

I. INTRODUCTION

Many people believe that data is a technological construct, that we encode information and knowledge inside data when interacting in electronic systems. Other people believe that data are the basis of science: hard numbers as the product of experiments. That data must be objective, reproducible, with the limits of precision known. Still other people believe that data are an observation of some kind. That data can be qualitative or quantitative, so long as it is a recorded observation. People may use data in the singular or plural, not as a grammatical error, but as a reflection of how one understands this ultimately socially constructed concept.

These beliefs are incommensurate and largely incompatible. They influence thought, analysis, and self-reflection, and are strongly influenced by someone's background and workplace. In studying these different philosophies of data, held by people who work with data every day, I found it difficult to ignore my own philosophy of data. To state a philosophy by fiat destroys any possible evidence for multiple philosophies. For that reason, it seems better to set aside one's own philosophy of data and use the techniques and research results described below to question that philosophy.

This paper explores my experimental research into the philosophy of data. My research has two goals: to create a methodology to probe the philosophy of data for practical use, and to see if people really do have different philosophies of data.

The research had two primary goals, phrased as statements of interest to guide the abductive process of rapid hypothesis forming. I seek to explore the statement: "People have different philosophies of data" and the statement "My methodology can probe people's philosophies of data." These statements serve to focus attention and define a universe of discourse for the investigations.

II. JUSTIFICATION OF RESEARCH

Creating a philosophical basis for data requires significant justification. Information Technology (IT) researchers tend to spend a great deal of time and effort chasing after quite worthwhile new technologies without considering the philosophical implications of those technologies [1]. IT practitioners must serve as an interface between computing and people. They must understand what people actually want and must understand the reality *from which they desire that thing*. If IT people cannot understand the needs of the users and the reality that they live in, they cannot do their jobs. However, if they cannot then understand the philosophies encoded into programs via the many socially constructed protocols that a computer requires to be useful, they cannot understand what problems a computer system thinks it solves.

An understanding of the philosophy of data is not merely an academic question. Siloing in organizations, the practice of small groups talking mostly amongst themselves [2], may be partly due to different understandings of the nature of data. Imagine someone with one of the other philosophies described above talking to "a busy expert" about what he or she thinks the needs of a system should be. Without an awareness of the different definitions of data, the amount of effort needed to create a linguistic trading zone and actually communicate with this person about their infological needs is far more effort than a simple dismissal.

Exploring the philosophy of data is a gateway question. I seek to help IT practitioners to accurately model clients' views of reality, and then to entice them into other philosophical thoughts. The difficulty of modeling is that the client seldom explicitly states their understanding of reality. By building tools to probe those models, this experimental philosophy is both a vehicle for discovery and something that allows us to start feeling our way into the philosophy of IT.

III. PHILOSOPHICAL LITERATURE

A dominant philosophical theme in my research is the concept of a trading zone: two groups, not sharing a common

language, come to a place where they can evolve a locally functional language [3]. It is a way of communicating concepts between two groups without forcing either group to change what they know to be true. Both groups understand that the language is only *locally* true, not globally so. Trading zones between groups in an organization come from creating an evaluative accent to understand the other group's jargon.

Volosinov describes the theory behind an evaluative accent thusly: "Any word used in actual speech possesses not only theme and meaning in the referential, or content, sense of these words, but also value judgment: i.e. all referential contents produced in living speech are said or written in conjunction with a specific evaluative accent" [4]. Differences in this evaluative accent are very acute when it comes to the signifier "data" and that people apply their own evaluative accents based on their current job.

Combine trading zones with learned evaluative accents and we can see the roles of the philosopher and the IT practitioner. In the trading zone necessitated by two different evaluative accents of "data", my research participants coined terms for locally true definitions. They used the term "Raw Data" to mean readings directly taken from sensors, and the term "Derived Data" to mean calculations on those readings performed automatically. They used these terms *even if they had different philosophies of data.* The terms were part of a local language. The participants who were not part of the group however, did not use those terms.

This sub-discipline of the philosophy of data does not seek to have the same explanatory power of the Philosophy of Information. Floridi is investigating the philosophy of information from a traditional intuitive stance [5]. He is exploring deep questions of the universal nature of information, computation and AI. The philosophy of data should have more practical breadth than the philosophy of Information's theoretical focus because data is a socially constructed understanding that predates and is broader than modern technological usage. We are asking questions of how people perceive reality without exploring the deep questions of whether or not that perception is fundamentally true.

The most difficult and most vital component of any IT project is to understand the clients' realities. Many techniques have been found to probe these foreign understandings of the world. User interface and database design practices heavily informed my methodology. One user-centered design [6] practice that inspired this research is Joint Application Design. It is a method of gathering all the stakeholders in a room, enticing them to state their wants, and hoping they come to a compromise [7]. Philosophically speaking, a Joint Application Design (JAD) highlights the local languages *between* groups for the developers of the system. Through the stakeholders' conflict, the developers can start to gain multiple viewpoints into how the groups actually use the terms and what they actually want as described in their own reality. My intuition of the JAD's ability to discover local languages was one of the prompting points of the decision to use the tools of information technology to discover the philosophical background of the subjects.

The study of the philosophy of data is multidisciplinary. The techniques that information technologists use to probe the realities of their clients can also be applied by philosophers to the domain of Information Technology itself. These techniques are useful for the purpose of experimental philosophy. Experimental philosophy, instead of relying on the intuitions of philosophers, seeks to find validation [8] (or falsification [9], exploring paradigms [10], or refining research programmes [11], whatever works [12], *etc…*) in the expressed understandings of other people. This objective synchronizes extremely well with the techniques of IT.

Data modelers have long used the data flow diagram (DFD) to model the processes of interaction in organizations [13]. My research appropriated the technique to probe an individual's understanding of data flows. As the computing industry moved away from structured programming, the role of the DFD ceased to be a programming tool. Instead, it became a workflow visualization tool. The DFD now is used to build a consensus reality of all of the data manipulations present in the client's workplace. In my philosophical work, I re-appropriate the technique and use it to probe an individual's constructed reality. I used the technique to cause the subject of the interview to make rapid classifications of many different flows of data, information, and knowledge. Thus, the re-re-appropriated DFD forms the cornerstone of my methodology.

IV. METHODOLOGY

My work involved two distinct methodologies. The primary exploration technique used a new technique called a Social Data Flow Network, iteratively developed throughout an interview, as a way of extracting a participant's philosophy of data. Due to unusual interest, however, a survey protocol was developed to complement the interview by probing a larger audience through the internet.

A. Interview Methodology

When probing the philosophy of data of a relative stranger, two important tasks must be accomplished before engaging in any kind of theoretical discussion. The participant must be willing to talk and the participant must uncover his or her own actual understanding of data. Only then can we actually ask the participant to define data. If the definition stage occurs too early, the participant tends to fall back on dictionary definitions that do not echo their true understanding, because the definitions are so divorced from normal practice.

Interviews to probe the philosophy of data occur in three parts. The first component is that of the introduction and demographics. Following that, we move to the Social Data Flow Network creation. Then we conclude with a theoretical discussion.

The introduction and demographic section is vital. It serves to break the ice. Participants usually are quite nervous in these interviews as there is a sense that "strong academic rigor" will be required of them because of the word: "Philosophy." By allowing participants to talk safely about their background, a number of important interview goals are fulfilled. Their discussion highlights an important facet of their work that will serve as the basis of the Social Data Flow Network. It also

provides the normal comforting elements of an icebreaking question [14].

The second stage begins when the participant is comfortable in the interview. This stage is the creation of the Social Data Flow Network, a term created to differentiate the diagram from the typical Data Flow Diagram and ultimately ignored for the practical term bubble diagram. The process of creating the bubble diagram starts with a very short description of an entity. An entity, in these terms, is any person or thing that can take in, manipulate, or transmit data, information, or knowledge. This definition marks the first major departure from the DFD. As the ultimate goal of this process is to force the participants to categorize data, information, and knowledge flows, the rigor of the normal DFD creation process is unnecessary and detrimental. Entities are drawn as ovals and flows are drawn as lines with arrows.

A useful technique employed to differentiate important examples was to highlight the difference between a pen and a computer. A pen is a classic example of a Heideggerian ready-to-hand tool [15].[1] We should not consider a pen an entity, as it is a transparent component of the participant's expression. However, a computer *should* be considered an entity as it performs transformations outside the scope of the participant's mind. If the participant is having trouble figuring out a scenario they would like to diagram, the entity dictionary section can be expanded into a brainstorming section. They should be prompted to brainstorm all the entities commonly involved in their work: people, machines, and important documents.

From this stage, the interviewer should encourage the participant to define two entities to start out their bubble diagrams. One of these entities should involve one of their job roles. They should not be allowed to define an entity as "myself." Instead, they should label the self-entity according to the type of work that it does. This differentiation is important because people can serve as multiple entities if they play different, discrete, roles in an organization. Once these entities are established, flows should be drawn from the entity representing the person to the other. This flow should be categorized in two ways. Above the line, the content of the flow should be written. Once there, the participant should be asked to categorize the flow as "Data, Information, Knowledge, or Other." They are welcome to use combined terms, but should be encouraged as to explain how something is both "Data and Information" as there are a number of ways that that combination can be interpreted. Participants should always expand the category of other, when used, into a specific category like "Wisdom".

The intent of the bubble diagram portion of the interview is to cause people to develop real definitions of their philosophies of data through categorization. The first entity that they define defines not only a role that they play in an organization, but also hopefully the role through which they will categorize the data flows. My research suggests that participants can have different discrete philosophies of data depending on what role they are visualizing at the time: my first two interviews were conducted against the same person, but produced remarkably

---

[1]   N.B. Never discuss Heidegger in the interview. It does not help.
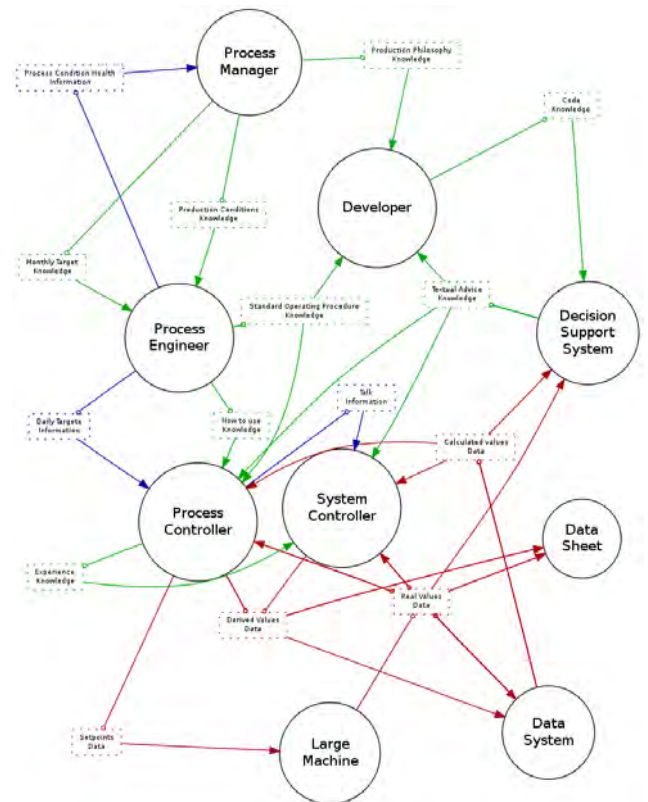

Figure 1.   A sample Social Data Flow Network, anonymized and reproduced with permission from the participant.

different categorization results. In the first interview, we explored the data dynamics of leading a group in a Facebook game. In the second, we explored the participant's work experience programming a decision support system.

With the bubble diagram complete, the true purpose of the interview can finally be revealed: to engage in a theoretical discussion of data. The participant should be encouraged to use the completed bubble diagram to articulate their own ontology of data, information, and knowledge. As most people believe that there is some sort of relationship between the three terms, participants should be encouraged to define all three terms, even though information and knowledge are outside the scope of research. By defining the boundaries and interactions between categories, participants will then usually define what is and is not data.

B.   *Survey Methodology*

The survey, being non-interactive, could not feature the iterative techniques of the formal SDFN. The SDFN is only possible with the interview providing the correct methodology, structure, and scaffolding of the technique. A first attempt at allowing survey participants to "self-scaffold" by reading very long theoretical descriptions of the technique was a miserable failure. In the sixteen attempts to take the survey as part of a pilot test, only two were partially successful. Both of the successful takers relied on prior knowledge of the methodology. This failure suggested that a direct translation of the interview into survey form would not be the most efficient

technique, assuming results were desired. However, the technical survey-presenting software, LimeSurvey, performed very well. As open source, professionally supported software, I can wholeheartedly endorse LimeSurvey for online surveys. Its internal database structure abstracts data at the necessary granularity for real anonymization, not just presentation-level hiding of details.

In rethinking the survey, the first item under consideration was, "what question am I trying to answer?" The failure of the first attempt required the rethinking of objectives, not the simple rearticulation of methodologies. The objective was rephrased as "how do people categorize different data-driven interactions between people?" The interview used the participants' own experiences as these interactions, but the objective does not require that level of specificity. Their own experience was neither sufficient nor required to understand how people categorize data, merely a useful foil in context of the interview. Reflection showed that it was possible to abstract away person-specific descriptions of flows and entities and all of the terminology to look at the core of the question: how does a specific *role* categorize data, information, and knowledge?

The second survey asked people first, to vividly imagine a role they play in their professional lives and to then describe that role in lieu of a traditional demographics section. By self-describing a role, the most important demographic aspect of the survey was captured (how do different roles understand data?) without any of the excess, potentially identifying demographics, that would otherwise be included in normal surveys.

Then, participants were asked to categorize twenty-six very short scenarios, each emphasizing a different flow. Some scenarios used were: "Alice receives a *letter* from Bob.", "Bob receives an e-mail from Alice, it is a *record of the daily temperatures* outside her apartment for the last week." And, "Bob attends a symphony with Alice and enjoys the *live music*."

The participants were asked two questions per scenario. The first was to categorize the scenario, with emphasis on the bolded element, as data, information, knowledge, or other. It is vital, in this instance, to include the other, so that survey-conscious participants do not simply assume a standard hierarchy and try to give "correct" responses. Furthermore, various test-takers used other to explicitly differentiate wisdom, art, or "none of the above."

The second question encouraged participants to: "please, in a sentence or two, explain why you categorized the scenario as you did." This open-ended question allowed participants who wished to engage in self-reflective behavior to do so without otherwise forcing an intimidating cognitive burden on them. As will be discussed in the results, this survey format provided a fantastic venue to confirm the initial results of the interviews.

V. RESULTS

Interviews were manually transcribed from recording devices. Pauses and corrections were edited out, to provide for a cleaner landscape for analysis. As the intent is primarily philosophical instead of linguistic, speaker errors and inadvertent statements are not interesting to the results.

Using recursive analysis, a process designed to iteratively summarize small portions of interviews (to minimize accidental data loss), over 80,000 words of interview transcript was summarized. Table 1 presents the philosophy of data of participants in the interviews and survey. As a check on the methodology, the bubble diagram was analyzed independently of the interview transcript.

Figure 1 shows a bubble diagram rendered in Graphviz as part of the transcription process. Graphviz allows for the

TABLE I. RESULTS FROM THE RECURSIVE ANALYSIS OF INTERVIEWS AND SURVEYS.

Source	Philosophy of Data
SDFN	Data as activity causing communications
SDFN	Data as unprivileged communications
SDFN	Data as facts
SDFN	Data are precise and scientific
SDFN	Data as recorded results of sensors with the world
SDFN	Data as bits
SDFN	No analysis possible
SDFN	Data are numbers about reality
SDFN	Data as unpersuasive ideas
SDFN	Data are numbers about the world
SDFN	Data as objective facts about the world
Interview	Data as instructions
Interview	Data as facts
Interview	Data are factual numbers
Interview	Data is a discrete observation
Interview	Data as bits
Interview	Data are technical numbers
Interview	Data are factual numbers
Interview	Data are numbers about reality
Interview	Data as observations
Interview	Data as measured observations
Interview	Data as contextualized hard numbers
Survey	Data as observations
Survey	Data as factual scientific observations
Survey	Data as electronically stored observations
Survey	Data as records of activity
Survey	Data is objective facts without observation
Survey	Data are specific and precise observations of physical phenomena, and can be stored as bits
Survey	Data are numbers without context
Survey	Data are facts without context and Data has no intrinsic meaning
Survey	Data are observations of the world without context
Survey	Data as symbols without meaning
Survey	Data is an unanalyzed sign, set of signs, or communication
Survey	Data as statements without structure (context)
Survey	Data as structured records
Survey	Data as small, measurable, descriptions of the UoD
Survey	Data as observations without interpretation

"gvpr" language to automatically process graphs. This allows for trivial coloring of the graphs and is far superior to manual layout methods in Visio. I have found that the "neato" and "twopi" layout methods are both necessary, depending on the intricacy of the graph.

VI. ANALYSIS

This research, being primarily exploratory, had two statements of intent designed to guide the abductive process. Furthermore, I intentionally discarded formal "rules" about what data has to be. Participants were allowed to use the term as it came to them naturally, be it singular or plural. If no clear usage was found, the term "as" was used in the summary.

The first, "People have different philosophies of data" was phrased so that analysis would try to detect differences between philosophies. While in the future, with this work as a foundation, it should be possible to create falsifiable statements predicting philosophies of data according to various demographics like educational background and workplace duties, the imposition of strict hypotheses at this time would have unduly biased the results. The statement of interest would serve to focus the analysis in useful directions.

The second statement of interest, "My methodology can probe people's philosophies of data" served in a similar role. Its presence was a *memento mori* of a kind: I was using untested methodologies to explore a new philosophy. The statement of interest required self-reflection onto the efficacy of the methodology and forced me to assess if I was actually capturing real philosophies of data.

Over a set of ten interviews, three distinct philosophies of data arose: data as bits, data as hard numbers, and data as recorded observations. These philosophies seemed to stem mainly from educational background. Despite these different philosophies however, the terms "raw data" and "derived data" were used by members of the group as a strictly defined, though implicit, local language.

The philosophy of "data as bits" seems to be the product of understanding data as something inherently technological. The data as bits philosophy considers that computers are the ultimate arbiters of data: they take in signs as input from human and transform them into data, communicate the signs as data, and reproduce the signs as data. Humans then, through analysis, transform that computer-mediated data into information and knowledge. Books are only frozen communications and do not count as data-transmitting devices. Data is an electronic storage media that contains encoded human produced information and knowledge.

The philosophy of "data as hard numbers" seems to be espoused mainly by scientists. They understand data to be the product of measurements carried out in a scientifically precise manner. Whether the measurements are for experiments or for the process control of manufacturing, they serve as data: objective, reproducible, unambiguous measurements with a precisely understood set of meta-data. The meta-data: provenance, location, time taken, and amount of error, does not count as data. They merely are factors influencing the analysis of that data. Data must be objective, and should stand on its own as a persuasive mechanism.

The philosophy of "data as observations" seems to be an engineering philosophy. This camp understands data as the recorded product of observations. That everything produces data and our knowledge allows us to filter out "irrelevant" data. Important data is then turned into information by contextualizing it against other data through relationships provided by the observer's knowledge. Some of them believe that instead of a hierarchical process, data, information, and knowledge exist in a feedback cycle. Data is subjective, a recorded instantiation of observation.

The surveys produced less conclusive results as a number of responses indicated erroneous understanding of the scenarios or the purpose of the survey. Considering the much wider groups surveyed: intelligence officers, database developers, and more research scientists; their responses do seem to fall into the rough classifications of the interviews. Obviously, far more research and refinement of the survey methodology is required before any positive statements can be made. At this time, however, the results from Table 1 are quite promising.

VII. APPLICATIONS

Both the SDFN and the Philosophy of Data have academic and commercial uses. The SDFN can be a fantastic tool for exploring meaning in domains where theoretical definitions can overwhelm the functional definitions of terms as understood by participants. In Information Technology work, it may be especially valuable as a tool for modeling an organization's current reality, moving from the theoretically objective current procedure of DFD to the "personal reality capture" of the bubble diagram. This capture may allow practitioners to not only see what kind of data flows need to be implemented, but to discover holes in an organization's understanding of itself and group silos, where members of a group communicate only inside the group. Rendering the reality of an organization will make subsequent database design and customization far easier.

Academically, communications and linguistic theory may employ the SDFN technique as an extension of their prototyping theory [16]. While it is inspired from a different background, prototyping seems to be following similar patterns in intent. Categorization may also apply to other approaches in experimental philosophy, offering a novel interview technique to probe people's moral philosophies.

The Philosophy of Data, as a foundation for further research, is useful academically. It not only serves as a useful argumentative basis on the theory of multidisciplinary thought and ventures, but as a way of prototyping synergistic interactions between ventures. It can also serve academically as a basis for consulting services; as it offers an ontologically neutral way of probing a practical aspect of an organization's philosophy. This is useful both in conjunction with other ethnographic research techniques and alone, identifying how an organization thinks about its collective memory (data in databases) and the dysfunctions of communication and ontological interpretation between small groups.

Finally, the philosophy of data will be quite useful to usability engineers, as it allows for a theory behind the "stuff" they are presenting. While most usability engineers are well

versed in communicative and semiotic theory, little attention is paid to epistemological and ontological questions of the substance that they are transmitting.

VIII. CONCLUSION

The philosophy of data seeks to understand how different individuals understand data. This understanding is not part of some grand scheme to present a unified theory of data, but a tool built to allow better technologies and to provide a way for businesses to function more efficiently.

The ideas of silos between different individuals and groups being partially caused by linguistic disconnects in the most fundamental medium of business exchange (data), is not the only conclusion of this research. Better, it should serve as a sample of a potential product: improving efficiency by reducing both epistemological and ontological errors. While the primary business of technologists is in producing technology, I hope this research is a reminder to us all that there are real philosophical bases for anything we create, and that, by assessing how other people construct their understanding of the world, we can make our technologies more useful to the people who will use them.

This research seems to have successfully investigated both statements of interest. Considering statement two, "my methodology can probe people's philosophies of data" the remarkably varied philosophies in Table 1 seem to support the assertion. A poor methodology would most likely produce either low detail or similar results. However, as no statistical analysis is applicable to this analysis, it is impossible to assert that the revelations are statistically significant, merely that they suggest a fascinating field for new research.

The success of statement two and the remarkable diversity of philosophies discovered suggest that statement one has also been satisfied. While there is, of course, the danger of intentionally separating meanings in an attempt to prove statement one, the large subjective-objective gulf between filterable observations and "hard measurements" seems difficult to reconcile with the same philosophy. The qualitative-quantitative gulf also seems suspect if part of one unifying philosophy.

These three philosophies, "data as bits", "data as hard numbers", and "data as recorded observations" are certainly not the only ones that exist, nor did any of my interview participants fall squarely into one camp. While the idea of three competing philosophies of data conflicting between the technical, the scientific, and the engineering ideologies is appealing, we must understand that these are three points on a spectrum of possible philosophies, as our understanding of data is socially constructed from interactions and education.

There is much opportunity for additional research in this area, both testing my results against other research groups and exploring to see if there are any philosophical trends throughout an organization or over time. From these results, the study of the philosophy of data in organizations can also provide value back to those organizations by solving some of the communications problems between small groups.

It is my hope that other research avenues will attempt using the SDFN. It should not be confined to a fledgling field of philosophy and will be useful in any investigation into how people think about things.

REFERENCES

[1] E. Burton Swanson and N. C. Ramiller, "Innovating Mindfully with Information Technology," *MIS Quarterly*, vol. 28, no. 4, pp. 553-583, Dec. 2004.

[2] L. Bryan and C. Joyce, "The 21st-century organization," McKinsey Quarterly, vol. 3, pp. 23–33, 2005.

[3] P. L. Galison, Image and logic: A material culture of microphysics. University of Chicago Press, 1997.

[4] J. Maybin and O. University, Language and literacy in social practice: a reader. Multilingual Matters, 1994.

[5] L. Floridi, "Open problems in the philosophy of information," Metaphilosophy, vol. 35, no. 4, pp. 554–582, 2004.

[6] K. Vredenburg, J. Y. Mao, P. W. Smith, and T. Carey, "A survey of user-centered design practice," in Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, pp. 471–478, 2002.

[7] E. Carmel, R. D. Whitaker, and J. F. George, "PD and joint application design: a transatlantic comparison," Communications of the ACM, vol. 36, no. 6, pp. 40–48, 1993.

[8] R. Carnap, "Testability and meaning," Philosophy of science, vol. 3, no. 4, pp. 419–471, 1936.

[9] Sir K. Popper, The logic of scientific discovery. Hutchinson, 1959.

[10] T. S. Kuhn, The structure of scientific revolutions. University of Chicago press Chicago, 1996.

[11] I. Lakatos, J. Worrall, and G. Currie, The methodology of scientific research programmes. Cambridge Univ Pr, 1980.

[12] P. Feyerabend and P. K. Feyerabend, Against method. Verso Books, 1993.

[13] A. J. Oppel, Databases demystified. McGraw-Hill Professional, 2004.

[14] C. Hogan, Practical facilitation: a toolkit of techniques. Kogan Page Publishers, 2003.

[15] M. Chalmers and A. Galani, "Seamful interweaving," in Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and techniques - DIS '04, p. 243, 2004.

[16] J. R. Taylor and J. R. Taylor, Linguistic categorization. Oxford University Press New York, 2003.

15 References

This chapter contains two sets of references. In-text citations are sorted in text-appearance order and are intended to be used during reading. Works consulted are sorted in alphabetical order and are intended as an easily searchable list to explore my influences.

In the in-text citations, works are numbered in the order they appear in the document, with repeat citations retaining the original number. DOIs are offered instead of hyperlinks where available (though both are clickable in the PDF version of this document) as a hedge against "bitrot" and broken links.

15.1   In-Text Citations

[1]   John Scott. Social network analysis. *Sociology*, 22(1):109–127, 1988. doi:10.1177/0038038588022001007

[2]   Peter Gorm Larsen, Nico Plat and Hans Toetenel. A formal semantics of data flow diagrams. *Formal aspects of Computing*, 3(1), 1994. <citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.57.5990&rep =rep1&type=pdf>

[3]   Sherry X. Sun, J. Leon Zhao, Jay F. Nunamaker and Olivia R. Liu Sheng. Formulating the data-flow perspective for business process management. *Information Systems Research*, 17(4):374–391, 2006. doi:10.1287/isre.1060.0105

[4]   Norbert Wiener. Cybernetics. *Bulletin of the American Academy of Arts and Sciences*, 3(7):2, 1950. doi:10.2307/3822945

[5]   Wolfgang Nejdl, Klaus Tochtermann, Ralf Klamma, Marc Spaniol and Yiwei Cao et al.. *Innovative Approaches for Learning and Knowledge Sharing*, volume 4227 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. doi:10.1007/11876663

[6]   J.A. Barnes. Graph theory in network analysis. *Social Networks*, 5(2):235–244, 1983. doi:10.1016/0378 -8733(83)90026-6

[7]   Christophe Reffay and Thierry Chanier. Social network analysis used for modelling collaboration in distance learning groups. In *Intelligent Tutoring Systems*, pages 31–40, 2002. doi:10.1007/3-540 -47987-2_8

[8]   Colleen Jones. Rediscovering communication. In *UX matters*, 2007. <www.uxmatters.com/mt /archives/2007/08/rediscovering-communication.php>

[9]   Louis R Pondy. Organizational conflict: Concepts and models. *Administrative Science Quarterly*, 12(2):296–320, 1967. <www.jstor.org/stable/2391553>

[10]   Bruno Latour. *Reassembling the social: an introduction to actor-network-theory*. Oxford University Press, Oxford, 2005. <books.google.com/books?id=DlgNiBaYo-YC>

[11]   Janet E. Burge. Knowledge elicitation tool classification. 1998 <http://web.cs.wpi.edu/Research /aidg/KE-Rpt98.html>

[12]   G. Kelly. *The psychology of personal constructs*. New York: Norton, 1955.

[13]   F.B. Tan and M.G. Hunter. The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, pages 39–57, 2002. doi:10.2307/4132340

[14]   P. Cossette and M. Audet. Mapping of an idiosyncratic schema. *Journal of Management Studies*, 29(3):325–347, 1992. doi:10.1111/j.1467-6486.1992.tb00668.x

[15]   M. Jelinek and J.A. Litterer. Toward a cognitive theory of organizations. *Advances in managerial cognition and organizational information processing*, 5:3–41, 1994.

[16]   C. Eden. On the nature of cognitive maps. *Journal of management studies*, 29(3):261–265, 1992. doi:10.1111/j.1467-6486.1992.tb00664.x

[17]   K.E. Weick. *Making sense of the organization*, volume 1. JOHN WILEY PROFESSIO, 2000.

[18]   W.J. Orlikowski and D.C. Gash. Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS)*, 12(2):174–207, 1994. doi:10.1145/196734 .196745

[19] K. Daniels, L. de Chernatony and G. Johnson. Validating a method for mapping managers' mental models of competitive industry structures. *Human Relations*, 48(9):975–991, 1995. doi:10.1177/001872679504800901

[20] Peter L. Berger and Thomas Luckmann. *The social construction of reality: A treatise in the sociology of knowledge.* Penguin Books, Harmondsworth, Middlese, 1967.

[21] H. Markus. Self-schemata and processing information about the self.. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology*, 35(2):63, 1977. doi:10.1037/0022-3514.35.2.63

[22] J.R. Wilson and A. Rutherford. Mental models: Theory and application in human factors. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 31(6):617–634, 1989. doi:10.1177/001872088903100601

[23] M.D. Cohen and P. Bacdayan. Organizational routines are stored as procedural memory: Evidence from a laboratory study. *Organization Science*, pages 554–568, 1994. doi:10.1287/orsc.5.4.554

[24] J. Rasmussen. Cognitive control and human error mechanisms. *New technology and human error*, pages 53–61, 1987.

[25] Russell L. Ackoff, Herbert J. Addison and Sally Bibb. *Management F-Laws.* Triarchy Press, 2007. <books.google.com/books?id=FCr2H97V4YwC>

[26] Russell L. Ackoff. From data to wisdom. *Journal of Applied Systems Analysis*, 16(1):3 – 9, 1989.

[27] Jay H. Bernstein, Elin K. Jacob and Barbara Kwasnik. The data-information-knowledge-wisdom hierarchy and its antithesis. *Journal of Information Science*, pages 68–75, 2009. <arizona.openrepository.com/arizona/handle/10150/105414>

[28] Ilkka Tuomi. Proceedings of the 32nd annual hawaii international conference on systems sciences. 1999. hicss-32. abstracts and cd-rom of full papers. In , pages 12. doi:10.1109/HICSS.1999.772795

[29] Fiona McNab and Torkel Klingberg. Prefrontal cortex and basal ganglia control access to working memory.. *Nature neuroscience*, 11(1):103–7, 2008. doi:10.1038/nn2024

[30] Chaim Zins. Knowledge map of information science. *Journal of the American Society for Information Science and Technology*, 58(4):526–535, 2007a. doi:10.1002/asi.20505

[31] Chaim Zins. Classification schemes of information science: Twenty-eight scholars map the field. *Journal of the American Society for Information Science and Technology*, 58(5):645–672, 2007b. doi:10.1002/asi.20506

[32] Chaim Zins. Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58(4):479–493, 2007c. doi:10.1002/asi.20508

[33] Chaim Zins. Conceptions of information science. *Journal of the American Society for Information Science and Technology*, 58(3):335–350, 2007d. doi:10.1002/asi.20507

[34] P.L. Galison. *Image and logic: A material culture of microphysics.* University of Chicago Press, Chicago, 1997. <books.google.com/books?id=6Gcu92U8CwYC>

[35] Greg Wilson and Carl G. Herndl. Boundary objects as rhetorical exigence: Knowledge mapping and interdisciplinary cooperation at the los alamos national laboratory. *Journal of Business and Technical Communication*, 21(2):129–154, 2007. doi:10.1177/1050651906297164

[36] Valentin Nikolaevich Voloshinov. Language and ideology. In *Language and literacy in social practice: a reader*, pages 271, 1994. Multilingual Matters. <books.google.com/books?id=cCo2Ray4B7kC>

[37]   Danah Boyd. Social steganography: Learning to hide in plain sight. In *apophenia*, 2010. <www
       .zephoria.org/thoughts/archives/2010/08/23/social-steganography-learning-to-hide-in-plain-sight
       .html>

[38]   Valentin Nikolaevich Voloshinov. *Marxism and the Philosophy of Language*. Harvard Univerity Press,
       1929. <www.marxists.org/archive/voloshinov/1929/marxism-language.htm>

[39]   Nicholas Chrisman Chrisman. Trading zones or boundary objects: Understanding incomplete
       translations of technical expertise. In *Social Studies of Science Annual Meeeting*, pages 28–31, 1999.
       doi:10.1.1.24.6239

[40]   Harry Collins, Robert Evans and Mike Gorman. Trading zones and interactional expertise. *Studies
       In History and Philosophy of Science Part A*, 38(4):657–666, 2007. doi:10.1016/j.shpsa.2007.09.003

[41]   Michael E. Gorman. Levels of expertise and trading zones: A framework for multidisciplinary col-
       laboration. *Social Studies of Science*, 32(5-6):933–938, 2002. doi:10.1177/030631270203200511

[42]   Susan Leigh Star and James R. Griesemer. Institutional ecology, `translations' and boundary objects:
       Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. *Social Studies of
       Science*, 19(3):387–420, 1989. doi:10.1177/030631289019003001

[43]   Albert Atkin. Peirce's theory of signs. In Edward N Zalta, editor, *The Stanford Encyclopedia of
       Philosophy*. , Winter 2011 edition, 2011. <plato.stanford.edu/archives/win2010/entries/peirce
       -semiotics/>

[44]   Daniel Chandler. *Semiotics for Beginners: Strengths*. Routledge, London, 1st edition, 2002. <www.aber
       .ac.uk/media/Documents/S4B/sem10.html>

[45]   Douglas L. Medin and William D. Wattenmaker. Concepts and conceptual development: ecological
       and intellectual factors in .... In *Concepts and conceptual development: ecological and intellectual factors
       in ...*, pages 317, 1989. CUP Archive. <books.google.com/books?id=uFY7AAAAIAAJ>

[46]   Kenneth A. Bruffee. Social construction, language, and the authority of knowledge: A bibliograph-
       ical essay. *College English*, 48(8):773, 1986. doi:10.2307/376723

[47]   Deb Roy. A mechanistic model of three facets of meaning. In Vega, Glenberg and Graesser, editors,
       Symbols, Embodiment, and Meaning, 2008. <citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.72
       .1880&rep=rep1&type=pdf>

[48]   John Seely Brown and Richard P Adler. Minds on fire: Open education, the long tail, and learning
       2.0. *Educause review*, 43(1):16–32, 2008. <webpages.csus.edu/~sac43949/PDFs/minds_on_fire.pdf>

[49]   Donald A. Norman. Affordance, conventions, and design. *interactions*, 6(3):38–43, 1999. doi:10.1145
       /301153.301168

[50]   Leon Festinger. *A theory of cognitive dissonance*. Stanford University Press, Stanford, 1957. <books
       .google.com/books?id=voeQ-8CASacC>

[51]   Thomas R. Shultz and Mark R. Lepper. Cognitive dissonance reduction as constraint satisfaction..
       Psychological Review, 103(2):219–240, 1996. doi:10.1037/0033-295X.103.2.219

[52]   Hubert L. Dreyfus. *A companion to Heidegger*. Blackwell, Oxford, 2004. <books.google.com/books?id
       =mhiAcSw201kC>

[53]   C. Addison Stone. The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal
       of Learning Disabilities*, 31(4):344–364, 1998. doi:10.1177/002221949803100404

[54]   Mark R. Leary and June Price Tangney. *Handbook of self and identity*. Guilford Press, New York,
       2005. <books.google.com/books?id=vafgWfgxUK8C>

[55] Geroge Lakoff and Mark Johnson. *Metaphors we live by.* The University of Chicago Press, Chicago, 1981. <books.google.com/books?id=W72CmHdnc3EC>

[56] Anna Dubois and Lars-Erik Gadde. Systematic combining: an abductive approach to case research. *Journal of Business Research*, 55(7):553–560, 2002. doi:10.1016/S0148-2963(00)00195-8

[57] Miri Levin-Rozalis. Searching for the unknowable: A process of detection—abductive research generated by projective techniques. *International Journal of Qualitative ...*, 3(2), 2008. <ejournals.library .ualberta.ca/index.php/IJQM/article/viewArticle/4467>

[58] Dan Sperber and Deirdre Wilson. *Relevance: communication and cognition.* Wiley-Blackwell, 1995. <books .google.com/books?id=2sOKgpYuX4wC>

[59] Terry Pratchett, Ian Stewart and Jack Sidney Cohen. *The science of Discworld.* Ebury, 2002. <books .google.com/books?id=gxZFAAAACAAJ>

[60] Christine L. Borgman. The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of Man-Machine Studies*, 24(1):47–64, 1986. doi:10.1016/S0020-7373(86)80039-6

[61] Margaret Anne Storey, Franco David Fracchia and Hausi A. Müller. Cognitive design elements to support the construction of a mental model during software exploration. *Journal of Systems and Software*, 44(3):171–185, 1999. doi:10.1016/S0164-1212(98)10055-9

[62] Edgar F. Codd. *The relational model for database management: version 2.* Addison-Wesley Longman Publishing Co., Inc., Boston, 1990. <portal.acm.org/citation.cfm?id=SERIES11430.77708>

[63] Susan E Hurley, Theresa M Saunders, Rachna Nivas, Andrew Hertz and Peggy Reynolds. Post office box addresses: a challenge for geographic information system-based studies.. *Epidemiology (Cambridge, Mass.)*, 14(4):386–91, 2003. doi:10.1097/01.EDE.0000073161.66729.89

[64] Donald J. Cunningham. Beyond educational psychology: Steps toward an educational semiotic. *Educational Psychology Review*, 4(2):165–194, 1992. doi:10.1007/BF01322343

[65] Yves Smith. Mers exposed ii: General counsel tells whoppers in testimony before virginia house. In *Naked Capitalism*, 2011. <www.nakedcapitalism.com/2011/01/mers-exposed-ii-general-counsel -tells-whoppers-in-testimony-before-virginia-house.html>

[66] Yair Wand and Ron Weber. Research commentary: Information systems and conceptual modeling?a research agenda. *Information Systems Research*, 13(4):363–376, 2002. doi:10.1287/isre.13.4.363 .69

[67] Robert Frodeman and Jennifer Rowland. De-disciplining the humanities. *ALIF*, 29(Spring):62–74, 2009. <csid.unt.edu/files/Frodeman final - Alif.pdf>

[68] Eliezer Yudkowsky. A day of very low probability. In *Harry Potter and the Methods of Rationality*, 2011. Fanfiction.net. <www.fanfiction.net/s/5782108/1>

15.2 Works Consulted

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis, 16*(1), 3 – 9.

Ackoff, R. L., Addison, H. J. and Bibb, S. (2007). *Management F-Laws*. Triarchy Press.

Alexander, J. and Weinberg, J. M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass, 2*(1), 56–80.

Anonymous (2011). datum, n.. In *Oxford English Dictonary Online* .

Atkin, A. (2011). Peirce's theory of signs. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Winter 2011 edition.

Ballsun-Stanton, B. (2010). Exploring the philosophy of data. Technical Report BlueScope Steel.

Ballsun-Stanton, B. and Bunker, D. (2009). Philosophy of data (pod) and its importance to the discipline of information systems. In *Americas Conference on Information Systems (AMCIS) 2009* .

Barnes, J. (1983). Graph theory in network analysis. *Social Networks, 5*(2), 235–244.

Baudrillard, J. (1994). *Simulacra and Simulation*, volume 50 of *The Body, in theory*. University of Michigan Press.

Berger, P. L. and Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. Harmondsworth, Middlese: Penguin Books.

Bernstein, J. H., Jacob, E. K. and Kwasnik, B. (2009). The data-information-knowledge-wisdom hierarchy and its antithesis. *Journal of Information Science*, pp. 68–75.

Borgman, C. L. (1986). The user's mental model of an information retrieval system: an experiment on a prototype online catalog. *International Journal of Man-Machine Studies, 24*(1), 47–64.

Boulos, M. N. K., Hetherington, L. and Wheeler, S. (2007). Second life: an overview of the potential of 3-d virtual worlds in medical and health education.. *Health information and libraries journal, 24*(4), 233–45.

Boyd, D. (2010). Social steganography: Learning to hide in plain sight. In *apophenia* .

Brown, J. S. and Adler, R. P. (2008). Minds on fire: Open education, the long tail, and learning 2.0. *Educause review, 43*(1), 16–32.

Bruffee, K. A. (1986). Social construction, language, and the authority of knowledge: A bibliographical essay. *College English, 48*(8), 773.

Burge, J. E. (1998). Knowledge elicitation tool classification. .

Carmel, E., Whitaker, R. D. and George, J. F. (1993). Pd and joint application design: a transatlantic comparison. *Communications of the ACM, 36*(6), 40–48.

Chandler, D. (2002). *Semiotics for Beginners: Strengths*. 1st edition London: Routledge.

Chesterton, G. K. (1996). *Heretics*. Online edition Project Gutenberg.

Choi, N., Song, I.-Y. and Han, H. (2006). A survey on ontology mapping. *ACM Sigmod Record, 35*(3).

Chrisman, N. C. (1999). Trading zones or boundary objects: Understanding incomplete translations of technical expertise. In *Social Studies of Science Annual Meeeting*, pages 28–31.

Codd, E. F. (1990). *The relational model for database management: version 2*. Boston: Addison-Wesley Longman Publishing Co., Inc..

Cohen, M. and Bacdayan, P. (1994). Organizational routines are stored as procedural memory: Evidence from a laboratory study. *Organization Science*, pp. 554–568.

Collins, H., Evans, R. and Gorman, M. (2007). Trading zones and interactional expertise. *Studies In History and Philosophy of Science Part A, 38*(4), 657–666.

Cossette, P. and Audet, M. (1992). Mapping of an idiosyncratic schema. *Journal of Management Studies, 29*(3), 325–347.

Cunningham, D. J. (1992). Beyond educational psychology: Steps toward an educational semiotic. *Educational Psychology Review, 4*(2), 165–194.

Czarniawska, B. and Mazza, C. (2003). Consulting as a liminal space. *Human Relations, 56*(3), 267–290.

Daniels, K., de Chernatony, L. and Johnson, G. (1995). Validating a method for mapping managers' mental models of competitive industry structures. *Human Relations, 48*(9), 975–991.

Dreyfus, H. L. (2004). *A companion to Heidegger*. Oxford: Blackwell.

Dubois, A. and Gadde, L.-E. (2002). Systematic combining: an abductive approach to case research. *Journal of Business Research, 55*(7), 553–560.

Eden, C. (1992). On the nature of cognitive maps. *Journal of management studies, 29*(3), 261–265.

Ehn, P. (1993). Scandinavian design: On participation and skill. In Schuler, D. and Namioka, A., editors, *Participatory design: principles and practices*, pages 319. Hillsdale, NJ.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford: Stanford University Press.

Floridi, L. (1995). From data to semantic information. *Entropy, 5*(2), 125–145.

Floridi, L. (2002). On the intrinsic value of information objects and the infosphere. *Ethics and Information Technology, 4*(4), 287–304.

Floridi, L. (2004). Outline of a theory of strongly semantic information. *Minds and Machines, 14*(2), 197–221.

Floridi, L. (2005). Is semantic information meaningful data?. *Philosophy and Phenomenological Research, 70*(2), 351–370.

Frodeman, R. and Rowland, J. (2009). De-disciplining the humanities. *ALIF, 29*(Spring), 62–74.

Galison, P. (1997). *Image and logic: A material culture of microphysics.* Chicago: University of Chicago Press.

Garfield, E. (1980). Is information retrieval in the arts and humanities inherently different from that in science? the effect that isi's citation index for the arts and humanities is expected to have on future scholarship. *The Library Quarterly, 50*(1), 40 – 57.

Gorman, M. E. (2002). Levels of expertise and trading zones: A framework for multidisciplinary collaboration. *Social Studies of Science, 32*(5-6), 933–938.

Gray, R. L. (2003). Brief historical review of the development of the distinction between data and information. In Ross, J. and Galletta, D., editors, *Proceedings of the 9th Americas Conference on Information Systems*, pages 2843–2849. Tampa, Florida.

Greene, J. C. (1987). Stakeholder participation in evaluation design: Is it worth the effort?. *Evaluation and Program Planning, 10*(4), 379–394.

Harper, D. (2011). Data. In *Online Etymology Dictionary* .

Howard, C. (2011). Philosophy and the development of a biologically derived artificial intelligence:an examination of cognitive and neurological methodology. Master's thesis, University of New South Wales.

Hu, W. and Feng, J. (2005). Data and information quality: an information-theoretic perspective. *Computing and Information Systems, 9*(3), 32–47.

Hurley, S. E., Saunders, T. M., Nivas, R., Hertz, A. and Reynolds, P. (2003). Post office box addresses: a challenge for geographic information system-based studies.. *Epidemiology (Cambridge, Mass.), 14*(4), 386–91.

James, W. (2007). *Pragmatism: A New Name for Some Old Ways of Thinking.* Minneapolis: Filiquarian Publishing, LLC..

Jelinek, M. and Litterer, J. (1994). Toward a cognitive theory of organizations. *Advances in managerial cognition and organizational information processing, 5*, 3–41.

Jones, C. (2007). Rediscovering communication. In *UX matters* .

Jones, R. (2006). *Internet Slang Dictionary.* Lulu.com.

Kelly, G. (1955). *The psychology of personal constructs.* New York: Norton.

Kipping, M. and Armbruster, T. (2002). The burden of otheress: Limits of consultancy intervention in historical case studies. In *Management consulting: emergence and dynamics of a knowledge industry*, pages 267. Oxford.

Knobe, J. (2007). Experimental philosophy. *Philosophy Compass, 2*(1), 81–92.

Lakoff, G. and Johnson, M. (1981). *Metaphors we live by.* Chicago: The University of Chicago Press.

Larsen, P. G., Plat, N. and Toetenel, H. (1994). A formal semantics of data flow diagrams. *Formal aspects of Computing, 3*(1).

Latour, B. (2005). *Reassembling the social: an introduction to actor-network-theory*. Oxford: Oxford University Press.

Leary, M. R. and Tangney, J. P. (2005). *Handbook of self and identity.* New York: Guilford Press.

Levin-Rozalis, M. (2008). Searching for the unknowable: A process of detection—abductive research generated by projective techniques. *International Journal of Qualitative ..., 3*(2).

Maier, N. R. F. (1970). *Problem Solving and Creativity in Individuals and Groups.* Brooks/Cole Pub. Co..

Markus, H. (1977). Self-schemata and processing information about the self.. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology, 35*(2), 63.

Marshall, P., Price, S. and Rogers, Y. (2003). Proceeding of the 2003 conference on interaction design and children - idc '03. In , pages 101.

McNab, F. and Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory.. *Nature neuroscience, 11*(1), 103–7.

Medin, D. L. and Wattenmaker, W. D. (1989). Concepts and conceptual development: ecological and intellectual factors in .... In *Concepts and conceptual development: ecological and intellectual factors in ...*, pages 317.

Nejdl, W., Tochtermann, K., Klamma, R., Spaniol, M. and Cao, Y. et al. (2006). *Innovative Approaches for Learning and Knowledge Sharing*, volume 4227 of *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Norman, D. A. (1999). Affordance, conventions, and design. *interactions, 6*(3), 38–43.

Norman, D. A. (2002). *The design of everyday things.* New York: Basic Books New York.

Norman, D. A. (2005). Human-centered design considered harmful. *interactions, 12*(4), 14.

Noy, N. F. (2004). Semantic integration. *ACM SIGMOD Record, 33*(4), 65.

Orlikowski, W. and Gash, D. (1994). Technological frames: making sense of information technology in organizations. *ACM Transactions on Information Systems (TOIS), 12*(2), 174–207.

Orlikowski, W. J. and Robey, D. (1991). Information technology and the structuring of organizations. *Information Systems Research, 2*(2), 143–169.

Owen, W., Landry, J. P. and Mckinney, D. (2000). Introducing information technology students to a new major : The role of an introductory course sequence. In *The Proceedings of the Information Systems Education Conference 2000*, pages §117. Philadelphia.

Pesce, M. (2011). The new toolkit. In *Blackwell Companion to New Media Dynamics* .

Phillips, J. (2011). Experimental philosophy. In .

Pondy, L. R. (1967). Organizational conflict: Concepts and models. *Administrative Science Quarterly, 12*(2), 296–320.

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson & Co.

Pratchett, T., Stewart, I. and Cohen, J. S. (2002). *The science of Discworld.* Ebury.

Rasmussen, J. (1987). Cognitive control and human error mechanisms. *New technology and human error*, pp. 53–61.

Reffay, C. and Chanier, T. (2002). Social network analysis used for modelling collaboration in distance learning groups. In *Intelligent Tutoring Systems*, pages 31–40.

Roy, D. (2008). A mechanistic model of three facets of meaning. In Vega, Glenberg and Graesser, editors, *Symbols, Embodiment, and Meaning* .

Schuler, D. and Namioka, A. (1993). *Participatory design: principles and practices.* Routledge.

Scott, J. (1988). Social network analysis. *Sociology, 22*(1), 109–127.

Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication.* The University of Illinois Press.

Shortis, T. (2001). *The language of ICT: information and communication technology*. Routledge.

Shultz, T. R. and Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction.. *Psychological Review, 103*(2), 219–240.

Smith, Y. (2011). Mers exposed ii: General counsel tells whoppers in testimony before virginia house. In *Naked Capitalism* .

Snow, C. P. (1993). *The two cultures.* Cambridge University Press.

Sosa, E. (2006). Experimental philosophy and philosophical intuition. *Philosophical Studies, 132*(1), 99–107.

Sperber, D. and Wilson, D. (1995). *Relevance: communication and cognition.* Wiley-Blackwell.

Spolsky, J. (2008). my style of servant leadership. In *Inc.* .

Stake, R. E. (1995). *The art of case study research*. SAGE.

Star, S. L. and Griesemer, J. R. (1989). Institutional ecology, `translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. *Social Studies of Science, 19*(3), 387–420.

Stone, C. A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*(4), 344–364.

Storey, M. A., Fracchia, F. D. and Müller, H. A. (1999). Cognitive design elements to support the construction of a mental model during software exploration. *Journal of Systems and Software, 44*(3), 171–185.

Sun, S. X., Zhao, J. L., Nunamaker, J. F. and Sheng, O. R. L. (2006). Formulating the data-flow perspective for business process management. *Information Systems Research, 17*(4), 374–391.

Tan, F. and Hunter, M. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*, pp. 39–57.

Thagard, P. (2004). Computing in the philosophy of science. In Floridi, L., editor, *The Blackwell guide to the philosophy of computing ...*, pages 371.

Tuomi, I. (1999). Proceedings of the 32nd annual hawaii international conference on systems sciences. 1999. hicss-32. abstracts and cd-rom of full papers. In , pages 12.

Voloshinov, V. N. (1929). *Marxism and the Philosophy of Language.* Harvard Univerity Press.

Voloshinov, V. N. (1994). Language and ideology. In *Language and literacy in social practice: a reader*, pages 271.

Voss, J. (2011). Revealing digital documents. concealed structures in data.

Wand, Y. and Weber, R. (2002). Research commentary: Information systems and conceptual modeling?a research agenda. *Information Systems Research, 13*(4), 363–376.

Weick, K. (2000). *Making sense of the organization*, volume 1. JOHN WILEY PROFESSIO.

Wetherbe, J. C. (1991). Executive information requirements: Getting it right. *MIS Quarterly, 15*(1), 51.

Wiener, N. (1950). Cybernetics. *Bulletin of the American Academy of Arts and Sciences, 3*(7), 2.

Wilson, G. and Herndl, C. G. (2007). Boundary objects as rhetorical exigence: Knowledge mapping and interdisciplinary cooperation at the los alamos national laboratory. *Journal of Business and Technical Communication, 21*(2), 129–154.

Wilson, J. and Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 31*(6), 617–634.

Yudkowsky, E. (2011). A day of very low probability. In *Harry Potter and the Methods of Rationality* .

Zins, C. (2007b). Classification schemes of information science: Twenty-eight scholars map the field. *Journal of the American Society for Information Science and Technology, 58*(5), 645–672.

Zins, C. (2007d). Conceptions of information science. *Journal of the American Society for Information Science and Technology, 58*(3), 335–350.

Zins, C. (2007c). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology, 58*(4), 479–493.

Zins, C. (2007a). Knowledge map of information science. *Journal of the American Society for Information Science and Technology, 58*(4), 526–535.