Kierkegaard, digitally wrapped

Karsten Kynde

Since 1997 a new, complete edition of all extant materials from Kierkegaard's hand is being published by the Søren Kierkegaard Research Centre in Copenhagen, under the name of Søren Kierkegaards Skrifter (Søren Kierkegaard's Writings, hereafter abbreviated SKS). From the very start of the project, the aim has been to publish a book as well as a digital version from the same base data, created by the researchers at the Centre. Apart from Kierkegaard's texts, it comprises explanatory notes, text critical accounts, and more, which are in the book version contained in the K-volumes (commentaries).

I The Art of Philological Mark-up

For a text to become published, it must be *marked up*. This is regardless of the publishing medium. In former days, an editor would supply a manuscript with instructions for the typesetter, often in red chalk, on how to further process the text. The instructions would include such things as typeface, font size, margins, pagination, or the placement of notes, in general referred to as mark-up. Today, most texts are conceived electronically, i.e. the mark-up is placed implicitly in some text processing system. In order to make this accessible in a variety of media, and in order to clarify these non verbal signals to ourselves, we should make sure that the mark-up is expressed explicitly.

Consider the following example. The illustration fig. [#1#] shows page 39 of Kierkegaard's journal *DD* from 1837-39.² The first philological task when creating a new edition, is to read and interpret the text. For instance, the heading motto *Nulla dies sine linea*, is written in so-called Latin hand (i.e. the sole line that most of us can actually read) as opposed to the neo-gothic hand that Kierkegaard preferred. In the mark-up this fact may be registered by a set of *tags*, surrounding the text:

```
<ant>Nulla dies sine linea</ant>
```

ant is the name of the tag (in print, a similar phrase would appear in a Roman or Antiqua typeface, hence the name). The text is also appearing larger than the rest of the entry, which is signalled by yet another tag set:

```
<gra str="+2"><ant>Nulla dies sine linea</ant></gra>
```

In this case *str* is the name of an *attribute* giving the size relative to the normal as "+2" (*grade* is the typographical term for font size and *str* is short for Danish *størrelse*, size). *Tags* and *attributes* are concepts of the *Extensible mark-up language* or *Xml* for short.³ As may be seen, the basic idea is

¹ Niels Jørgen Cappelørn, Joakim Garff, Anne Mette Hansen, Jette Knudsen, Johnny Kondrup, Alastair McKinnon, eds., *Søren Kierkegaards Skrifter (SKS)*, Copenhagen: Gad 1997ff.

² Kierkegaard Archive, Royal Library Copenhagen, A pk. 2 læg 3; *Pap.* II A 208-214.

³ Tim Bray, Jean Paoli, C.M. Sperberg-McQueen, Eve Maler, François Yergeau, eds., *Extensible Markup Language* (*Xml*) 1.0 (*Fourth Edition*), W3C Recommendation, 2006 [1998], http://www.w3.org/TR/xml.

to enclose mark-up in angel brackets and leave only the plain transcription of Kierkegaard's text untouched. A pair of corresponding tags starts and ends an Xml *element*; what is in between the tags is called the *content* of the element.

Nowadays, most texts are conceived electronically, and one may ask: Why not simply use an office word processing system such as *Microsoft Word?* Firstly, the result of such a process, the *file* in which the document is stored, is usually written in a proprietary format, i.e., the file format is readable only by that specific word processing program, the documentation is not open for public inspection, but is the property of the vendor, e.g. Microsoft.⁴ Secondly, such systems are often conceived as what is termed *wysiwyg* systems, an acronym for *What You See Is What You Get*. As such, they are almost entirely focused on how the final text appears in print. For instance the above mentioned Latin phrase would following modern English style customarily be expressed in italics. But italics may as well be used to express something quite different in the manuscript, such as emphasis by underlining. We would like to mark different things differently, and to postpone the decision on how to render the particular phenomenon until a later phase of editing. Furthermore we want to be able to express facts that are not immediately visible to the reader. E.g. the phrase mentioned must in *SKS* be explained to the non Latin reader—so the mark-up is supplied with yet a set of tags:

<kom id="dd-395"><qra str="+2"><ant>Nulla dies sine linea</ant>/kom>

meaning that the accompanying commentary volume should contain an explanatory note for this.⁵ The place of the text, the *lemma*, and the particular note are linked by the mutual reference *dd-395*.

Readers acquainted with the internet may recognize the tags as close to those used in Html,⁶ and may ask: Why not use Html? Xml represents a generalization of Html, and without getting deeper into the history of the two languages, one can say that *extensibility* in this context means that the vocabulary has been emancipated and become the choice of the user of the language (specifically the format designer), who has gained the freedom to decide exactly what to call and how to interpret the tags. As we shall see, this also leaves the obligation to fill up the 'semantic gap' between the pure Xml text and the final product which is the target of the edition, e.g. a book or something that can be viewed on a computer screen. an Xml application consists of a vocabulary and a set of *rules*; the format designer should carefully define the tags in order for them to be accepted as legal markup. In fact a formal grammar is constructed for the usage of the elements. In Xml, such usage rules are termed document type definitions (dtds) or schemas. For *SKS*, this has lead to a comprehensive description of the *Kierkegaard normal format* (Kn1).⁷

Now, here we might actually have chosen differently. A 'text encoding initiative' (Tei) have existed since the nineties, providing means for philological tagging of primary texts. Since 2002 its has even been based on Xml. The choice of constructing our own dtd rather than buying a ready-

⁴ Actually, newer versions of *Word* supports an open standard Xml format, *Standard ECMA-376 Office Open Xml File Formats*, December 2006, http://www.ecma-international.org/publications/standards/Ecma-376.htm. This partly invalidates this argument, at least as far as *Word* is concerned.

⁵ The note informs that the Latin text means "no day without a line", KJN, vol. 1, p. 527.

⁶ Dave Raggett, Arnaud Le Hors, Ian Jacobs, eds., *HyperText Markup Language*, HTML 4.01 Specification, W3C Recommendation, 1999, http://www.w3.org/TR/html4/.

⁷ Karsten Kynde, "Kierkegaard Normal Format 1", in *Kierkegaard Studies. Yearbook 2003*, Berlin/New York: Walter de Gruyter 2003. www.sk.ku.dk/kynde/KSYB03.pdf. This version was preceded by a Kierkegaard Normal Format 0, which was not Xml based.

⁸ TEI Consortium, eds., *TEI P5: Guidelines for Electronic Text Encoding and Interchange*, 2007, http://www.tei-c.org/Guidelines/P5/. Proposal P1 was published 1990.

made one is based on a highly pragmatic reasoning. If we should express the explanatory note reference as earlier mentioned in Tei, we may write

```
<xref doc="notes" from="id(dd-395)">Nulla dies sine linea</xref>9
```

Due to its higher generality, the Tei mark-up tends to become a bit more verbose than Kn1. The example assumes that some system or public entity with the name *notes* has been declared elsewhere to mean the commentary volume. Thus *xref doc="notes"* becomes the more general description of an *external reference to a document known as "notes"*, which in Kn1 is shortcut to *kom*. Likewise, the *from* attribute with the prefix *id()* must explicitly express that we are referring to an identifier in that document. The attribute is called *from*, because it is also possible to define an end-point *(to)* of the reference. In favor of choosing a more widespread coding scheme, counts that an experienced Tei reader without any further introduction may get an idea of what is meant here, whereas he in the Kn1 case inevitably would have to consult the manual.

Let us consider an other example, within the area of the philological discipline. Painstaking reading of the manuscript of fig. [#1#], reveals that the second paragraph in the main column (dated *I April*) ends with the words *og beundrer sin egen dybe Bas*. Later on, the author has overwritten the period with a comma and added (with a slightly darker ink): *som om det ikke var de Samtliges Samsyngen men Eens egen Stemme alene, der saa mægtigt fyldte Kirken*. ¹⁰ The *SKS* editors want this fact to appear from a text critical note, reading

Bas, som om (...) Kirken] ændret fra Bas.

ændret fra meaning *changed from*, the *SKS* code for a correction which the author has made on a later stage (which can be told from the darker ink and from the fact that the writing extends into the margin column). The Kn1 encoding is

Explanation: *tn* is the text critical note element, in turn consisting of a subtraction text (*sub*), and the replacing addition text (*add*). The former is typed *aef*, determining the note text (*ændret fra*), as well as the genealogy (late correction). Furthermore, it is specified where to put an ellipsis (...) in the printed note text.

Compare this with the Tei encoding¹¹

```
og beundrer sin egen dybe <app>
```

⁹ The example follows TEI P4. The *xref* is expressed differently in the newest proposal. Alterations in public standards are yet an issue to take into consideration when choosing a coding practice.

¹⁰ and admires his own deep bass. /, as if it were not the singing of all in unison which so powerfully fills the church but his own voice all by itself. (translated by Alastair Hannay, *KJN* 1, p.243, see fig. [#2#]).

¹¹ TEI Guidelines, 11, "Representation of Primary Sources", and 12, "Critical Apparatus".

The encoding may be done in several other ways, by which we just emphasize that the use of a general tool as Tei does not free the editor from a detailed description on how the encoding should be carried out, understood and finally rendered visually. The elements *del* and *add* encourages the transcriptor to classify the text alterations positionally, i.e., according to the placement on the manuscript page. *SKS* classifies alterations according to the genesis of the text, hence it would be reasonable to leave out these two pairs of tags and supply an attribute *type="aef"* to the *rdg* (pronounce *reading*) element. Note the impact of the wrapping on the content: The coding scheme determines the ways a critical note may be composed.

We conclude this section by pointing out one deficiency using Xml. A basic rule of Xml is that to a start tag must belong an end tag, and the tags must be placed parenthetically, i.e.

```
<gra><ant>Nulla dies sine linea</ant></gra>
and not
```

* <gra><ant>Nulla dies sine linea</gra></ant>

This way elements are nested in each other, like Chinese boxes. The document constitutes a tree, the root being the outermost element and the leaves the untagged text. This fundamental property of the Xml language ensures an efficient computer implementation. But is this also the most efficient way for the human reader to conceive a 'text' in general? Consider the concept of a page and the concept of a subsection or a chapter of text. These two can generally not be defined as Xml elements and fit nicely inside each other, usually page break will appear completely unstructured among the sections (creating a long row of well known typographical problems). One solution to this problem is to use *milestone tags*. Milestone elements have no contents and hence mark a point rather than a portion of the text. Thus the Kn1 tag

```
<kor id="39"/>12
```

actually merely an abbreviation of the empty element

```
<kor id="39"></kor>
```

marks the page shift to the manuscript page 39 rather than the page content. In *SKS* this problem becomes even more delicate as some pages of the original manuscript are missing. The text has to be restored from alternative, indirect sources, such as copies or earlier printed editions. This fact obviously has to be marked and eventually conveyed to the reader, but again, we have to accept that

¹² kor indicates a correlation with another text source, by default the page number of the manuscript. Other source references are *SKS* book page numbers and for the journals notably the *Pap*. references, which preceding editors so liberally has entered in the manuscripts (see fig. [#1#]) thereby creating our ultimate reference system.

the resulting sectioning is asynchronous with the with the logical sectioning of the book. Again, we have had to invent a less elegant milestone encoding for Kn1:

```
<lin id="1">
...<altbeg id="p1"/>...
</lin>
<lin id="2">
...
</lin>
<lin id="3">
...<altend/>...
</lin>
</lin>
```

An alternative solution to the milestones would be to close the elements of the secondary structure (the restored pages) together with the primary structure (a paragraph or entry) and re-open the element together with the following primary element(s). This leads to an unpleasant repetition of encoding; but it does make it easier to determine the scope within both structures: If a section of text is selected, containing no actual page break (as the second *lin* above), the page number is not explicitly evident.

The problem of multiple textual hierarchies is addressed by the *Multi Element Coding System* invented by Claus Huitfeldt and used by the Wittgenstein Archives of Bergen.¹³

II Versioning—The Book

So far the digital transcription of an edition remains gray theory as long as we have not dealt with rendering the text to the final reader. In the *SKS* case, it has from the very start been an ambition that the edition should appear in a book as well as a digital version, hence we may talk about the rendering in print as one possible *versioning*. In practice, this work is carried out by external specialists, the type setters, toward whom the Xml document type definition becomes the interface. ¹⁴ Fig. [#2#] shows the English translation of the book version, typeset from at marked-up version similar to the Danish original. ¹⁵

The manuscript page 39 is by *SKS* subdivided into five entries (plus one half, which is printed on the next page) numbered DD:95 thru DD:99. The entry numbering appears in the margin of the print (in bold); the numbers in regular font are page numbers from the Danish *SKS* book; the title and number are showed above the top bar, and so is the year of the dating, *1838*. All of which must obviously appear from the encoding:

¹³ C.M. Sperberg-McQueen & C. Huitfeldt, "Concurrent document hierarchies in MECS and SGML", in *Literary and Linguistic Computing*, 14(1), Oxford: OUP 1999, http://llc.oxfordjournals.org/cgi/content/abstract/14/1/29.

 ¹⁴ The graphic design of *SKS* is carried out by Bent Rohde and typeset by BookPartnerMedia, Copenhagen.
 ¹⁵ Niels Jørgen Cappelørn, Alastair Hannay, David Kangas, Bruce H. Kirmmse, George Pattison, Vanessa Rumble, K. Brian Söderquist, eds., "Journal DD", in *Kierkegaard's Journals and Notebooks (KJN)*, Vol. 1, p. 243, Princeton: PUP 2007.

```
</ant>
</lin>
...
</hs>
</opt>
```

(the small italics of the margin are line numbers).

The authors practice to fold the page and primarily write on the inner column, adding remarks in the margin, is followed—as far as possible—in the print. The marginal note, identified by *SKS* as [a], has no reference mark from the author, so we place it at approximately the same vertical point as in the manuscript. As it appears aligned with the line starting with *hvilken* (*which*), we have put an 'invisible' reference point at this place. *hs* and *ms* here means *main column* and *margin column*, resp.

The space below the bottom line is in the Danish *SKS* reserved for critical notes. In the English translation, this place is instead used for translator's notes, of which there are none for this particular page.

III The Digital Version

The aim of the digital version of *SKS* ¹⁶ has been to re-use all the material produced for the book version, in a way suitable for the computer. We have tried to avoid merely to transform the book to the screen whenever this would hamper the expedient use. This has admittedly not always been possible. First of all, Kierkegaard wrote books, and must be read and appreciated by reading books. The digital version should be used for searching, cross reading, and as an easy way to access a huge authorship, but no one expects the readers to apprehend large portions of text reading from a screen. Also, for the texts authored by the editors, i.e., explanatory notes etc. we must accept the fact that the focus is on the appearance in a book. Partly because the book version was published prior to the digital one, partly because this is the way philologists think. The basic idea of the *SKS* notes is, for instance, to explain a word or a phrase at some point of the text in the specific context of that point. A word may mean something slightly different in another context, and a person mentioned may posses other virtues, relevant elsewhere. This editorial choice appears obvious to the book reader,

¹⁶ Niels Jørgen Cappelørn, Joakim Garff, Anne Mette Hansen, Johnny Kondrup, Karsten Kynde, Kim Ravn, eds., *Søren Kierkegaards Skrifter (SKS)*, digital version 1.3, 2008, http://sks.dk.

but the computer user may wonder on the numerous duplications that become evident when searching in the notes department.¹⁷ Once again, note the subtle impact of the encoding on the contents of both versions.

The web site of *SKS* version 1.3, 2008, *http://sks.dk*, contains 9,807 Html formatted files, corresponding to 29 of Kierkegaard's published works, 49 journals, and companion commentary material such as explanatory notes, critical accounts of the texts, illustrations, calendars, and maps. The original Xml files are transformed into Html for the use on the web and split into smaller units, the journals thus into single entries. Technically, this is done by creating a style sheet for each type of text, e.g., journal text, published work text, explanatory note, critical account, etc. ¹⁸ After the style transformation, the resulting files are cut into appropriate units. This is done offline, i.e. once and for all by a huge batch building the actual web site. This means that the Xml files are still the ultimate base for the digital version, no content editing is done posterior to the transformation.

This also means that this technical detail may be altered in future versions. Xml is designed to be displayed on the net, and modern browsers may do the styling directly from the Xml file. This calls for a solution that splits the files offline, but leaves the styling to be done at read time. Again, the chosen solution is based on highly pragmatic reasoning, really not worth discussing here (such as different browsers that interpret style sheet just a teeny little bit differently).

Fig. [#3#] shows a screen shot of the by now quite well-known journal entry DD:96.¹⁹ For this example, we have checked three of the boxes in the blue bar at the top of the window in order to show the manus page (39), the *Pap*. reference (II A 209-210), and references to the explanatory notes (*). Clicking the asterisks causes a separate window to pop up with the notes. The on and off showing of the page numbers obviously can not be done in the book version, and the *Pap*. references has to be looked up in a concordance in back of the volume. Unique for the digital version is the effect of hovering over the name *Poul Møller*, which shows a normalized version of the person's name. For this purpose we enhance the tagging in the base file

```
<pers norm="M&o-;ller, Poul Martin">Poul M&o-;ller</pers>
```

Names of places are tagged similarly, and so is Bible references of which Kierkegaards religious writings are particularly rich. This is also used to compile an alphabetic index for these topics. The tagging is also carried out in the notes section, with the extremely useful effect that even references that are not given but indirectly by the author (e.g. "her") will appear in the index, insofar a reference is given in the notes (e.g. "Regine Olsen"). Such an index could feasibly have been compiled for the book version as well; it is not, partly because an index for one volume would be of limited use, and a collected index would have to wait for the entire edition to be finished.

The vertical alignment principle of the marginal notes mentioned in connection with the book, is obeyed in the digital version. This is done after numerous experiments and discussions in the editor group. The basic idea is, that Kierkegaard's notes are an inherent part of his writing, and the reader should not be allowed accidentally to skip them. This rules out most of the traditional screen solutions in which some hypertext linking is used to display notes on demand from the reader. The technical solution is to make use of the JavaScript programming facility which is a de facto standard

7

¹⁷ "This is the pattern that for instance the Søren Kierkegaard edition still maintains *even* in its electronic form", Hans Walter Gabler, "Henrik Ibsens skrifter under utgivelse", in *Nytt Norsk Tidsskrift*, 4, 2007, p. 364 (my translation and highlighting).

¹⁸ James Clark, ed., *XSL Transformations (XSLT)*, Version 1.0, W3C Recommendation 16 November 1999, http://www.w3.org/TR/xslt.

¹⁹ The web site is graphically designed by Karen-Margrethe Österlin.

for internet browsers, same facility which implements the on and off showing of the numbers and asterisks. The height within the window of the (possibly invisible) reference point in the text—in fig. [#3#] shown as a pair of empty brackets []—is measured, and the margin note, initially placed at the top of the window, is moved vertically down to fit this height. The interested reader may check this by disallowing JavaScript in the browser and observe the result. The solution comes with the price that the user can not alter the width of the columns, as this would destroy the alignment computation.

Footnotes in the journals are more seldom, i.e. notes that Kierkegaard places at the bottom of the same column. In *SKS*, they are placed at the end of the entry. In the published works, however, they appear frequently, and here the solution is less obvious, see fig. [#4#]. We find a placement at the end of the relevant section as too far away from the actual occurrence, so the note is put directly at the spot. This may, however, leave the impression that the following text starts a new paragraph, which is not (necessarily) the case.

The variety of notes is a salient point of *SKS* in any version. The digital solution to two types of author's notes has been mentioned, and so have the explanatory notes. Fig. [#5#] finally shows the text critical note. By checking the proper box, right square brackets reveals wherever the text has a notable variant. Clicking the bracket highlights the entire lemma together with the note. This arrangement too has been subject to diligent discussions, and again in close in-fight with the book version. The obtained solution is closely connected to the book note, thus obeying the rules of genetic style, yet it appears immediately at the spot so that the reader may easily understand the scope and meaning of the event.

This presentation has been focused on the *SKS* texts as such. Indexes and concordances has been briefly mentioned as well. Obviously a search machine is an indispensable tool too, *SKS* hardly introduces anything new here. An ever increasing field of future extensions for the digital version are the *resource files*. One such is a catalog of Kierkegaard's library compiled at his death, ²⁰ an important tool in the Kierkegaard research, which we are looking forward to implement with cross references to the texts and notes of *SKS*.

Illustrations:

_

²⁰ Karsten Kynde, Kim Ravn, *From the Bibliomanic Nerd to the Resource File*, Paper presented at the Society for Textual Scholarship's 14th Biennial International Interdisciplinary Conference at New York University March 13 -17, 2007, http://www.sk.ku.dk/kynde/NY2007.pdf

Nulla dies fine linea. fullow py the for Einest famile my his lite fright Jan Try fair invoyers i den Litter Evige Rufe in Sele saint the grows figor bod; for low Ali: no soff with the growth will be for in storm of from on free gulle law no wing I felly daw lagor trick i hors 2.1 april. I A210 Ped: Uto Blingen, lighton nood man former Eistend I og men de glammer Dryck og de fils ognyn Manigfat og brundsom find ogne dybr Gost, fam om ert de un de Bameleje Panelyngun nur fine agan Nemum alana, in fa. Ty fat und lille Carl gas Bjortht of all on, at the is night Boulfir, jag aghan at flyn far i fambler in the Porter of gas after at glored mig le. "It folker jag of a far me? and of fages faw. Mit trongramed: frof on. The sund far notworking is its bygan. Men now It is a bygan. Men word is the bygand of at jag med a wahling the most Inoles and mundomobles for an Bit ing far fare gas a liquante mark men fare fare gas a liquante mark ment trops gas. Haar set fortner for en ogle & for son i fans Dorto Reine , sa er set fort son ball. · Dist. -Jugarniel en ern at jour foundant + man for, at the an feeling, not for men flower to the state of the man for the the state of the sta

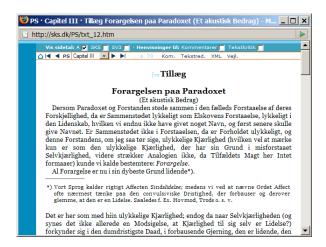
1. Page 39 of Kierkegaard's journal DD. The various arches and strokes are actually mark-up originating from the first editors of Kierkegaard's journals in 1865

252 Nulla dies sine linea. 95 1838. [a] This morning I saw half a score Again such a long time has passed in which I have of wild geese fly away in the crisp been unable to collect myself for the least thing-I must now make another little shot at it. cool air; they were right overhead at first and then farther and farther Poul Møller is dead. away, and at last they separated into two flocks, like two eyebrows over my eyes, which now gazed into the land of poetry. there are some people who make the same mistake in 1 April their relationship to the development of the age as a person singing in church who forgets the organ and the rest of the congregation and admires his own deep bass, as if it were not the singing of all in unison which so powerfully fills the church but his own voice all by itself. 1 April I sat with little Carl on my lap and talked about how much I really looked forward to the old sofa in the new apartment into which I planned on moving. He said, I like it very much too. When I asked him why, he was naturally unable to answer. But isn't it curious how I dwell with a strange sadness on reminders from a time I have never experienced, and now, again, I see him going in the same direction. 25 When the world grows dark for a real Xn in his hour of death, it is because the sunlight of eternal bliss shines too strongly in his eyes.-

2. Page 243 of Kierkegaard's Journals and Notebooks, translated from SKS



3. Journal entry DD:96 of the *SKS* digital version. The top orange line designates that this renders an unpublished manuscript. The same color appears at the back of the book version



4. Section of *Philosophical Fragments*, chapter III, appendix. The top blue line signals together with the text justification that this renders a published work



5. Text critical note of Journal entry DD:97