

UKCOLUMN

Where there's no will, there's no way: Why artificial intelligence will never rule the world

Alex Thomson of UKColumn interviews Jobst Landgrebe¹ and Barry Smith²

March 21, 2023 - Video on [Odysee](#) - Video on [Youtube](#)

Alex Thomson: I'm joined today by [Barry Smith](#) and [Jobst Landgrebe](#) to discuss their fascinating book [Why Machines Will Never Rule the World](#), published by [Routledge](#) in its Philosophy imprint. The subtitle of the book is *Artificial intelligence Without Fear*. Anyone who's watched UK Column News for a while will know that in almost every episode, references are made to the supposedly imminent takeover of one profession or the other by artificial intelligence, and there's certainly a lot about [fear](#) in UK Column News and special broadcasts.

So, with that subtitle as an excellent promise of what's to come—that we can discuss the subjects [without fear](#)—I'm delighted to welcome you both. I will read the blurb from the back cover of the book and then it's over to the both of you to introduce yourselves as you see fit.

Barry, I suppose it might be over to you first, because people listening to that may have wondered what ontology is, for starters—and it does come into the book repeatedly. So perhaps you could explain something of your career. What it is that's taken you to upstate New York, where you're

¹ Jobst Landgrebe is a scientist and entrepreneur with a background in philosophy, mathematics, neuroscience and bioinformatics. Landgrebe is also the founder of Cognotekt, a German artificial intelligence company, which has since 2013 provided working systems used by companies in areas such as insurance claims management, real estate management, and medical billing. After more than ten years in the AI industry, he has developed an exceptional understanding of the limits and potential of AI in the future.

² Barry Smith is one of the most widely cited contemporary philosophers. He has made influential contributions to the foundations of ontology and data science, especially in the biomedical domain. Most recently, his work has led to the creation of ISO/IEC 21838, an international standard in the field of ontology, which is the first example of a piece of philosophy that has been subjected to the ISO standardization process.

speaking to us from today, and something about ontology and data science: how the two relate and what they are good for.

Filleting the scams

Barry Smith: ‘Ontology’ started life as the Latin translation of the Greek word ‘[metaphysics](#)’, and that may give some people some idea of what ontology is. It’s the study of being, in the traditional definition. But in a more modern definition, it’s the study of the *kinds of beings that there are* and of the relations between them. And this study was involved in the very birth of artificial intelligence, as the first attempt to create artificial intelligence in computers consisted in attempts to replicate the ontologies of ordinary human beings.

The idea would be that if we understand how ordinary human beings classify entities in the world, and if we can transmit that information to a robot, then the robot would be able to navigate its way through the world in a way which is similar to how humans do it. Now, with that, ontology became established not just as part of philosophy, but as part of computer science; and it’s been of growing importance, I would say, since around 1970, when these experiments were first made in Stanford.

They all *failed*, incidentally—no-one succeeded in creating a robot on the basis of an ontology—but it was certainly an important stepping stone in the development of artificial intelligence.

The great successes of ontology are not to be found in AI, but rather in biology and medicine. What happened was that at the point in time when the Human Genome Project was beginning to be completed, medicine in particular and the life sciences in general began to realise that they were faced with a gigantic avalanche of new data, new technology, new devices, new kinds of experimental methods, which they knew absolutely nothing about.

These data and methods were based upon a new kind of *chemical* information, and the problem was to find ways of translating this chemical information—which consists of incredibly long strings of letters—into a language which a clinician or a biologist could understand. The key [response] to that was something called the Gene Ontology, which is a collection of terms (nouns and noun phrases) used by biologists to describe biological phenomena, which have now been used to tag sequence data—gene sequence data, protein sequence data, RNA sequence data, and so on—over many years, resulting in an investment of several billions of dollars.

This Gene Ontology thereby serves effectively as the bridge between old biology and clinical medicine on the one hand and the new *chemical* biology which was unleashed by the Genome Project. I was one of the people involved at a crucial stage in the development of the Gene Ontology in turning it into something which is logically coherent.

I’m a philosopher by training and I know something about logic. The people who built the Gene Ontology knew a lot about genes and a lot about genomic data. But they didn’t know very much about logic, and so they built an ontology which was full of logical gaps: logically embarrassing steps and missing items and unclear items. And I showed them, basically, how to do a better job of the logic of the ontology, and that gave me a certain influence in the world of bio-ontology.

That led me to becoming involved in critical work on other biological and medical artefacts which were being created to keep pace with developments in computer science. I was very critical of some of this work: some of it was scams—that is to say, [it was based on] claims about the

computational powers of new medical terminologies which were unsupported in the terminologies themselves.

I was one of the few people who could speak out about these scams. In connection with my work along these lines, I was approached by Jobst—who at that time was actually working for one of these scam organisations; he can justify himself in a minute—and he wrote to me pointing out that he agreed with everything I was saying. And, for a time, he worked as my mole in this world.

I have been working since then in various other kinds of ontology efforts. But I think that's probably enough to give your audience some idea.

Alex Thomson: So are you currently in an academic post? I know that you're in Buffalo, New York.

Barry Smith: Yes, I'm a professor of philosophy, of computer science and engineering, of bioinformatics. and of neurology. But my union card says 'philosopher'.

Alex Thomson: What a fascinating combination. And Jobst Landgrebe, who is going to introduce himself now, has no less a fascinating combination. He's the kind of figure that there is perhaps more of on the European continent than in the English-speaking world, because, as I read, he has qualifications in philosophy, mathematics, neuroscience, but also biology and chemistry.

So he's got all of the *ologies*. He's got fields which have a lot of formulas in them; but also fields that require a lot of writing and thinking on particulars using language.

Also, I think you're properly trilingual, aren't you? Because as well as being a German who uses English for professional life, your *langue maternelle* is French—but you're speaking from your native Germany. So why don't you give us something about your background and what it was that impelled you to study all of these different disciplines?

Jobst Landgrebe: Thanks first of all for inviting us here, Alex. I will try to make it short. So in the end, after I finished *Gymnasium*, which is the equivalent of high school, or grammar school in England, I really didn't know too well what to do. And so I first started philosophy, which I then gave up because I was shocked by the state of university philosophy at the beginning of the 1990s. This was also a time when Barry was starting to think that he should diversify within philosophy. So I wasn't completely wrong with my impression.

So then I switched over to medicine and biochemistry, which I finished in 1998. And then I started a postdoc at the Max Planck Institute of Psychiatry and Neurogenetics and did experiments which yielded so many data points that this brought me into mathematics. So then, I didn't get a degree in mathematics, but I studied it long enough to start publishing a lot of mathematical papers, and that's how I became a biomathematician.

But my interest in philosophy never really died off—and so, when I later got involved in biomedical terminology and ontology, I discovered Barry's writings and then got in contact with him. I also, from the late Nineties, started to work in what was already called artificial intelligence at the time, but I prefer to call it 'statistical learning' or 'machine learning'.

I used it a lot in the biomathematical research that I was doing from the late Nineties, and I've always used it since then in different fields and applied it to different areas. And then, because I

used AI professionally as a *technique*, as a form of *applied mathematics*, I then also started thinking about it, and in the end this led me to write first papers about the topic of artificial intelligence and then later on also our book.

Alex Thomson: So you had a definite writing aim in mind, and the title really brings that out, doesn't it: *Why Machines Will Never Rule The World*. One or other of you will have proposed that title to the other. We've heard from both of you that you dislike scams and [low standards in academia](#). I have recorded already one three-part series and transcribed it [with my father](#) on low standards in academia, and there's another one in the can coming out.

So you're certainly not the only figures who've become disenchanted. When did it come about? Apart from the world of science fiction, which I think we might get into as well because it's far from peripheral here, but [rather] in the world of mainstream hype—commercial hype and academic hype—which I know are symbiotic: they feed off each other.

When did people start popping up saying, "It's not long until machines will rule the world," and—perhaps even more to the point (you hint at this at various points in the book)—*what kind of people* were they who are making this claim? Did they have the rounded view of life and learning that both of you have?

Jobst Landgrebe: I think that claims that machines could become as intelligent as human beings have been made since the Fifties. The [famous paper by Alan Turing](#), where he describes the Turing Test, also discusses this possibility. Turing himself believed that this would come about somewhere in the future—but there was pushback, I think, in each wave of AI.

There were these claims which were made, and they were made more aggressively with each wave. In the first wave, which was the one in which Turing participated, at least at its beginning, the claim was [simply] made. In the second one, it was also made and led to a rebuttal by [Hubert Dreyfus](#), who wrote a book against the possibility of artificial intelligence in the early Seventies.

And then in the third wave, of which we are now contemporaries, it is made in the most aggressive way. I think one of the leaders making the claim has been Kurzweil. Ray Kurzweil was technology director at Google: a really important man in the development of optical character recognition. So he really did good technical things. But *as an engineer*, I think he misses the understanding of what the human mind is and what it can do; and therefore his claims are unfounded.

The mind and the two kinds of systems

Alex Thomson: This segues me nicely to the next question which I'd like to hear both of you talk about: What is the [mind](#)?

I know that's a pretty massive question, but you do dare to broach the topic in the book, and you do so [apophatically](#), by saying what the mind is *not*. You say that the mind, or the brain—which [terms] you use roughly interchangeably—is not a machine, and you then bring out some characteristics which it does have.

So over to you on that: what is the mind? And perhaps you might want to have another pop—I think it's quite legitimate—at the Kurzweilists of this world who say that the mind is reducible to being a mechanical device and thus can be mathematically modelled.

Barry Smith: I'll have a go at this. The whole theme of the book is that there is a distinction between two kinds of systems.

One kind of system is the **mechanical system**, as you [Jobst Landgrebe] used this term, and computers are mechanical systems. Your laptop is a mechanical system; a toaster is a mechanical system. And mechanical systems are built by humans. We understand how they work because *we understand physics*, and we can predict the behaviour of the systems on the basis of what we know about their parts and the way they're put together.

On the other hand, there are **complex systems**—and *all organisms are complex*; systems and the oceans of the world are complex systems; and so on. We can talk a lot about complex systems.

Now, the problem is that the *claim* of the artificial intelligence enthusiasts, shall we say, is that the mind itself is just a mechanical system and therefore sooner or later we will be able to understand how it works in just the same way we understand how a computer or a laptop or a toaster or a car works.

And we argue—on the basis of a quite complicated series of arguments, some of which are grounded in mathematics—that this is not the case: no organism, not even the simplest organism, is ever going to be able to be understood in the way that we understand the workings of mechanisms.

And so this means that *we cannot understand in particular how the brain works*. This is one of the consequences of our general thesis. The brain is—I don't want to say it's a mystery, because that will upset Jobst, since he knows a lot about neurology and neurobiology—but the brain is such that we will never be able to understand it in the way that we understand mechanisms.

Now, what that means is that there are two kinds of intelligence. There's the kind of intelligence that can be achieved by using mechanisms, such as computers—that's *artificial intelligence*. And then there's the kind of intelligence which we might call *general intelligence*, which is the kind of intelligence exhibited by an organism, specifically the human being. And given this gap, which is a necessary gap which will never be eliminated (and that's probably the weakest point in our argument), this means that that can never be *artificial general intelligence*.

Just one other point relating to this which grew out of the discussion of Kurzweil, who was one of the very first people to talk about [the Singularity](#): our book is really addressed to those people who are *worried* that the Singularity is near (I think that was the [title of one of Kurzweil's books](#)).

Alex Thomson: He's been saying that the Singularity will be here by 2030. At least, somebody once gave me a transcription job to do in which he told an audience some years ago it would be here by 2030, just to put that in context.

Barry Smith: It will *never* be here.

Jobst Landgrebe: I think early on, he said it would come earlier—but he keeps postponing it, you know, like the Parousia [Second Coming of Christ], which was [postponed](#) [by enthusiastic predictors] also ...

Barry Smith: The idea behind the Singularity is that once we have a computer which is as intelligent as human beings, that computer will be able to program, devise, somehow assemble even more intelligent computers, in a kind of snowball effect, to bring about gigantically intelligent computers who will take over the galaxy. And that idea, I know from experience of people I talk to, causes real fear in a large group of people. Our book was aimed to be an antidote to that. It's groundless.

And just one other thing: I am actually not so happy with our subtitle, because people assume that the world is going to be "without fear" wherever artificial intelligence is involved. The problem is that there are still going to be *people*, and there are going to be *evil people*, who use artificial intelligence to do things which people might reasonably be afraid of.

And so it's artificial intelligence *in itself* which we should not be afraid of. But once humans start using artificial intelligence—for instance, in scams, but also in creating super-powerful weapons or creating social control mechanisms, then we quite reasonably should be afraid of it. But that means we're afraid of *what the humans are doing with AI*; not with AI itself.

Alex Thomson: This is underappreciated by people who hear, perhaps with reference to their own profession, that some of the work is going to be taken over by AI—or, as you are helpfully directing us to say, 'artificial general intelligence', a more specific term used in the book. People are hearing that AI will take all these roles on, such as claims adjustment, which I read in the book occupies a quarter of a million people in Germany, although that seems to be rapidly dwindling with this computerisation.

Now, in a sentence, in a nutshell, what your book points out is that AI, in order to improve itself, needs a continual dialogue with *people*—just picking up there on what Barry said about people and their intent behind AI. And that forms part of your key argument: the syllogism right there in your Introduction [that] there will never be the Singularity super-duper AI.

Why not? Because it's going to have to be designing a more competent and more amazing successor than itself, and in order to do *that*, it is going to have to have natural conversations with very intelligent human designers, saying, "Now I need to build the better version of me." And it's going to have to instil confidence in the human colleagues—no longer programmers, but colleagues and informers—that it knows what on earth it's doing.

And the last stage of your syllogism is: in order to do that, it would already have to *be* artificial general intelligence; it would already have to be a human mini-me, which can't happen. Have I got the syllogism more or less right?

Barry Smith: That was very beautiful, actually.

Jobst Landgrebe: That's one variant of how it can be put, and this variant rests on the insight that machines cannot use propositional thinking. So when a machine like, now, ChatGPT—which is creating a huge hype again of how 'dangerous' and 'almost already fully intelligent' this AI seems to be—when such an AI is used to create utterances, then it is *not uttering anything*, but it's just producing a sequence of symbols which it doesn't understand.

And this is what many people don't realise. They think, "This is impressive and this sounds almost right"—but they don't realise that this is just a sequence model that that is not thinking at all. It doesn't have any intentions; it doesn't have any self-awareness; it's just a syntactic reckoning

machine like the one about which Ada Lovelace wrote, built by her colleague Charles Babbage in the middle of the nineteenth century.

Of course, the machines now are much more powerful. They can compute a billion times faster than those old machines of 150 years ago. But they are *still based on the same principle*, and so they do not and cannot develop consciousness or think. They are just performing mathematical operations which were defined in the 1930s by Alan Turing and Alonzo Church, and they are just performing combinations of these mathematical procedures.

This is not thinking at all. And so, to call the machine 'intelligent' is a marketing trick used to create hype, but this is just applied mathematics.

Absent the will, no personality

Alex Thomson: Let's thrash out these terms, then, because 'intelligence' is the term that's been marketed now.

Your book is in some ways complementary to one written about forty years ago now, which I've tipped off to Jobst, called *Architect or Bee?* by Mike Cooley—an educated man coming from the labour movement, organised labour in Ireland and then Britain—pointing out that there are many more orders of magnitude involved in human consciousness and cognition than merely intelligence, and we could go into some of the vocabulary.

It varies from author to author, but if we start very classically with the Platonic scheme, a character, a personality, is made up of mind, feelings, and will. And *intelligence is a subset of mind*; it's not the whole of a character. *Feelings* you don't ignore, but you don't particularly mention in the book, either. *Will* is an absolutely core concept in all three parts of your book: the first part on the mind, the middle part on the hard maths of what can be modelled and what are complex systems; and the third part bringing them together: what can maths do to model a mind.

But *will*—not that we really meditate upon this, unless, like you, we look at the philosophers—*will* is absolutely key to having even the most basic conversation, isn't it? Because in a conversational exchange, there is a will to compete, a will to live in harmony, a will to understand each other. This is what's missing when you have a game of chess or a conversation with a computer, isn't it? It doesn't *want* to be your friend; it just does, in a symbolic syntactical sequence, what it's told. So, if you speak to an AI bot in [healthcare](#), it doesn't want to heal you. If you speak to a bot in a [courtroom](#), which I understand according to some reports is happening in China, the bot doesn't want to do justice; it doesn't have feelings or will for justice.

And then there's also, in the Introduction, questions about the terms 'consciousness' and 'cognition'. If I understand correctly, consciousness as a study is what has become, in philosophy, [phenomenology](#). You play a lot of tributes to [Edmund Husserl](#)—not a very well-known philosopher in the English-speaking world, although I think educated viewers will have heard of him—and a couple of Husserl's near-contemporaries, possibly disciples or colleagues: [Max Scheler](#) and [Arnold Gehlen](#).

Jobst, this is your domain; these are German thinkers and they have a happy marriage of being [realists](#)—so they're not [nominalists](#), they don't think that everything's in the mind; they think that what's in the mind *is* connected, even physically, to the real world—but they are [nevertheless] phenomenologists; they're studying consciousness; they don't go off into the long grass of [later](#)

[twentieth-century Franco-German philosophy](#) that says [nothing really exists](#). That might sound irrelevant to some of our viewers, but that's the philosophical glue in this, isn't it?—that we now know, after a lot of study that consciousness, cognition, is a lot more than just the mind.

Jobst Landgrebe: Before I answer, Barry should also answer this, because Barry is a phenomenologist. This is also the foundation, or one of the foundations, of our friendship, because when I first contacted him, I really mentioned phenomenology; and Barry is now doing ontology, but it's very much based on the work of Husserl, and I would like to mention this.

But yes: phenomenology I see as the peak of the development of Western philosophy. I think it is the highest form that philosophy has taken on since it came about in [Pre-Socratic](#) times; and it is still, as one can see in our book, extremely useful as a tool to understand reality, and that is what we do in the book.

So we use Max Scheler and Edmund Husserl as two very important philosophers as a foundation for our work. And the reason they can be used so well is that they are not only realists but they are also able to *provide an explanation for phenomena that cannot be derived from experience*. So if you ask yourself, "What is the mind? What is consciousness? What is intelligence?", it's not sufficient only to look (as positivism proposes) at empirical data.

Phenomenology provides a philosophical framework for how to deal with concepts or entities which are not made of matter, and so—because we need this when we discuss intelligence and consciousness—it is so useful to use this philosophy.

The main philosophical failures of the twentieth century derive either, in positivism, from the inability to understand these immaterial entities, or, in the case of [Heideggerianism](#), from an unwillingness to accept rationalism, realism, [Aristotelian thinking](#). I think phenomenology provides a foundation for a mature and realistic view of the world that is also in harmony with common sense.

And this is also, I think, one of the foundations of our book: that it is really written and thought through in harmony with common sense.

Alex Thomson: And, as a bridge to Barry's own response to that: Barry, you mentioned that *circa* 1970, so half a century ago now, the big drive—perhaps slightly informed by the science fiction of Isaac Asimov and Robert Heinlein and such like—the big drive was to have a robot that, as you very aptly said, could navigate its way around the world, or even just one profession in the world. Have we got there even now? Is there any prospect, given what's just been said by Jobst, of a robot, an artificial general intelligence, navigating its way around the world?

Barry Smith: Well, I think that the key—and you recognise this—to all of this is the will. We can build robots that can navigate their way around Disneyland, for instance, because it's a controlled environment and only a certain limited range of phenomena can be encountered. And that's the key difference between artificial intelligence and artificial *general* intelligence.

We have intelligent machines, but they're intelligent *only in special worlds*, where everything is simple. So this we call 'narrow intelligence'. But we don't have machines that can navigate in the ordinary real world that humans navigate in, where conditions are changing all the time, where strange phenomena can happen, where we're called upon to to make snap decisions in relation to phenomena that we've never encountered before.

And I think that the key here is the will in the following sense: some of the environments in which machines can navigate are created by games: video games, or just chess or Go. And computers can, as we know, can beat humans when playing games like chess and Go. And so people assume that the computers must *want to win*—and therefore that computers can *want*, and therefore computers can want to take over the galaxy: that they could want to win the game of life with real people.

But actually, if you look at the way computers do things like win at chess, they don't do it because they *want to do* anything. They do it because they've been trained, in a certain highly-specialised way, to have what a human would call a reward system. This is what the AI people call Reinforcement Learning. But you can build a reward system which will imitate having a will *only if you can assign a reward computationally*—which means just by doing a certain piece of arithmetic.

Now, you can't assign rewards to contributions to a conversation—and if you don't believe me, just try it: try and set up a reward system which you and your friend or your wife can agree on, and then you'll give each other rewards for each step in the conversation. You will *never* succeed, because a conversation is a realisation of a complex system, namely the people involved.

So the AI people can indeed create something that *looks very much like* a will, but only in those very narrow areas where we have what we have in games: namely, a strict set of rules which allow allows the calculation of rewards so that you can train the machine to get higher and higher rewards by having the machine play *itself* millions of times.

And that's what they do; that's how they create machines that can play chess or Go better than humans. And most of the really impressive—and there *are* many really impressive achievements of computers in recent years—most of the successes are in areas like games, including mathematically equivalent areas of logic games. *All* of the successes are in what we call narrow areas; that is to say, areas where we have something like a logic system or a simple system which we can understand—in other words, something mechanical.

Jobst Landgrebe: The key here is the closed world. 'Closed world' means that the attributes and what is called the phase space, which you can imagine like a multi-dimensional Cartesian coordinate system—that this phase space is somehow predictable and either you have a real game situation, like the games Barry mentioned, in which this is how the games are; or you can also model a complex system in reality in this way, if you only model it *partially*.

So these partial models of complex systems can be very successful at modelling *certain very regular patterns* of complex systems. For example, the traffic pattern in a big town is very regular; there's a lot of traffic in the morning and in the afternoon; and then there are other traffic patterns throughout the week, and there are cycles depending on the season, and cycles depending on bank holidays. And all of this creates a regularity, so that you can actually get an AI to model this very well. So even complex systems can be modelled with AI methods, *if they have a regularity*.

But the problem of complex systems is that they have a lot of *irregularities*, and these make them inaccessible to modelling with AI. And that's why, whenever complex systems interact and create *unexpected outcomes*, which happens in every conversation—or in many, many other human interactions as well, even just movements of crowds—these irregularities happen. And whenever this is the case, the AI fails, and that's why the automation potential of AI can only be applied to very regular events.

And so that's why, for example, a self-driving car cannot drive freely in Disney World. It could drive in an *empty* Disney World, but as soon as people are running around, it will not drive freely. It can then only drive in a certain limited area where everything is controlled; but as soon as this chaotic nature, or irregular nature, of the behaviour of complex systems comes into play, the AI will always fail. And it's not possible to train it [out of that failure]: neither by using types of mathematical logic nor the stochastic algorithms which now dominate AI—which are called deep neural networks, and which have created the big hype.

Alex Thomson: So, in a nutshell, it's the people factor. AI can navigate a world in which there are things; but people and their irritating unpredictable desires and ways of expressing themselves is going to be beyond it. It is always going to say no when an unexpected conversational remark is made. Hence, if you go to a computer to heal you or to judge your case in the law, if it doesn't understand your behaviour from a mathematical model, it's going to tell you that *you* are the problem; you do not compute.

Jobst Landgrebe: It's not only human beings; it's nature in general, because *most* natural phenomena are complex. So it's also animals; the weather; and the whole way our world is structured is a complex system world. Therefore, *the machine can't cope with the real world*, because it can only cope with simple system settings.

And so, if you think of, for example, using a machine as a robotic policeman, this will completely fail, because the machine would just not be able to cope with the real world, because no situation that it encounters is like the situation it was trained for in the laboratory.

So it will just miserably fail, and that's why the fear that machines will be used in this way is wrong. There *are* legitimate fears, though; but this one, for example, isn't [a legitimate fear], because it just doesn't take into consideration that machines *can't be made to act autonomously* in such a setting.

Robocop-out and Ch(e)atGPT

Alex Thomson: But it *is* happening, isn't it? I don't know which of you would like to answer, but UK Column News, just in the last couple of months, has covered multiple jurisdictions across North America—some in East Asia as well—where robocops are being let loose in certain situations. And the lawyers are the men, or the women of course these days, who are writing the algorithms for them; and they're always telling the robocops to err on the side of not getting the police department sued, which isn't very promising as a set of rules of engagement. But it *is* happening. So perhaps we need to have a footnote here on the claim there is no need to fear in the AI era.

Jobst Landgrebe: Barry, let me quickly answer this. What's happening is that they are now using robots for enforcement activities of police. But these are not *autonomous* robots. They are like the toys that we used as children, with remote control. Like remote control toys. So you can actually have a robot that goes into a danger zone with a remote control that has sensors and a camera and can explore it. But it's not acting autonomously, nor can it in any time to come.

So yes: robots *are* being used in law enforcement settings. But *this is not AI*; this is just a sensor on wheels. And this sensor on wheels may also become armed, soon, and be able to shoot or detonate, and all of this isn't nice; but it is not that this robot acts autonomously.

And also, I've not seen that anywhere there is usage of automated AI in any court system or legal system. There are, of course, AI tools in social media surveillance which are being used, which are very primitive. But there's nothing like this in the actual workings of the justice system.

Barry Smith: I think that there are working examples in traffic law, and there is some academic literature which demonstrates that the use of not robots but *computers* in simple traffic law situations is both cheaper than using people and also more often correct than using people, where 'correct' means applying the law to a given traffic situation.

And I think what one needs to say—and this is something that you did in your own AI work, Jobst—is that those systems only work if you have humans in the loop who are there in the ten per cent or whatever it is of cases where the computer is not confident that it's giving the right assessment because it doesn't have the right data or because there's something which confuses it.

And I want to use this as a segue to talking about [ChatGPT](#), which is the current target of hype. It is indeed possible to generate impressive-looking material out of ChatGPT, but I've been trying all day yesterday and all day today to work out what's going on when I enter into ChatGPT simple questions about two people called Barry Smith, both of whom are philosophers. One lives in London; he's called Barry C. Smith, and he also does work which is vaguely phenomenological. The other one lives in Buffalo, and that's me, and I don't have a middle initial.

Now, when e-mail was first introduced—both Barry C. and I are old enough to have been around before e-mail—when e-mail first started, we used very occasionally to get [spam] e-mails “from girls” who thought that they were in love with some kind of compound Barry Smith, which included features of him and features of me. It didn't happen very often, but it did happen.

It doesn't happen any more, because the [spam-generating] e-mail systems are now using AI—in fact, almost certainly in such a way that those kinds of ambiguities happen almost never. But ChatGPT is *still* making exactly the same mistakes. So it thinks I'm from London, and it makes a number of mistakes like that; it thinks that my job is in Leipzig, which it was ten years ago. It *still* thinks that my job is in Leipzig.

Jobst Landgrebe: It doesn't think!

Barry Smith: So I tell it, “You are making a mistake, please correct this information.” And then I ask it exactly the same question again a few seconds later, along with the complaint that it's making a mistake, and it says, “I am very sorry that I made this mistake”—*and then it repeats exactly the same false information.*

Jobst Landgrebe: If I may react to this immediately: this is a typical instance of what has been described for a long time for stochastic learning—that *stochastic systems can't be corrected at will*. They are trained on a huge set of data, which creates patterns that direct them in how they should then create new sequences of symbols, which is what they then print out on the screen, their output. And so they can be retrained, but *there is no guarantee that they will then learn the right things.*

So, when you give a high load of a certain type of language to these models, you can indeed induce a certain learning effect. For example, the chatbot [Tay](#) by Microsoft five or six years ago, when it went online at Twitter, was retrained by the users to utter extremist and sexist language.

It had to be shut off because it was flooded with these utterances by users and then it copied their behaviour.

So this can be induced. But *to teach a stochastic model to answer a certain special question exactly is almost impossible*, especially when the model is very big. So these models are just approximative sequence models. They don't understand anything, and what Barry just told us is a good example of this.

Alex Thomson: And for those who aren't regularly hearing language of 'stochastics', I know David Scott, one of our presenters who's very keen on [economics](#), will say that in some contexts it's just posh talk for 'guesswork'—but you're here talking about it in terms of the hard maths at the centre of your book, aren't you? You're talking about it as referring to the kind of system where the whole world is happening at you and you just have to deal with it without computing in a straight syntactic line as programmed.

Jobst Landgrebe: When I just said 'stochastic system', I mean really that the training type of today's AI systems, the ones that create the big hype, are *trained* using stochastic approaches.

This means that they're basically given a huge set of data from which they learn patterns, or from which patterns are distilled' and these patterns are then applied in novel situations. This works really well *if and only if* the pattern of data that was used for training is the *same* as the pattern that is then encountered later upon the [commencement of] usage of the model.

And so, when that's the case, the models can be wonderful and can be super-successful and can perform better than human beings. But if there's a deviation from the training pattern, then they fail miserably and it's very hard to correct them.

Barry Smith: I want to draw a quite general conclusion in regard to the general thesis of our book, from what we're talking about now. AI can only work with simple systems, but these [AI-compatible] systems can be huge. *The English language is not a simple system*. But you can create a simple system which is a *model* of the English language and which is very very large and powerful. That's how Google Translate works. It turns the different languages into simple systems, and then it can build codes which enable them to be translated between each other.

Now, ChatGPT has created a simple system out of knowledge on the internet, basically—which is certainly *not* simple and which is changing all the time. But it does this by creating a *temporal cut*. It doesn't have information after (I think it's) 2021, and that means it could never replace Google, because Google will often be required to answer questions about what was happening four minutes ago or four hours ago—and ChatGPT, if you ask it for questions about very recent affairs, will say, "I apologise. I only have information up to 2021."

The reason it has to have a cutoff is because otherwise, it's not going to have anything static which it can build a *simple model* of so that it can use computational tools to process in the way that Jobst just described. And that's yet another reason why ChatGPT is going to be making all kinds of mistakes. And it will be making similar kinds of mistakes even when it's [GPT-4](#) which is being used as a basis, which will have many more data.

And I think we should underline what Jobst has just said: many people in the AI world *think* that if we just have more training data, then we will crack ever more aspects of intelligence that we have hitherto not been able to crack. But that's not true. The *size* of the training data is not relevant.

What's relevant is its *representativeness*; it has to be representative of the entire target set—which is open-ended, not a closed world—and that is always impossible when we're dealing with systems which involve organisms like human beings.

Jobst Landgrebe: The core property that makes this so is that the processes that happen in animate systems are **non-ergodic**. That means that they *don't create repeating distributions*. A simple example is the pattern of each single wave that occurs at the English coast. Since England emerged a long time ago, there have been a huge number of waves which arrived at England's coast. But none of them is like another; each one is unique. And so this is a good example of a system that people don't think so much about, which seems rather simple—but it's highly complicated.

And even if you were to make models at a very high resolution of millions and billions of waves, still we couldn't predict what the next wave would look like at the molecular level. This gives you an example of complex systems in *inanimate* nature, and that tells you why mathematical modelling of complex systems is always only an approximation. This approximation can be very powerful; in some contexts, it can also be dangerous. We can talk about usage of AI and warfare if you're interested in this, but it still is based on design and usage by humans.

Machine translation and interpreting: AI can't bridge semantics to syntax

Alex Thomson: We've rather nicely hit upon a couple of the other concepts I wanted to bring in here, one relating to language—artificial translation and artificial interpreting; that is, voice translation—and the other being this key distinction between inanimate and animate, which, in many fields of study, including linguistics, is a key concept, literally meaning *not having a soul* and *having a soul*, respectively. Because Jobst was just sketching out there: when things become animate, there is no predicting the individual components that go into them.

A moment ago, Barry was talking about machine translation—and that's a term that's more common in my profession of translation and interpreting; it's more accurate than to say 'artificial translation'. People who haven't any background in this will have got a sense from the first part of this interview that a machine is a model: for the purposes of this discussion, a model of reality—it is *not* reality.

So, for the whole time since (your book mentions 2014) Google Translate—and its competitors like DeepL by Langenscheidt (being German, it's better of course)—since they came to maturity, just getting on for a decade ago, people have been struck by how well they do in some areas and how awfully in others.

Now, I'm a Bible translator and a literary translator. The things I do would never go [straight] through DeepL or Google Translate—although I have to confess that, like the whole of the profession, I have the dirty secret of using them for the legwork.

I remain responsible for the end product, much as we like to compare ourselves with pilots and surgeons, in that you have to have tens of thousands of hours getting the habit into your muscle memory after you get your qualifications before you're any good to fly solo, right? So a machine can do the simple operation—the flying in clear air, or doing the not-too-challenging parts of the operation, or doing the boring bits of a translation—but it can't do any of the human-intensive bits; you have to be looking over the machine's shoulder.

And so, about ten years ago, I started to get pinged as a freelancer by companies promising to build translation—and, in the end, voice translation, [interpreting systems](#)—that would be good enough to be used at the United Nations. And they all fall over. They're all predicated upon this idea that you, the paid monkeys, are going to fill in syntax cards and in the end the whole English language—we've just heard it's a very complex system—is going to be reduced to syntax, and all these annoying variations that these pesky humans use in their wording is going to be reduced.

So a sentence that starts with 'Although' has to be rephrased—*by a human*—to one that starts 'X is true but Y is not true'. Then, supposedly, the machine has got the full syntactic model. But, your book points out, the machines fall flat at a much earlier state than that. And this goes right back to NSA and GCHQ in the 1950s, using SYSTRAN, which is still available as a machine translator, to process a lot of Russian intercepted material. To get from the semantics—in other words, what the speaker or writer is getting at—to the syntax, isn't going to work.

Let's take a very simple sentence of German, such as I might hear at a conference that I'm interpreting for. Let's say that the speaker says at some point, in a new paragraph:

Es gibt eine Menge Gründe warum das nicht geht.

Supposedly, that's syntactically reducible, so that [the first phrase,] *es gibt*, will be reduced to 'there is' or 'there are'. But that's not how I'm going to *replicate* it, generally. I might actually, like some conference interpreters, use [in my written notes] some of the [logic symbols](#) that are featured in your book. I might use an upside-down capital E (\exists —'there exists'), a logic symbol, in the margin of my notes, to tell me that at this point I have to [predicate](#) something; I have to say it exists.

But when it comes to the end of the speech [as delivered by the original speaker] in German, if I'm working in consecutive mode and I then have to speak English to the audience, I'm going to be reading the room at that point *before* I decide how to interpret it. If I see that it's a bunch of non-native speakers of English, or native speakers who are half asleep and who need a bit of a jolt, I might go into Jack-and-Jill mode of English and say:

There are several reasons why not—

and I would use the *intonation*, as well, that says: "Listen up, guys; this is a new point."

If, [on the other hand,] I'm in full flow and I see that the audience is with me and hanging on the edge of their seat, I might *start with the predicate*, and at the end of the sentence say:

—and there's no end of reasons why that won't be the case.

So there's an infinite number of ways of saying things from German to English, and it's not at the syntactic level—the computable level—that the problem resides. It's at the level of what am I getting at, what are my feelings, and above all: what is my *will* in this conversation, what am I trying to *achieve*?

And so, you're quite confident, the pair of you, that given another decade, we *won't* see more of this [asymptotic](#) shoot upwards? You quote an author in the Introduction as saying that what looks like a curve towards infinity often turns out just to be one of those long [S-shaped curves](#), and it

might flatten out—and you’re fairly convinced philosophically that that’s the way it’s going to be? We’re not going to carry on for a few years and then find that interpreters are completely replaced?

Barry Smith: I gave a talk a couple of weeks ago in São Paulo to a big Brazilian software conference. I gave the talk in English, but there were two simultaneous interpreters, translating my talk into Portuguese. (Most of the audience had headphones.) Both of them wanted to have lunch with me! Basically, they wanted to kiss my feet, because I had shown why they would still have a job in the next five or ten years. And we truly do believe that. So GPT–4 will create something which in some respects is *slightly* better than GPT–3, but it will still have some of the *same problems*.

And I think we need to deal with the distinction between *different sorts of audiences* when we evaluate these phenomena. When we look at GPT, we’re trying to find ways in which it goes wrong; or when we look at Google Translate, we’re trying to find ways in which it goes wrong. But *most* people are happy if it *seems to go right*, and they will be happy over and over again, because they *won’t be looking for the errors*. In fact, [even] I didn’t notice—when I first asked ChatGPT about *myself*—the errors. It was only when I looked more carefully on the next day that I realised that [the bot] was confusing me with another Barry.

Jobst Landgrebe: I would like to add that I think, in machine translation, we are already seeing a saturation. The saturation comes from the problem that we have not only the level of syntax, which machines can only deal with in a limited way; but we also, of course, have the layers of [semantics](#) and [text pragmatics](#). There is also the *context* which a sentence creates for every other situation in which a text occurs, and so on. To interpret this correctly, you need—and you described it really well, Alex—you need *intentionality*.

So your *will*, and your intentionality derived from your will, allow you to understand: “What does the situation mean for me? What does it probably mean for the others?”

You need [intersubjectivity](#); you need to put yourself into the shoes of the other, think ‘Aha!’, what is the relevant part for them, and so on. And if you are a *good* interpreter—and especially also in simultaneous interpretation—you’re able to do this very fast, and the good results mean that you have captured the intentionality both of the speaker and also the probable intentionality of the listener, which will let you interpret in a different way. Even if you interpret the *same speaker* to different listeners, you will give different interpretations and different translations. And *all of this you can’t model mathematically*.

And so therefore, this profession is not endangered at all. Basically, Google Translate [as used sensibly] for translators—and I translate a lot of text as well, and have always done it in my life—I use it now [merely] as a kind of dictionary that can also help to translate *phrases* or to give approximations for phrases and sometimes sentences. But of course, the *real work of the translation* always has to be done by the human who understands what’s going on, who understands the *situation*.

And so this is an area where AI can’t replace human beings.

Where AI works best is when you have a situation that’s completely repetitive, like an assembly line, or also warfare, in which certain patterns of destruction can be repeated and automated. There are now approaches in warfare to use armies of drones that are not as precise as a human being in their destructive work, but that will basically clean out the whole area of tanks.

These systems are now being developed and they will soon be deployed. And they will be very, very effective, and menacing, and terrible. But *they will not be intelligent*; they will just enable a new form of destructive welfare. These things are happening; we have to think about them. They may have to be regulated; we have to cope with them. But it's much better to do this from a point of view of truly understanding what is mathematically happening.

Learning is irreducibly human

Alex Thomson: Shall we, in the final section of this interview, then, talk about the hierarchy of sciences? Because Jobst has just ended there by saying what's *mathematically* possible. In the Introduction, you talk about three levels of impossibility, of which mathematical impossibility is the most interesting to me.

Physical impossibility has to do with the laws of physics of course; and *technical impossibility* has to do with the state of our development as humans.

The third, *mathematical impossibility*—and this only really dawned on me very recently when reading a [history of science](#) written by an objectivist, a devotee of Ayn Rand. I have question marks over her and her school, but be that as it may, this objectivist summary of science—feel free to disagree with the conclusion here, gentlemen—said mathematics itself is very deeply *human*; that the problem with certain philosophical traditions was that it wasn't thought to be human, but it *is*, because it's how *we* relate things in the world to each other.

Being an objectivist, this historian of science [David Harriman] went further, and said that mathematics is how *we* relate *quantities* of things to others in the world. So it's a model in our mind. So: maths—model—mind, getting closer and closer together than the layman might care to think.

In other words, mathematical impossibilities aren't just things that the universe says 'no' to. They're things which won't go because *we*—at least, under God, if you're a believer—*we* are in a sense running the universe. *We* are understanding what works and *we're* winding the machines up and telling them to go. So if something's a *mathematical* impossibility, there is no way around it by waiting for further technical genius developments. Am I right?

Barry Smith: I agree with some of what you were saying there. In fact, we're writing, at the moment, a paper which will be the beginning of a series of papers (if everything works out) which defends the view of mathematics along precisely the lines you just described. Mathematics, like physics, *is* a part of human culture; it develops historically with time. But hand in hand therewith goes the *discovery of necessary laws*; and some of these necessary laws will be *internal* to mathematics, and some of them will be laws pertaining to the *application* of mathematics.

And so, one necessary law which we've been talking about all morning—or all afternoon, in your case—has to do with what computers can do. Computers can only execute programs if those programs require *computations* in the mathematically-defined sense, which are called [Church-Turing](#) computations.

[This is] [computability](#) in the Church-Turing sense, which is very limited to a very small number of very boring, trivial operations—but which, when you have a big enough and a powerful computer, can be applied to bodies of data which have trillions of data points, and so they can achieve great things.

So ChatGPT achieves the great things which it seems to offer by simple—very, very simple—steps, applied to long, long, long vectors of ones and zeros. There's no knowledge, no will, no semantics; there are just very simple steps applied very quickly in a very powerful way to these long strings of ones and zeros.

And it will *always* be thus. There is *never* going to be a computer which is not working like that. That's what Turing held, and that's what everyone will continue to hold.

And that is the weak point in our book, because many people who are visionaries will say, "Oh, mathematics may discover a new way of computing which is not Church-Turing computing. It may be some kind of organic computing or a non-digital computing, analogue computing. And when we have that, then we will have artificial general intelligence."

And that's where we draw the line. That's where we have to say, "Well, maybe you're right ... but we're not holding our breath."

Jobst Landgrebe: I would go a bit further than Barry.

First of all, mathematics is of course a part of our culture. But also, its precondition is the structure of our brain, and so the limitations we have in mathematics are—mathematics is much less limited than Turing computability—it's true: compatibility is a subset of mathematics. Mathematics is broader, and we *could* think of machines that could do more mathematics than today's Turing machines can. This might really evolve.

However, we still have the *limitation of mathematics itself*. Where does this limitation come from? I think it comes from the *human mind*—or the mind-body continuum, as we say in the book: that the human mind is limited structurally *by its biology* to a certain level of complexity. And the laws—the necessary laws that Barry has just alluded to which we have in mathematics and physics—I think these necessary laws are related to, are determined by, the *maximum complexity* that we can mathematically figure out or imagine.

And so, if we look at the most advanced part of physics, like [quantum field theory](#), we very, very clearly see that quantum field theory is limited by our mathematical capabilities. Now, I believe that these are consequences of the structural limitations of the brain that we have, and that evolved in the process of human evolution. And so, yes: this evolution might go on, but we are not to expect *exponential* changes of our mathematical capabilities, and even if we had them, we would still be completely overwhelmed by the number of variables and the complexity of the relationships that occur, for example, in the human mind—or even in the brain of an animal.

So I think that the mathematical limitations are here to stay, and that it's much better as a scientist to select fields of study with these limitations in mind. And that is what all the clever and great physicists of the twentieth century and also twenty-first century, and all the good mathematicians, have done.

They have focused their minds on problems that are accessible to mathematical thinking, and they all know—the great ones all know very well—how limited this is. They have all said it. And those who make us believe that artificial intelligence can be created, and that machines can become more intelligent than human beings, are *not* these mathematicians and physicists.

There's one exception: Stephen Hawking. But Stephen Hawking, I think, never took the time to think artificial intelligence fully through. And he also has a tendency for sensationalism. But other than him, I know of no really great physicist or mathematician who has ever believed in the feasibility of artificial intelligence—because all of them, by their own experience of what they do when they do mathematical models of reality, know about the limitations.

Alex Thomson: Is it perhaps part of the problem that physics and mathematics have become much less experimental—much less *doing in the real world*—for perhaps a century, and have become much more deductive, much more theorising, for which obviously there is a place; but could it be that the induction, the keeping of a whole model of the world in our mind and refining it as we find new facts, has fallen by the wayside?

It was there, obviously, as we built up from classical geometry to algebra, astronomy, physics, [chemistry]: the whole line [of sciences that spawned each other in previous centuries]. You mention in the book that biology is an odd man out in this, because, as you mentioned a moment ago, and in the book, even animals—let alone people—have a mind-body continuum, so you can't model what the body is feeling.

Even animals are placing themselves to some extent in the person of the other animal, predator or prey; and that can't be modelled, you know. But even where it's just inanimate phenomena in the world, the sciences have at this point started navel-gazing with their theories, so that they're unable even to see what they're missing by not inducting more.

Jobst Landgrebe: Before Barry answers this, just one very important remark here. AI as it practised today is highly inductive. The applied artificial intelligence research that has led to ChatGPT and many many other applications is highly inductive, because it's using empirical material to *automatically* create mathematical algorithms or equations. So it *is* highly inductive.

But you are still, I think, pointing at a very important problem, and that is that the reflection of science by theoreticians has become *detached from reality* to a certain extent. In the humanities, this is a huge problem that has been ongoing for quite a long while now, but even in physics itself, there are areas where physicists have detached themselves from experimentation and are claiming that they can produce pure theories of validity. This is a very dangerous trend in physics.

Yet I don't think this explains the hype around artificial intelligence. This hype rather comes from *people*: from practitioners, on the one hand, who don't understand well enough the mathematics of what they're doing; and on the other hand, of course, from entrepreneurs and politicians who want to exploit AI for certain purposes.

Alex Thomson: Are there no worse intentions at play than that, Jobst? I bounce that back to you, because in the previous interview you gave about your book with Jamie Franklin on the *Irreverend* podcast you made no secret of your Christian profession, and you talked about the different intentions there may be here.

I ask because the most famous science fiction scenario in the world, and I know it was influential in Germany as well, was Isaac Asimov's series *Foundation*, and of course the classic three volumes [at the core of it] were written in the 1940s and 1950s—they're part of this California-based exuberance of the post-war era, in which a number of questionable characters are writing science fiction: some go on to found cults or to become drunk or drug-addled maniacs, later including Aldous Huxley, of course, who oversees that scene and approves of it, as it were.

But others are more sober-minded, like Asimov, who has a better intent for mankind, although I have to say he was brought into science fiction by Robert Heinlein, who was a senior member of an out-and-out Satanist group, the [Ordo Templi Orientalis](#), and it was Heinlein who told Asimov after the War, “You’d better write science fiction”—interesting detail.

Asimov lives until the early Nineties, and at the end of his books, the [last sequel](#) he writes before his death, has as his favourite character the robot Daneel Olivaw, who in the mid-twentieth century is doing what Barry told us the robots failed to do, which was to navigate their way around the world.

Poor Daneel Olivaw, who is the real hero—even more than Hari Seldon—there at the very end of the whole series (here’s a spoiler if you haven’t read the series) summons the humans who’ve been trying to unite the galaxy in peace and says [in paraphrase]:

Well, I was behind the drive to artificial intelligence—I, who am myself a robot. I tried to get, through physics and through biology, everyone to will the same will, to get rid of this annoying problem that people have their own wills. I can’t cope with that. I’ve now managed to engineer the galaxy to a point where everyone wants the same thing. I even set up the environmental movements as a robot, so that people would feel an impulsion to want the same things.

Okay, you could say with goodwill that Asimov’s got good intentions here. But [given] that the whole end of this saga—which I think leaves its mark deeply on other sci-fi and on people who come out of that cultural milieu, like Kurzweil—is, “We have to get everyone pointing the same way; we have to have them sharing the same will and feelings.”

In other words, they become a single mind-body. So can we discount the possibility that there are some actively dark evil actors in the field who want AI, or who want to robotise the human brain, to do just what Asimov says his robot wants to do at the end?

Which is: having failed to build a smarter model of himself using circuitry at the very end, [the robot mastermind Olivaw] has to take hold of the most advanced [augmented] *human* in the galaxy he can find and basically steal his [biological] brain in order to go further with the plan. So how much of that is going on right now in AI?

Jobst Landgrebe: I would say, from the intentions perspective, there may be people who have this intention. And this is it, if you allow me to make one religious remark here, Barry: in [Matthew 4](#), the temptations that the devil sets up for Christ are around power and bread, right? Creating bread for everyone, and having one world power that governs the whole world—so that peace can be established and that we can create heaven on earth.

And I think that it is not for nothing that these are the big temptations, because this seems so nice. But we know that we can’t achieve this, and we know that by no means—even in North Korea—can we have everyone intend the same thing.

Even with the cruellest system, the most violence and abuse of power, this cannot be achieved. If there are people who want this [outcome], well, they are bound to fail, and the question is just how much pain will happen on the way to failure.

However, nothing of what you describe from the *Foundation* saga is currently happening, because all of it is so far away from mathematical, physical, and technical feasibility that it's just not happening.

To give you an example: this firm by Elon Musk [[Neuralink](#)] to create brain implant chips has failed so miserably that he is now sorry. He is now getting back to just creating good old [basal ganglia](#)-simulating chips, which have been around for twenty-five or thirty years.

So this is so far away from feasibility that it's just not even—the intent may be dangerous, but the way to *try* to create this [dystopia] will not go via AI but much more traditional ways of exerting power upon human beings.

Barry Smith: I'd like to volunteer a more modest thesis, more modest than what both of you just speculated about. I think that there are *limits to evolution*: there will never be nine-foot-tall humans, because you just can't make the biology work in such a way that those humans would be selected for. This applies not just on Earth; it applies on all conceivable planets. And it applies not just to height; it also applies to intelligence.

So there will never be *considerably* more intelligent people than the ones we have now. There will never be a people who are cleverer than Leibniz and Newton who will evolve and establish societies, neither here nor on other planets, and this has the beautiful consequence that we can explain [Fermi's Paradox](#): the reason why we don't see aliens regularly landing on Earth, flying in spaceships which would go [faster than the speed of light](#), is because they're as stupid as we are.

And all the governments on Earth are just as stupid as we are, so they'll never be able to create these big effects which Alex has nightmares about. It will always be just stumbling through from one bad outcome to another, and we pick up the pieces and move on to the next stumble.

Foundation isn't coming

Alex Thomson: Let's round off, then, with some practical encouragement. We've already consoled those who are dreaming dreams and nightmares.

People in their own line of work may be told, whether they're in management or at the coal face, that AI is replacing you or your boss or your underling or part of your job. It may be framed positively: "it's taking the donkey work off you," whether you're a soldier or a lawyer or a doctor. People may have their qualms about that, perhaps, having heard this; or at least they they'll understand it's not really feasible.

What kind of educated doubts should people sound in their own meetings at work in order to bring a measure of reality back to the conversation, and de-hype? Or perhaps, in the longer term, what kind of more rounded figures should people in their own line of work be seeking to produce so that, perhaps, the next generation of leadership will not be enticed by the hucksterism that's prevailed for so long with regard to AI?

Jobst Landgrebe: I think that there are two.

For the normal workforce person, the best thing is to look at the Industrial Revolution as it has happened in the last 150 years. The Industrial Revolution has, certainly, mechanised a part of human work—and I think most of this was really beneficial, because dangerous and very painful work has been taken on by machines. And new work opportunities have been created. I think that this will continue.

But it won't continue—and this is the second point—at a *speed* that is comparable to the mechanisation of labour in the second half of the nineteenth century. At that time, there was really a very fast mechanisation of human labour—for example, in cloth manufacturing and industry—and this is not happening.

[I say this] because we have the AI technology that is now being developed; we have had it now for twenty-five years now, and for ten or fifteen years now, there has been huge progress with neural networks. But *we haven't seen a big change in productivity*, and so that means that the usability of these technologies is fairly limited. Otherwise, they would now already be being used massively in various industries—and they *are* being used, but their effect on productivity is quite modest.

And so this is the second point: that it's part of historical development, but it's not as extreme and radical as it was thought.

And the other advice one can give is, of course, education, right? Because education always helps to cope with changing environments.

Alex Thomson: Barry, you're involved in education; so your contribution on this is most welcome as the last word.

Barry Smith: For a long time, people were arguing that just as simultaneous interpreters would soon be out of the job because of computers, so ontologists would soon be out of a job because AI can create the ontologies and do a better job than humans.

All I can say is that all my students get jobs immediately, and some of them are earning—straight from PhD—more than I am. Because there is a need for human beings who can build ontologies; and I take comfort from that.

I think if you have a coherent problem to solve in relation to complex systems—and *every* science, every data-gathering effort, every new experimental method, is a complex system—you're going to need humans to work out *how to make use of the results*.

So I don't worry, and I tell my students not to worry.

Alex Thomson: This has been an extremely rewarding conversation about *Why Machines Will Never Rule the World* by Jobst Landgrebe and Barry Smith, published by Routledge in its Philosophy series. But don't be put off by that if you've never read a book knowingly from the philosophy section of a bookshop or library before.

Make this the one, because if you read it chapter by chapter—and there's a guide in the Introduction to which order is best for you, depending on what kind of reader you are—you will get through it, even if you have to skip a couple of equations. So I very much recommend that people read that book, and I hope that we'll be talking to both Jobst and Barry again.

The UK Column is an independent multimedia news website [supported by its members](#).

© 2006-2021 UK Column. For [permission](#) to use any content on its website please [contact us](#).