DANIEL ANDLER

Professor Emeritus at Sorbonne University, member of the Académie des sciences morales et politiques, philosopher

Patrick Nicolet, CEO of Linebreak Ltd., former Group Chief Technology Officer of Capgemini

Without further ado, Daniel, set the scene.

Daniel Andler, Professor Emeritus at Sorbonne University, member of the Académie des sciences morales et politiques, philosopher

I've just published a 400-page book on AI, and people sometimes ask me how I could have written it so fast, considering how recently AI has been in existence — a year or perhaps a bit more. As a matter of fact, AI wasn't born with ChatGPT. The basic idea in rudimentary form has been around for at least 2 centuries if we go back to Charles Babbage and Lady Lovelace, and even longer if we go back to Jacquard and Pascal, Hobbes and Leibniz. In its modern form, AI was launched by Alan Turing in 1950 and was baptized in 1956 during a meeting in Dartmouth, USA.

How can a little history help grasp the present situation? First, it dispels the notion that present-day AI systems came out of the blue, the outcome of a revelation that overnight changed the fate of mankind; rather, it is the outcome of a long and windy process, during which it ran into difficulties that forced it to abandon its initial assumptions and undergo a radical rethinking: instead of taking mental processes to be a kind of logic, it started seeing them as a kind of perception. Instead of trying to mimic the kinds of thought that we entertain consciously, AI aimed for the sort of information that neurons can process, information to which we have no direct access. Instead of trying to directly turn a von Neumann architecture into a thinking machine, it chose to educate neural nets.

Another reason for remembering the birth of AI is the name it chose for itself, which masked an ambiguity—was it aiming for intelligence or something else? From Day 1, there were two projects behind the project. One was to create a computational system that would think like humans and be intelligent in the sense where humans are intelligent. The other was to find ways to automatize the solution to as many kinds of problems as possible, from chess to translation, from pattern recognition to robot navigation... On the face of it, these are distinct goals. Yet the basic insight was that thinking is really nothing but the ability to solve problems. A fully intelligent system would be one that could solve all kinds of problems, and conversely, the more problems a system could solve, the closer it would come to full intelligence. So, AI set out to automatize one problem after the next. It turned out to be more difficult than expected. AI systems could not figure things out from scratch: they needed rich input, too rich to be spoon-fed to von Neumann computers by human programmers. So, they turned to

neural nets that could learn by themselves from examples. After a slow start, neural nets met with smashing success.

But here's the thing: the systems that AI built, whether old-style reasoners or new-wave perceivers, were special-purpose problem solvers, a population of specialized algorithms that did not add up to anything remotely resembling human intelligence. It seemed that one of the two goals that AI had set for itself in the beginning had been dropped. The mainstream of the profession took that as a fact of life and still does: there are enough problems or tasks waiting to be automatized, or to be automatized more efficiently, to keep AI engineers busy.

But the dream of a machine that would be genuinely intelligent, a true thinking machine, one that would possess "artificial general intelligence" or AGI, or again "human-level intelligence" is alive again. The advent of LLMs, of Generative AI, has tipped the balance: the ability to compose on command coherent and often relevant text and images of any kind and on any topic is not only, as everyone was quick to realize, a true game-changer in terms of applications in countless domains; it also makes it more plausible that AGI might be within reach in just a few years.

But this is based on two assumptions that are implausible. The first is that the current victorious trend is bound to continue until the entire repertory of kinds of problems which the human mind can solve has been conquered by AI. The second assumption is that once that happens, human-level intelligence will have been reached.

As for the first, least unplausible assumption, there are two grounds for caution: first, the current spectacular systems are far from perfect, and far from fully understood. They are too fragile a basis for predicting future success. The second problem is that even if the present successes do herald further progress, they don't support the idea that problems of all kinds are within reach; in fact, it is pretty clear that those which are obey some severe constraints. As for the promise that human-level intelligence is within reach, I claim that it is completely idle. I can only offer two arguments today. The first is that the most visible scientific leaders of AI today all agree on the need for some new insight, in the absence of which AI will plateau. AI today may be on the eve of a turning point similar to the neural net revolution, but it doesn't know yet where to turn. The second is the observation that human intelligence is only very partially a matter of problem solving, and I can't see how AI as presently conceived can do anything but solve problems.

These assumptions are not only implausible; they are also potentially harmful. They send the profession on a wild goose chase, that of artificial, fully autonomous thinkers, instead of sticking to what I take to be AI's major calling, which is to provide humankind with powerful, trustworthy auxiliaries that can help us overcome some of the present technical, scientific, social and political challenges, as well as facilitate daily tasks for which help is really needed. And they facilitate a major falsification: making mechanical systems pass off as silicon-based genuine human beings. The irony is that some people worry about the so-called 'existential risk' posed by human-level, and in short order by superhuman-level intelligence. As I see it, this worry is misplaced; what does worry me though is the combination of the unfounded belief that AGI is around the corner with a misplaced priority given to the goal of having AI implanted in as many contexts as possible, for the sake of making use of such a wonderful tool, regardless of the broader consequences.

The central challenge today is to turn AI into a regular engineering discipline, one which produces, in a well-understood fashion, trustworthy artefacts with built-in guardrails against improper use.

Patrick Nicolet

Thank you, Daniel. I will summarize and from what I understand, and we agree on the panel, that what we see today is specific AI and while general AI remains a possibility we do not see a path to it today. We talk about Artificial Intelligence so much today because, as you described, there is a breakthrough. Before we could have inputs into Artificial Intelligence, which were simple and then became more complex, but the output was always simple. With ChatGPT we can now produce complex outputs. With a complex input and complex output you can take text, images, video, sound and produce the same, which we could not do before. That is the breakthrough. It has an impact not only when you play with your kids but also in the enterprise, and as we heard on the panel with Virginie Robert yesterday, it can interfere in democratic processes.