Technological Disputation

Memo by Francisco De Abreu Duarte

Andrea Calderaro (RSC Visiting Fellow, EUI) and Andrea Renda (STG Professor, EUI) discuss the quote "*How Artificial is Artificial Intelligence?*", inspired on the work of Michael Jordan, 'Artificial Intelligence—The Revolution Hasn't Happened Yet'.

This event has been organised by the Technological Change and Society Interdisciplinary Research Cluster

Andrea Calderano began by observing that, although AI is at the centre of the discussion lately, we take too many facts for granted. As a social scientist, with a keen interest in computer science, Andrea conveyed the feeling that both traditional coding languages and more advanced algorithms still aim at achieving the same. They are processes to produce a certain output, based on instructions given by a human. This means that there is more 'old' than new about AI technology. In any event, Andrea C. did acknowledge some fundamental changes in the technological landscape. He highlighted the role of data and the somehow outdated discussion on Big Data, as well as the role of computational capacity as an absolute gamechanger (especially the rise of quantum computers).

He pointed out that the role of a social scientist is to translate complex challenges into more understandable concepts for the broader society and to provide policymakers with solutions to address such challenges. He questioned whether the concept of AI, as general umbrella concept, would really do such a great job in achieving those objectives.

Andrea C. then turned to Michael Jordan's piece applauding the fact that experts on the field are now criticizing some dimensions of AI. By labelling every single reality as AI, he argued, we risk disempowering policymakers by not giving them enough space to focus on what matters. According to Andrea, we should focus on three dimensions:

- The role of computation capacity and quantum computers on geopolitics, which he described as a cybernetic 'atomic bomb';
- The black-box problem of algorithms and algorithmic transparency;
- The need to focus on the ethics of big data.

Andrea C. further defended the idea that social scientists and other disciplines should cooperate. However, he cautioned that such cooperation should be done beyond the general umbrella of AI, instead focusing more on the more practical aspects of the technology.

Andrea Renda started by referring to the timely moment for such a discussion. In about a week time, the EC will release its official proposal on AI technologies. He then alluded to Magritte and his famous painting 'ceci n'est pas une pipe' to contextualize the question of how artificial AI really is. He put it in a simple manner, as observed later by a EUI Professor, by arguing that AI is neither artificial nor intelligent.

According to Andrea R., what truly distinguishes a human from a machine is the capacity to tell reality from non-reality. Professor Renda shared an example which illustrated the topic:

Some years ago, I was engaged in a project at Berkeley which aimed at understanding how fully automated checkout-free stores were, and the potential challenges associated with the absence of human interaction. The idea was to look at those experiences and to understand how automated they truly were. The conclusion was that, although they seemed absolutely automated to the end-user, there was a hidden factor at play. In fact, those technologies were being constantly monitored and updated by manual human labour, thanks to independent contractors mostly located in developing countries such as Madagascar.

This led Andrea R. to conclude that not everything that 'glitters is gold'. In fact, humans play still a fundamental role in accompanying these AI systems, either by choosing which data to feed machines, or through ex-post oversight on their performance. He further remarked that those workers were often under precarious working conditions, with short-term contracts and a job that could become redundant over time, as machines are adequately and extensively trained.

Discussion and Remarks

Attendees to the disputation provided their thoughts on Andrea Calderaro and Andrea Renda.

A EUI researcher opened the discussion by noting that Michael Jordan and our speakers were right to describe AI as a vague catch-all and ML as an important part of it. He further remarked that AI research often consists of many independent communities, such as autonomous vehicles or natural language processing that share the same paradigm (statistical learning/pattern recognition), but which advance independently.

He then focused on such a diversity and remembered that diversity is also a piece of knowledge, some which is being shared in popular courses like the one offered by fast.ai or the Coursera course on Machine Learning offered by Stanford.

He concluded that the widespread of the concept of AI might also be related to the hunger of science funders and business investors for buzzwords. He showed optimism towards the

future, believing that this problem will solve itself as business and science funding move on to new buzzwords. Some of this can be seen on the landing page of fast.ai which claims to want to make neural networks 'uncool' again.

An attendee then followed with a comment on humans in the loop. He argued that perhaps AI was not as artificial as one would assume. On the contrary, there seem to be too many humans in the loop and were those that made the machine work. He pointed to the fact that ILO was already looking at the working conditions of these people. According to him, humans are not being replaced by machines but by other humans, underpaid and working in terrible conditions. He finished by cautioning: 'if you do not see a human in the loop, you just need to see the bigger loop and you will find him'.

A EUI Professor then proceeded to ask how we could transmit these ideas to policy makers. He confirmed that many of the cluster members agreed with this idea of an AI which is not intelligent nor artificial, but which is still dangerous. The difficulties were then to transmit this to the policymakers in charge.

Andrea Renda took the opportunity to clarify some of his previous points. He replied to the last question of the EUI Professor explaining how the European Commission had seen an opportunity to lead the debate on responsible AI. At the time of the drafting of the proposal, Andrea R. explained, the debate was not as crowded as it is today. In any case, the plan was overly ambitious (von der Leyen wanted a binding proposed legislation in the first 100 days of her mandate, when not even 1000 would suffice to solve the matter completely).

Professor Renda then concluded that the debate towards an AI regulation had been rushed a bit, and some choices had been made. An important discussion was whether to focus the regulation on specific techniques (such as machine learning), an option that he found to be far from optimal. This was an attempt to adopt a more pragmatic approach. The problem with the entire debate, he argued, was that it could lead to a vicious cycle. First, there is a growing hype around the topic which influences policymakers; then such urge rushes the debate; ultimately the pieces of legislation contain some problems which generate even more hype.

This, Andrea R. concluded, often leads policymakers to legislate on things they did not fully understand.

A cluster member continued the discussion by playing devil's advocate. Maybe it is good that AI is such a general concept, she argued. It might attract other disciples to join the debate and chip in. If we narrow too much the debate, we risk losing interested people who could add to the debate.

Another EUI Professor agreed with her and defended also a more nuanced position. The Professor added a different perspective to the debate by arguing that AI might be more artificial and intelligent than what everyone was assuming until then. He provided two reasons:

Firstly, it was artificial because AI was in fact an engineering discipline. This meant that AI created things, there were things being built by it. This rendered it artificial.

Secondly, the fact that it is not generally intelligent (as a general AI) does not mean that it is simply not intelligent. Intelligence depends on how we assess it. Depends on whether the ultimate goal is to maximize efficiency or something else.

He then remarked that there had been very good results on narrow/specific AI and that we should be ready for more coordination between AI systems. The Professor concluded in a futuristic note: 'now machines are learning from humans; but maybe in the future they will learn from other machines'.

A participant provided an interesting comparison between AI and bureaucracy. The researcher argued that we should be clearer regarding what we mean by non-intelligence. Were we referring to animal intelligence? The fact is that many human activities do not require fantastic levels of intelligence and we still consider them human such as mechanical or bureaucratic activities. So maybe AI is not generally intelligent, but there are many situations in which human activity is also very instruction-bound and does not require some general level of intelligence.

Andrea Renda replied to him agreeing with his interesting point. He counter argued claiming that, although it is true that humans are part of infrastructures that go largely beyond their individual control, bureaucratic processes are rule-based and have some safeguards. Andrea R. pointed out how transparency in administrative procedures was crucial, something clearly lacking in many AI systems.

He remarked that, when one wants to regulate a certain sector, it is important to foster the debate that the cluster member referred to. The problem is then where to draw the line between what is AI and what is not. Even the definition provided by the High-Level Expert Group on Artificial Intelligence (in which both Andrea Renda and another EUI Professor participate) was quite vague. Even within the technical community, you will see that different sectors have different views on the novelty of AI (e.g. financial services have longed used software and do not see such an incredible revolution with AI).

Andrea R. concluded with a personal story. One time he met a promising Microsoft researcher who told him that he had "fixed fairness". Microsoft people were convinced they had managed

to code it. The story highlighted how different the worlds of science and social sciences were sometimes. You can code many things, but not fairness. It showed that for some disciplines, fairness was about both process and outcomes. For data scientists, fairness is all about procedures.

The Professor followed-up with a question. Might this discussion on whether AI is bright or dumb simply distract us from the real problem? For policymakers, the problem with AI is not philosophical - on whether it is natural or artificial. It is all about use cases. He argued that maybe we could focus on practical problems such as whether we want Facial Recognition AI technology to be deployed, or AI weapons, and leave the theoretical abstractions for later.

He then conveyed that, when discussing these matters on the High-Level Expert Group on Artificial Intelligence, members were always very keen on discussing high-level questions such as whether AI was intelligent or whether human rights principles should apply to it. However, as soon as someone started talking about concrete real-life cases, nobody wanted to jump in to discuss them.

He then concluded on a strong note: the question is to talk about the hard issues, to talk about the practical issues.

Another EUI Professor then questioned whether trying to regulate AI as a general discipline might have been a mistake. There are so many questions for specific sectors which we could focus on that we could have maybe started there.

A PhD researcher followed on this and attempted to craft an explanation. Maybe the European Commission was trying to have another GDPR moment, by crafting a general piece of AI regulation to be followed by sector-specific legislation. He cautioned, however, for the dangers of this approach if poorly coordinated.

Andrea Calderaro then agreed about the advantages of a practical approach to AI. He however refuted the idea that such approach ought to be done under an AI umbrella. According to Andrea C., the extensive debates on big data and ethics apply perfectly to the new technological developments. If we start labelling everything again and start from scratch it will be very distracting, he argued.

Andrea Renda concluded the session responding to the researcher and his point on the GDPR. He argued that sometimes the EU is a victim of its own success. Although the GDPR is indeed a great feat, one cannot forget that such horizontal approaches come with risks, and that enforcement and compliance remain important problems when it comes to the GDPR.

Andrea pointed out that the Commission is deeply convinced of being a pioneer on this matter, by proposing an all-encompassing AI regulation. However, other countries such as the U.S.A or Canada are taking a much more sector-specific approach. The question is then why they have not invested in a general regulation and whether it was a conscious decision.