EDITORIAL

AI: Interrogating questions

Ramon López de Mántaras, Carles Sierra, Pere Almeda



Araya Peralta

It can no longer be denied that Artificial Intelligence is having a growing impact in many areas of human activity. It is helping humans communicate with each other—even beyond linguistic boundaries—, finding relevant information in the vast information resources available on the web, solving challenging problems that go beyond the competence of a single expert, enabling the deployment of autonomous systems, such as self-driving cars or other devices that handle complex interactions with the real world with little or no human intervention, and many other useful things. These applications are perhaps not like the fully autonomous, conscious and intelligent robots that science fiction stories have been predicting, but they are nevertheless important and useful, and most importantly they are real and here today.

The growing impact of AI has triggered a kind of ‘gold rush’: we see new research laboratories springing up, new AI start-up companies, and very significant investments, particularly by big digital tech companies, but also by transportation, manufacturing, financial, and many other industries. Management consulting companies are competing in their predictions on how big the economic impact of AI is going to be and governments are responding with strategic planning to see how their countries can avoid staying behind. Although all of this is good news, it cannot be denied that the application of AI comes with certain risks.

Several initiatives have been taken in recent years to better understand the risks of AI deployment and came up with legal frameworks, codes of conduct, and value-based design methodologies. Examples are the Asilomar principles for beneficial AI, the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, the technology industry consortium 'Partnership on AI' to benefit people and society, or the EU GDPR regulation which includes the right for an explanation. There is also a rapidly growing literature on the risks of AI and how to handle it. For instance, Luc Steels and Ramon Lopez de Mantaras organised in march 2017 a debate in CosmoCaixa (Barcelona). The main outcome of this event was the "Barcelona Declaration for a Proper Development and Usage of AI in Europe". [1]

The *Center for Contemporary Studies*, a think tank of the Government of Catalonia, is committed to analyze the great global challenges that society is facing. Aware of the impact of AI, this special issue of the magazine IDEES addresses different aspects of the debates around AI through a wide array of articles. These articles, therefore, elaborate on the impact of AI on scientific and industrial policies, economy, labor market, smart cities, the role of data—and particularly the importance of open data—, as well as the importance of ethics in AI. It also includes articles reflecting on the future, including the personal view on AI of a science fiction film director raising some futuristic questions about the role of the relations between humans and artificial intelligences. Some topics, such as AI for Good or Governance, have not yet been covered, but will be included in the near future due to their great importance.

Moreover, AI will certainly have a deep impact on our life and could potentially lead to the paradigm shift of our time. Being aware that every big transformation has its consequences on the economic, social and political sphere, the debates around AI technologies are therefore more relevant than ever. Mindful of this complexity, IDEES special issue on AI aims to pitch into the debate and raise some of the AI's main questions: what principles should govern AI? What impact will AI have on citizens and among our societies? What strategies or policies are being pushed by the leading countries? What institutions will have to tackle AI main challenges? In what ways AI will change the nature of our jobs? How AI will affect world geopolitics? Is AI going to reinforce or endanger democracy? These are just some of the questions ahead that open the discussion.

*

> The whole of the articles in this special issue #48 devoted to Artificial Intelligence will be published gradually in the following months. In the spirit of making the driver ideas of this project available to the readers, this editorial note refers to all the contents pertaining to it.

Race for the control of AI

AI is a global phenomenon that has triggered initiatives in many countries. Indeed, different regions in the world have identified AI as one of their strategic focus for the next decade. In this sense, we wanted to include on this issue a description of the AI main strategies, such as the ones in the USA and China. Professor Stone, in his paper provides a detailed analysis with plenty of document references to the major USA events on AI since 2016. After an apparent lack of activity during 2017 and 2018, 2019 has witnessed a new impulse from the US government. Six National AI Institutes will be created to support several strategic lines. The scientific community at different meetings convened since 2016 has identified these lines. The political strategy seems now on track and actively supporting the most dynamic country in the development of AI technology.

In this AI race, we also wanted to look into China, another extremely dynamic country in AI, who recently announced plans to become the leading country in AI. Professors Yi Chang and Chengqi Zhang describe the Science and Technology Innovation 2030 initiative of the Chinese government. It strongly focuses on AI technologies, as well as the New Generation Artificial Intelligence Development Plan, whose goals are to seize the major strategic opportunities for the development of AI, to create Chinese first mover advantage, and to accelerate building China into an innovative country and a leading power of science and technology in the world. It is worth pointing out to the special emphasis that these initiatives put on improving AI education establishing AI Schools in pilot institutions as soon as possible, and increasing the enrollment of Master and doctoral students in AI and related disciplines.

Ecosystems

Regarding the impact of AI on industry, the recent acceleration in the development of AI in the world fueled in part by many young startups. Catalonia is very active in startup creation and that is why we wanted to analyse the Government of Catalonia’s national strategy to coordinate the Catalan AI ecosystem. We believe that this ecosystem is strong and arguably the strongest in southern Europe. Indeed, Professor Karina Gibert addresses the extraordinary potential that Catalonia has to become a key region for AI. The paper starts by pointing out some historical facts about how AI started in Catalonia and its consolidation with the creation of the Catalan AI Association (ACIA). Next, the paper describes some highlights of the Catalan AI ecosystem, in particular the research and education activities as well as the existence of a private sector with companies involved in AI projects. Finally, the professional organizations such as the Official Professional College of Computer Engineers of Catalonia and the Catalan administration initiative to launch a strategic plan on AI.

Indeed, as of January 2019 there are nearly one thousand startups related to Industry 4.0, among them one hundred and sixteen in AI, thirty eight in Robotics and over two hundred in Big Data. In terms of number of startups, Barcelona is the fifth startup hub in Europe and number one in southern Europe. Daniel Villatoro was asked to describe the current Catalan

and Spanish startup landscape in the context of the Spanish and European AI strategies. The variety of applications and services provided by current startups shows how transversal is the impact of AI in our society. Barcelona is clearly identified as the third most attractive European city to create a startup nowadays. In a similar vein, Onn Shehory analyses the situation in Israel. Israel has been seen as the most dynamic nation in the world for the creation of startups (see the book “Start-up Nation” by Dan Senor and Saul Singer). For any parameter you look at, Israel becomes first in a per capita ranking. Professor Shehory describes with some detail the technology behind several disruptive start-up companies, selected from the more than 800 start-up companies in Israel based on AI technology.

Ethical challenges

As mentioned earlier, AI carries undeniable risks. The development of AI technologies that mediate our relationship with the world is generating an intense debate on how ethical are their decisions. AI can potentially be harmful and thus it may be pertinent to establish certain limits to these decisions. Pak-Hang Wong and Judith Simon address the ethical issues raised by autonomous AI and argue that the ethical reasoning of an AI cannot be separated from the analysis of the supporting values. We completely agree. In particular, they discuss the need to align the decisions of such autonomous systems and the social context where those decisions are taken. The consideration of multiple stakeholders in this alignment is key and defines their notion of distributed ethics of AI.

Carme Torras argues that the increasing interaction of people with all sorts of AI-based devices poses important social and ethical challenges. She aligns with the trend that technical university degrees should get open up to the Humanities, so that students become aware of possible sensitive issues they may face in their future careers and learn to reflect on and discuss these matters. Along this line, she advocates for ethics courses for technologists based on science fiction. She gives several examples of existing initiatives including her novel The Vestigial Heart, (MIT Press), which includes an appendix with 24 ethics questions and hints for a discussion on the situations featured in the book. These content is suitable for a course on Ethics in Social Robotics and Artificial Intelligence covering the following topics: how to design the ‘perfect’ assistant; the importance of robot appearance and the simulation of emotions for the acceptance of robots; the role of AI programs in the workplace and in educational environments; the dilemma between automatic decision-making and human freedom and dignity; and civil responsibility related to programming ‘morals’ in robots.

The future of AI is also the cornerstone of many discussions around the world. From an academic point of view, one important issue is the distinction between narrow AI, that is AI systems capable of performing a single task or a small set of highly related tasks, and general human-level AI. All the examples of AI we have nowadays are of the narrow type. General AI is a goal very hard to achieve because, among other things, requires providing common sense knowledge to AI systems and no one knows how to approach this extremely hard problem. In his article, Miquel Casas, citing a number of opinions from well-known figures in the world of technology and business, proposes the possibility of a future General

Artificial Intelligence. Based on several unlikely (in our opinion, extremely unlikely) but not impossible hypotheses, this intelligence could become a superintelligence that would achieve the so-called “Singularity” and therefore make the transhumanist dream come true. The author lists the potential pros and cons of this achievement and states that, in any case, we have the reins of the future in our hands. He also states that in order to prevent any company or country from monopolizing a hypothetical superintelligence, the search for a future general AI should be the endeavour of a project of international co-operation in order to guarantee that a possible superintelligence would benefit all humankind.

Military use of AI

The ethical implications of the concept of autonomy drives the discussion raised by two papers proposing the banning of autonomous lethal weapons from two angles: Toby Walsh from the perspective of an AI researcher and Roser Martínez and Joaquín Rodríguez from a philosophical perspective. Leaving the decision of who can live and who can die to a machine is considered by a large part of AI researchers and philosophers as morally unacceptable. Again, we completely agree with this view. To support the banning, Toby Walsh shows the pitfalls of the main arguments used to defend the use of autonomous lethal weapons. From a legal and philosophical perspective Martínez and Rodríguez describe similar pitfalls of some exaggerated expectations on AI technology used to support the development of autonomous lethal weapons.

Social and democratic impacts

The impact of AI in society is certainly a crucial aspect that requires extensive analysis at all levels. Lorena Jaume-Palasí, in her paper, argues that Artificial intelligence is a new form of infrastructure, one that is immaterial. Automating a process with AI implies building an invisible layer of software to permanently mediate interactions with and among all involved parties of the process. In this way, immaterial infrastructures are being built into sectors where an infrastructural dimension was unthinkable before. The main idea is that, since infrastructure is the architectonic expression of the politics of a society, AI is a technology affecting societies architectonically, and thus on a collective rather than individual level and, therefore, the implementation of AI requires societal thinking. A very interesting point of view indeed.

Along the same lines of societal impact, but from a regulatory point of view, Joana Barbany describes in detail the fundamental features of the Catalan Charter for Digital Rights and Responsibilities. This Charter is the result of a citizen participation process that has culminated with its approval in the Parliament of Catalonia. Artificial intelligence is part of the Charter. The Charter requires AI developments with algorithmic transparency and respect for the ethical principles of our society. The defense of an open society, democratic and respecting the individual freedoms that this Charter represents, is opposed to the power recently granted to the Spanish government on limiting the digital rights of citizens. This power has been established via the ‘digital decree’ approved in November 2019 at the congress of deputies of Spain.

Pompeu Casanovas reflects on the cultural change that technology is bringing about and its impact on ethics and law. He uses some recent examples of unethical social control using data to describe new forms of power. The rapid development of sociotechnical systems where humans and machines interact will necessarily force the adaptation of our normative and legal systems. He gives a few hints on how AI can actually be useful in the process of defining these updated regulatory systems.

Instead of thinking about how the emergence of AI will change society, César Rendueles turns the question around and asks himself what social changes have taken place that have led us to believe that a burst of AI will take place with shocking effects. His paper focuses on diagnosing a certain 'bubble' of artificial intelligence, in particular for the approaches to AI based on Big Data. He claims that the old epistemological debates about strong AI have disappeared from our contemporaries, not because these debates have been resolved but because of a decision to act as if strong.

The economics of AI

The paper by Joan Torrent-Sellens, analyzes the economic dimension of AI from several points of view. The first one considers AI as a technologically connected platform and dwelves on its relations with other technologies. Also, he studies AI from the viewpoint of a general purpose technology. Then, he addresses the impact of AI on productivity and labour, particularly as a source of labour efficiency. Finally, presents the main challenges that the application of AI generates on the theories and the modeling of the economy.

AI in education

The role that AI can play in education is another fundamental aspect with a tremendous impact on society. Richard Tong and Joleen Liang describe how AI can be applied to improve education. In particular on a high impact area: Adaptive Education and its key solution, the adaptive instructional system (AIS). They explain its history, design and basic mechanisms. The article then reflects on how Artificial Intelligence can be applied to the AIS from the architecture, application and computational model perspective.

Data, AI and governance

We believe that AI has to play a crucial role in empowering citizens and in scaffolding a more democratic and participatory society. In this vein, the application of AI in cities is the topic of a paper by Batlle-Montserrat, Delannoy, Kerr and van Cleempur. They voice their own experience in different municipalities across Europe. A number of areas in city management are already benefiting from the application of AI, including transport, education, safety, and citizen participation. We agree with the authors that the application of AI technology is a big challenge for many cities, not only from a technological point of view but also, and perhaps more importantly, from a cultural, legal and ethical perspective. The paper analyses some of those challenges in detail and proposes a number of

recommendations to city officials on how to deploy the technology.

Regarding the impact on Government and administration, Lourdes Muñoz describes in her paper the increasing impact of the open data strategy of governments and institutions. When data is open to public use and scrutiny, governments improve their transparency and citizens are empowered. Spain ranks second in the list of data openness in Europe and, according to this author, Catalonia has one of the most ambitious laws on data openness (Catalan Law on transparency 19/2014). The paper describes a few interesting Catalan projects using public open data and closes with a description of the challenges that Catalan institutions face. Along the same idea of data openness, Núria Espuny describes the strategy of data openness that they are following at the General Directorate of Transparency and Open Data of the Government of the Generalitat of Catalonia. A strategy aimed at making public data available to citizens, companies or organizations to generate social and economic value or to analyze and interpret patterns and trends to solve complex problems.

In his article, Javier Creus questions the superiority of human rational intelligence based on a series of observations regarding human behaviour in aspects such as climate and social emergencies. He argues that if, apparently, we are not superior —in everything— to plants and we are not different —on everything— to machines, then we have to think of intelligence as a system of continuities and complementarities, where all the intelligences are necessary and connected. It will hence allow the emergence of new technological architectures and social institutions that focus on the potential of AI for improving living conditions on the planet and people's living opportunities. These new architectures should be distributed, that is, without a centralized control and with a separation between data and applications based on them. The article ends by advocating the importance of data for the common good and its control by the citizens.

The cultural imaginary

When it comes to speculating about the future, science fiction authors are possibly the main actors. Such speculations have often been a source of challenging ideas. For instance, we think that Kubrick and Clarke screenplay of "2001 A Space Odyssey" actually anticipated quite a few AI achievements, although Arthur Clarke clearly was wrong picturing a computer, HAL, with general human-like intelligence. In fact, there have been several surveys given to AI experts at different AI conferences, asking when general AI would be achieved. The answers reflect a very wide diversity of opinions, ranging from "within the next ten years" to "never." In other words…we cannot really tell. Our own opinion is that we are really, really, far away from Artificial General Intelligence (AGI). We agree with the answer that Oren Etzioni, director of the Allen Institute for AI, gave to the question of when we would have AGI: "Take your estimate, double it, triple it, quadruple it. That's when." In his paper, Kike Maillo, Director of the science fiction movie "Eva", describes his personal experience regarding his own relation with science fiction as well as the relations of AI and science fiction in general. He explains that his interest for science fiction started when he was a High School student. For him the interesting thing was that when trying to design

machines similar to humans the fiction scientists had to first find out what makes us human and he conjectures that this is also the case nowadays for real AI scientists. Twenty years later, all these reflections led him to direct “Eva”. This film tells us about a future time where social machines proliferate. A future with robots intended to keep company and fight loneliness. The main character in the movie wants to design machines that are not always shown as slaves of our will, which is paradoxical because it is precisely obedience what defines any mechanism. This fact raises the main questions addressed in the movie and in his paper: Can machines that only, and always, obey our orders be good companions and be loved? Can we establish powerful relationships with agents that we do not feel as “equal”?

*

Given the public interest in AI and the eagerness of many organisations, both private companies and governmental institutions, to develop applications that affect people in their daily lives, we think that it is very important that society engage in open discussions. Indeed, as a society we need to understand the consequences of the use of artificial intelligence. We also need to debate calmly about the need or not to regulate it. We hope this special issue helps readers to have an informed and critical position that will empower them in order to engage in such open discussions.

REFERÈNCIES

1 — Under the auspices of the Biocat and l’Obra Social la Caixa with support from ICREA, the Institut de Biologia Evolutiva (UPF/CSIC) and the Institut d’Investigació en Intel·ligència Artificial (IIIA-CSIC).



Ramon López de Mántaras

Ramón López de Mántaras is an engineer, physicist and research professor at CSIC (Consejo Superior de Investigaciones Científicas). He also holds a PhD in physics from Université Toulouse III (France), in informatics by the Universitat Politècnica de Barcelona and a *Master of Science* in Computer Science by Berkeley University. He was also a professor at the Computer Sciences Faculty, in the Universitat de Barcelona and at Université Pierre et Marie Curie in Paris. He is considered to be one of the pioneers in the field of artificial intelligence in Spain, and one of the most renowned scientists. He has been the chief editor of *Artificial Intelligence Communications* and an associate editor of the *Artificial Intelligence Journal*, two of the most influential publications in the field of AI. He has been awarded with several prizes, such as the *Premi Ciutat de Barcelona de Recerca* in 1982, the *European Artificial Intelligence Research Award* in 1987 and the *American Association of Artificial Intelligence* (AAAI) *Robert S. Engelmore Award* in 2011. He is currently member of the Institut d'Estudis Catalans and, since 2000, he is a member of the ECCAI (European Association for Artificial Intelligence). He is one of the coordinators of the 48th issue of Revista IDEES on artificial intelligence.



Carles Sierra

Carles Sierra is a professor, researcher and director at the Artificial Intelligence Research Institute (Institut d'Investigació en Intel·ligència Artificial, IIIA) from the National Research Council (Consell Nacional d'Investigació) at CSIC. He is an associate professor at the Western Sydney University and chief editor in the *Journal of Autonomous Agents and Multiagent Systems*. His research focuses on distributed intelligent systems and in agreement technologies, concentrating specifically in the way agents interact and on how to design frames for interaction. He has participated in more than fourty research projects funded by the EU, and has been awarded with the Prize for Research on Autonomous Agents (Premi de Recerca d'Agents Autònoms) by ACM / SIGAI el 2019. He is one of the coordinators of the 48th issue of Revista IDEES on artificial Intelligence.



Pere Almeda

Pere Almeda is the director of the Institut Ramon Llull, a public body founded with the purpose of promoting Catalan cultre and language abroad. Previously, he has been the director of the Centre for Contemporary Studies of the Catalan Government and of the IDEES magazine. Jurist and political scientist, he holds a MA in Political Science and a postgraduate in International Relations and Culture of Peace. He is also an associate professor of Political Science at the University of Barcelona. He has collaborated and worked as advisor in different institutions such as the Catalan Parliament, the European Parliament or the Department of Political and Peacebuilding Affairs at the UN Headquarters. Has served as coordinator of the International Project of Sant Pau and Director of the Think Tank Fundació Catalunya Europa leading the project *Combating inequalities: the great global challenge*.