# ARTIFICIAL INTELLIGENCE, ROBOTICS AND DATA SCIENCE

**Topic Coordinators**
Sara Degli Esposti (IPP-CCHS, CSIC)
Carles Sierra (IIIA, CSIC)

## INTRODUCTION

The world we live in is increasingly interconnected and features smooth interactions between human beings and all sorts of systems and devices which are showing increasing levels of autonomy and intelligence. Thus, Artificial Intelligence (AI), robotics and data science are already part of people's everyday life and are changing people's working structures, relationships, and learning habits. We understand AI as the ability of a computer or robot to perform tasks usually associated with intelligent beings. In the vision expressed in this book, we include classical and modern approaches to AI, the technologies that come from them and, in general, all kinds of artificially intelligent entities and systems.

Living a life which is mediated by smart technologies poses new and unexpected social and ethical challenges. For instance, the controversy around the ethical implications of the concept of autonomy in AI has led many experts to demand the banning of autonomous lethal weapons (aka killer robots). This banning has received the strong opposition of major players such as China and the US. The United Nations plays an active role in this debate, while the EU amongst other players is figuring out how to position itself within this heated controversy; in the meanwhile, the scientific community is leaning towards a human centred development of AI, a shared position amongst recently funded Research Networks on AI. Ensuring the socially acceptable and ethically aligned development of AI is the central priority of the next decade.

National and regional governments, as well as public and private entities, consider AI an area of high strategic importance and R&D priority in the next decade. AI generates a global phenomenon reshaping how organisations and individuals interact. This consideration invites us to reflect on how the collaboration between humans and machines should enfold to ensure the flourishing of humankind.

The societal and economic relevance of AI has already triggered a number of policy initiatives around the world. Overall there is general agreement on aspects related to AI that all governments need to address. Most national governments such as the UK, France, Spain, New Zealand, or Australia, realise the strategic importance of AI for industry R&D and innovation. They all propose strategies that are rather similar: raising awareness and addressing social and ethical risks; acknowledging the essential role of AI in the public sector and in the provision of public services; and responding to the growing need to promote further research, training and education in this domain.

In 2020 the European Commission put forward a European approach to AI envisioning investments for EUR 20 billion per year over the next decade. The *White paper on Artificial Intelligence — A European approach to excellence and trust,* published in February 2020, presents policy options for creating an ecosystem of excellence and trust, namely to enable a trustworthy and secure development of AI in Europe, in full respect of the values and rights of EU citizens. Similarly, the *Spanish Strategy for AI* identifies six priorities: creating organisational structures for AI development and assessment; identifying AI application areas with high societal and economic impact; facilitating knowledge transfer from scientific research to industry; boosting AI education and training through multi-disciplinary programmes; building big data infrastructures and a digital data ecosystem; assessing the ethics of AI.

Values, rights and ethics are of such paramount importance that in 2019, the EU set up a group of experts to discuss the ethical implications of the technology: the high level expert group on AI, that provided a large number of recommendations, some of them aligned with the already mentioned Beijing declaration on AI and Education. After the aforementioned "White Paper on Artificial Intelligence" was released, different politicians announced multi-billion investments. Things seem to move fast, but there are doubts on whether Europe is doing enough to support AI R&I.

In the United States, after an apparent lack of activity during 2017 and 2018, AI received in 2019 a new impulse from the US government. Resources were allocated to create six National AI Research Institutes to work on trustworthy AI, foundations of machine learning, and AI for accelerating molecular manufacturing, discovery in physics, AI-augmented learning and innovation in the food system. The 2019 *US National Artificial Intelligence Research and Development Strategic Plan* identifies the following key priorities: the development of effective methods for human-AI collaboration; better understanding and addressing the ethical, legal, and societal implications of AI; ensuring the safety and security of AI systems.

China, another extremely dynamic country in AI, recently announced plans to become the leading country in AI. In particular, with the Science and Technology Innovation 2030 initiative of the Chinese government which 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 the transformation of China into an innovative country and a leading power of science and technology in the world. We would like to stress the important focus of these Chinese initiatives in Education.

The role that AI can play in education is a fundamental aspect of the impact of AI on our society. It is a major element in the Chinese strategy as mentioned before, and an essential part of some countries' AI strategies, like in the case of Finland. The way in which we educate children and citizens in AI is key for the acceptance of the technology and to prepare citizens to embrace and shape the change brought by smart technologies from a critical and constructive standpoint.

It is important to stress the relevance of the May 2019 Beijing *Declaration on AI and Education*, which was signed by more than 100 States. A consensus was reached on the urgency for Governments to prepare the workforce for the changes in training and skills required by the widespread adoption of AI systems. The reason for this call for better training and education around AI is to ensure AI technologies are used to empower teachers rather than replace them, by developing the appropriate capacity-building programmes for teachers to work alongside AI systems. The declaration wishes also to promote the equitable and inclusive use of AI irrespective of people's disability, social or economic status, ethnic or cultural background, or geographical location. It emphatizes gender equality, and aims to guarantee the ethical, transparent

and auditable use of educational data, among several other recommendations. The role of research-informed education in the preparation of our society for the integration of AI in our lives is nowadays a cornerstone of public policies, regulatory development and R&D funding strategies.

The current Covid-19 pandemia has further demonstrated that the deployment of AI technologies play a fundamental role in tackling today's global challenges; it has also expanded international collaborations and accelerated the pace of R&I on AI. An example is represented by the collaboration between the Chinese company Baidu and the Oregon State University working on predicting the secondary structure of the RNA sequence of Covid-19. Several other players, such as Deepmind or US digital giants, have been offering their expertise in AI and data analytics and their computing power.

## WHY A WHITE PAPER ON AI, DATA SCIENCE AND ROBOTICS?

Within this context, the Spanish National Research Council (CSIC) identified the need of assessing its internal capabilities and fostering the collaboration between different institutes and research groups working in the fields of AI, robotics and data science in order to identify open questions and challenging research priorities, and to sketch a strategic roadmap for the next decade.

The present white paper is the result of the collaborative effort between several research groups, which work on different aspects of AI, data science and robotics within CSIC. The works that led to this white paper started with a meeting convened at the Artificial Intelligence Research Institute (IIIA), part of CSIC, on 13 and 14 June 2019. During this preliminary meeting members of different research groups working on AI discussed the main challenges of the field and decided to create an informal communication channel named AI-HUB.CSIC for future co-ordination. This informal channel proved instrumental at the end of June to define the map of competences of CSIC on AI, as requested by the Ministry of Science to draw the Spanish AI Map, which complemented the Spanish Strategy on AI released in March 2019.

In July 2019, CSIC's scientific Vice-presidency announced the identification of 12 strategic thematic areas to be developed during the next decade (2021-30). One of them was Artificial Intelligence, Robotics and Data Science. In order to better identify research groups and individual researchers working in this area, all institutes of CSIC were invited to identify groups working on AI

foundations and applications across disciplinary domains. In total, 32 groups responded to the call for interest in the AI thematic area. This internal call offered an interesting overview of research activities in this domain. Group members discussed their research with their peers in a plenary meeting held at the IIIA on 28 and 29 November 2019. Researchers created working groups around sub-thematic areas which later became the chapters of this white paper. Two researchers were elected by members of each group to become the coordinator and adjunct coordinator of working group. They assumed the responsibility of coordinating communications within the group and with the coordinators of the thematic area. From December 2019 until May 2020, each group worked on developing the content of one chapter of this white paper. The list of authors of each chapter appears at the beginning of the chapter. At the end of the book, the complete list of CSIC researchers who have contributed to advance the field of AI, robotics and data science is available. Some researchers have contributed to the foundations of the area, others to application in a rich variety of domains.

## STRUCTURE OF THE WHITE PAPER

The white paper offers an overview of research carried out in groups and institutes located across Spain. More importantly, the paper sketches a preliminary roadmap for addressing current R&D challenges associated to AI.

The book is structured in eight chapters. Each one represents a sub-area studied within the CSIC's research community. All chapters follow the same structure: an executive summary, a comprehensive introduction and description of state of the art and current research challenges, an overview of CSIC's strategy and advantage position to tackle the challenges identified in each chapter. Over 50 challenges are discussed along this white paper. They represent a very exciting research program for a growing community within the CSIC with more than 150 permanent researchers contributing to AI.

Chapter One addresses the importance of the integration of three classical topics in AI: knowledge (or representation), reasoning and learning. Given the importance of these three areas in the history of AI, the chapter introduces the reader to the broader domain of artificial intelligence. The chapter identifies five challenges, which mostly delve into the integration of the three areas that have so far evolved quite independently. Among the challenges, special mention deserves "large-scale problem solving" that aims at overcoming

traditional limitations in complexity and size of classical AI solutions. Attacking real size problems will require renewing old methods and adapting methods proven successful in other areas of AI. Classical AI is coming back. Our future is that of a mixed society of people and artificial intelligence. A myriad of devices around us will need not only to make intelligent decisions, but they will also need to coexist and coordinate with one another to serve humans.

Chapter Two studies the challenges associated with the development of theories, and supporting technologies, for that envisioned future society. Finding ways to relate the micro world (individual autonomous agent interactions) with the macro world (the properties we seek in those large complex societies) is a key priority. This challenge requires collective and multidisciplinary effort in order to find solutions for current pressing needs—think for example of the massive adoption of autonomous vehicles and of the coordination problems arising from the coexistence of humans and autonomous systems).

Chapter Three focuses on the thriving area of machine learning. To a large extent current interest on AI derives from the spectacular results recently obtained by machine learning techniques. On the one hand, the colossal amount of data collected by administrations, government and corporations is calling for methods to use them in forecasting and analysis. On the other hand, the growing computing power at our disposal makes the use of computing-intensive AI methods feasible. The application of ML methods to a large variety of domains is reflected in the large number of groups, and thus contributing authors, involved in this chapter. Many applied challenges have been identified. Just to mention a few, the use of ML to forecast disease propagation has a tremendous potential in the mitigation of the impact of outbreaks like the current one of COVID19. There are also theoretical challenges like the Interpretability and Explainability of ML results that require urgent attention. Otherwise, we risk that these results are not accepted by our society. The embodiment of intelligence has been one of the most attractive business cards of AI.

Chapter Four presents the fascinating area of Intelligent Robotics. A number of exciting and difficult challenges are included in this chapter. One of them is particularly relevant in the context of the mixed societies mentioned in the second chapter: How to build robots that are easy to reprogram and to adapt to changing circumstances through a process of continuous learning. Robots, among many applications, will assist us and will therefore need to learn our

preferences, adapt their pre-programmed skills to our context and understand the consequences of their action under unforeseen circumstances. These are very difficult functionalities. Also, smart robotics has ethical implications that connect this chapter with the challenges discussed later in Chapter 6.

Chapter Five explores the contribution of computational cognitive models to the design and development of AI systems. These models aim at describing and simulating human cognition and behaviour by understanding the principles of intelligent action from the study of living beings and by reproducing these principles in the development of intelligent devices that mimic, simulate, and expand the cognitive and physical capabilities of living beings. Through the understanding of artificial consciousness and the design and development of computational psychological and social models and cognitive architectures, this chapter addresses difficult scientific challenges such as the possibility of creating behavioral human clones that can contribute to enhance the capabilities of cloned individuals, detect behavioural disorders and simulate preferable reactions and approaches that the cloned subject can imitate in order to adapt to difficult situations. Finally, the clone would ensure the persistence of a person's way of reasoning and memory after death; the clone would preserve these memories and hand them over to future generations.

Chapter Six discusses the ethical, legal, economic and social implications of mass scale deployment and adoption of AI systems. From a constructive standpoint, the chapter explores how to embed ethics in AI and in engineering through innovative multi-disciplinary collaborations and educational curricula that include the study of human and machine biases, the characterisation of moral agency in artificial entities, and the development of specific "robot laws" tackling the issue of the legal personality of autonomous systems. The effects of AI on the economy and society also demand special attention, especially the development of governance and R&D mechanisms to ensure the beneficial co-evolution of humans and artificially intelligent systems.

Chapter Seven addresses the demand for implementing high-speed low-power systems that can perform intelligent tasks, while consuming an affordable amount of energy and computing resources. By drawing inspiration from the capacity of biological brains to solve cognitive problems using low-power and low-speed noisy computational neurons, this chapter sheds light on the creation of high-speed sustainable hardware for AI; a type of hardware that would overcome some of the constraints of current hardware solutions based on traditional Von Neumann computer architectures. Research challenges explored

in this chapter include the development of systems featuring high-speed reaction, high-accuracy recognition, robust learning capabilities acquired by means of neuromorphic sensors and processors, Spiking and Artificial Neural Networks.

Chapter Eight focuses on smart cyber security and addresses the challenges of developing secure, safe and privacy-respectful AI systems ranging from the hardware to the application layer by enacting security-by-default principles and a more robust, formal analysis of adversarial machine learning models. By aligning AI functionality with information security requirements, this chapter tackles research challenge highly relevant to organisations and societies. An example of these challenges is the automatic detection of misinformation campaigns, especially those involving scientific findings, and the deployment of counter-strategies to debunk false information while reducing opinion polarisation.