Science & Technology
FORESIGHT
from society to research

Background Document

MAD4Future
(Models, Algorithms, and Data for the Future)

WG MAD4FUTURE

National Research
Council of Italy

MAD4Future (Models, Algorithms, and Data for the Future)

The CNR Foresight Project

The Science and Technology Foresight Project seeks to define a medium to long-term vision – 5 to 30 years – in order to elaborate coherent research strategies relevant to socially critical problems in the field of environment, health, food, energy, security and transportation.

Both, the holistic approach applied to the analysis of the topics, as well as the innovative format of the invitation-only workshops, enticed the participation of internationally acknowledged experts. This framework guaranteed all participants the necessary conditions to carry out an open interactive debate, consolidating a collective intelligence, which assisted in achieving a consensus on research priorities, knowledge gaps, and funding needs. This approach is designed to facilitate convergence towards common positions in order to address the social acceptability of future products and services and resultant market potential.

A strong consensus among the workshops' participants during the past four years has been reached regarding the urgency of scientific and technological breakthroughs in a number of issues, strongly correlated to each other, which need to be tackled. The issues can be summarized as follows:

(i) development of materials to perform different functions according to external environmental conditions and to respond to different requirements;
(ii) interaction between AI, data, models, and knowledge in order to design learning machines that provide satisfactory explanations to humans for their decisions, continuously learn to respond to unknown conditions, and robustly handle adversarial examples;
(iii) systems considered as a global entity with all kinds of dimensions and always analysed by parts with proper interfaces. The study of the interactions between the interfaces, the dynamics and the communication ways is of primary importance for understanding the functioning of the systems themselves, even if the nature of the interactions can be very varied and may require multi-scaling treatment.

The aim of the present document is to identify foresight priorities within the topic "Models, Algorithms, and Data for the Future", which indicate a roadmap for the discussion in future workshops.

The working group MAD4Future within the Foresight Project

Digital technology is pervasive in our lives. It has provoked major shifts to our habits in the last twenty years, especially fuelled by the diffusion of Internet connectivity. 'Everything connected' is a reality nowadays, thus generating a cyber-reality as complementary to the physical reality.

Connecting **'things'** to a network creates a '**system of things'**, able to collect data from the physical environment and to interact with it. The system is composed by physical and virtual (cyber) components, the former being observed and measured by sensor 'things' and the latter collecting and analysing data to derive meaningful information in a given context. This information can be used to interact with the system, thus modifying the physical environment by means of actuator 'things'. Connecting things generates a system, which can be connected to and be interdependent with other systems as well, thus forming systems of systems. The latter can operate in very large physical and virtual environments thanks to the Internet, virtually connecting everyone and everything all around the world.

All those connections, in turn potentially generating interactions, pose the challenge of **orchestrating** things, connections, computing power, and data. Orchestration can be seen as a set of techniques aimed at assigning tasks to be performed (the part to play, like in an orchestra) to different entities in a system (instruments). In the end, the entities composing a system are directly or indirectly connected to each other, thus generating interactions among physical and cyber-components, in the form of new connections or data exchanges. Those interactions, which can be thought of as data flows, may alter the behaviour of existing components and form newly connected ones. Those components, in fact, can be seen as *interactive adaptive units* [1], i.e., elements exhibiting some form of intelligence. The entity in charge of the overall orchestration, i.e., the director, *plays* a complex and large orchestra.

A different approach discharges the challenge of orchestration, instead allowing **emergent behaviour**. Even simple components, capable of few simple actions, can exhibit rather complex behaviours once in a collective environment. A system with emergent behaviour and functionalities evolves over time, but no director can be identified guiding such a process. Yet, the system is there.

Systems of systems are fed by enormous amounts of raw **data** being injected into it by millions of devices, such as sensors, smartphones, wearables, laptops, and any connected 'thing'. Those streams of raw data are analysed by means of very different **models** implemented in the form of **algorithms**, which should be efficient, energy-efficient, and able to derive high-quality information. Information and knowledge are key to the **future**, being extremely precious for humans in charge of complex decision making, and also for digital technology itself, which is becoming increasingly capable of suggesting and even autonomously accomplishing tasks that were once the prerogative of humans.

At the forefront of the technologies that may aid closing the data-to-knowledge loop, we can undoubtedly count Artificial Intelligence (AI), which has been recently named as the 21st century's most impacting technology since "helping us build a "*fifth element*" after air, earth, water and fire: a *data layer* that increasingly surrounds us, gradually virtualizing our environment and multiplying our possibilities as mankind" [2].

Artificial Intelligence (AI) can be defined as the science and engineering of making **intelligent machines,** whose application has started to pervasively permeate our daily life and contribute to optimise logistics and industrial production, better diagnose diseases, fully tailor recommendations to specific tastes or needs and better understand societal phenomena.

Dating back to the early 40's, AI has evolved passing through various definitions and experiencing alternating success and interest, with the recent renaissance mainly due to the data deluge that has boosted data-driven methods and data science (see the end-note[i]). At present, AI is intended as the broad science that collects the classic AI and Machine Learning, thereby delivering machines that exhibit intelligent behaviour along the three axes reported in Figure 1: **learning**, **reasoning** and **planning**, **acting in** and **adapting to the environment.**

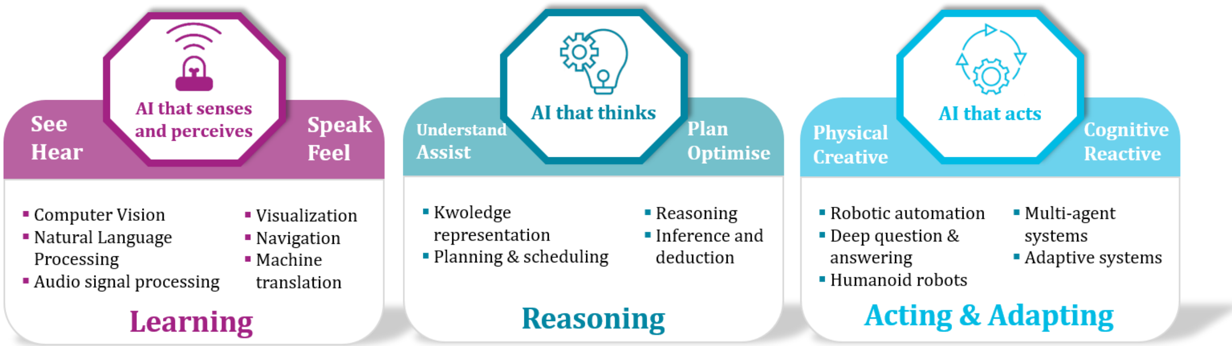

Figure 1. The three axes of AI capabilities (inspired by Grasso's infographics)

Most of the currently striking examples of AI, such as autonomous driving, gaming, virtual assistants, transformers, creative writing of journalistic reports or painting, strongly leverage the powerful learning capabilities emerged lately. Overall, the successful AI forms that exist today in our world are usually focused on specific tasks and mainly concentrated on one of the three axes in Figure 1[1]. Being trained with specific data or programmed with specific procedures to operate within a predetermined and pre-defined range of actions, these solutions are categorised as **Narrow AI**[2]. The main characteristics of these solutions is that they cannot be generalised to other tasks and their performance tends to decay when exposed to unexperienced conditions.

The current debate around the future of AI is grounded on different perspectives. From a scientific standpoint, researchers are working to define effective strategies to expand the capabilities, the efficiency and the efficacy of intelligent machines, while from the social and ethical standpoint scientists and stakeholders are engaged in assessing and regulating the impact that these technologies may have on society. The two lines of course overlay and exchange views and outcomes.

In the research communities, the great hype around data-powered intelligent machines is being followed by reflections on the inner limits of learning and data-driven models. Main concerns pertain to:

- quality, quantity, diversity and fairness of the data needed to train effective models;
- the yet unproved assumption that, with sufficient data, any complex model could be learnt;
- the peculiarity of these methods to mainly extract and rely on correlations among data, instead of causal relationships and a clear understanding of a phenomenon;

---

[1] Unless some hybrid solutions that are able to combine multiple abilities at once, such as autonomous driving and firstly attempts to build humanoid robots

[2] Opposite to human-like General Intelligence, which is the overall quest of AI

- the difficulties in identifying the latent variables, especially in models based on network of interconnected variables[3];
- the incapacity to cope with unseen cases;
- the lack of contextual reasoning capabilities.

Furthermore, models based on Deep Learning, one of the most powerful learning approach to date, are undoubtedly powerful computational machineries, based on a rather flat mathematical idea corresponding to a complex layered network of simple processing units whose connections and behaviour emerge from a training procedure that learns a structured representation of observations and data. The high number of dimensions of this structure and a certain lack of a clear declarative representation of knowledge bring along difficulties in generating the necessary explanations and/or interpretations, and in turn limit its ability to characterize the classification strategies and gather useful insights into the phenomenon at hand. Overall, many data-driven approaches are considered to be opaque, lacking immediate interpretability of their inner functioning [3] and they have been demonstrated to be susceptible to deceit when exposed to adversarial examples [4],[5], thus generating the lack of trust that is being tackled from an ethical standpoint. Opposite to this, knowledge- or model-based approaches appear more transparent and explainable, based on explicit modelling of the problem [6]. Nonetheless, such methods are affected by other challenges, as acquiring value knowledge and creating comprehensive models is rather difficult and, often, the resulting solutions are less powerful.

Some authors demystify the inner opacity of learning models in the fact that intelligence, for its nature, has only some parts exposed to a rational explanation. Some of it is *just* instinctual and based on intuition. This standpoint relies on the idea that current deep learning models mainly implement what Kahneman [6] defines as *system 1* (fast and instinctive), while inference-based, reasoning approaches mainly correspond to *system 2* (slower, deliberative and logical) [8]. The dichotomy between these two systems has emerged lately as a guiding principle to endow intelligent machines with the right mix of learning and reasoning abilities, that is with both data-driven and model- or knowledge-driven functionalities (see the second end-note of this document[ii]).

Overall, to really move from data to knowledge, thus truly understanding a given phenomenon or effectively implementing a systemic approach to a problem, we share the opinion that relying on data-driven or model-driven approaches alone is not profitable and, instead, a suitable fusion of the two is needed.

Moreover, besides the research per se, we further remark that the ultimate goal of intelligent machines would be **establishing a synergy with human users**, to produce around them a kind of intelligent *exoskeleton* that enables them to do something they cannot do on their own. For instance, and recalling the aforementioned metaphor of the music director, such a synergy may provide a powerful assistant to humans in several fields of life. In turn, AI can be thought as in charge of reducing the number of parameters (i.e., the complexity) a human director must consider in his/her activity, thus effectively

---

[3] When we represent a problem through a network of interconnected variables, a parallel question is how to select highly influencing nodes and quantify the extent with which these can disseminate information

providing support when facing complex and multi-faceted systems to solve the grand societal challenges. In this respect, an emerging field of research deals with how to make such an exoskeleton really helpful for the human hood, establishing the main principles for a Humane-AI in order to create favourable regulatory environments and establishing global standards and best practices to ensure that all AI-based technologies provide a really positive added-value to society. This is the main subject of the debate on the future AI being carried out from the societal and ethical standpoint. The ultimate goal is establishing a priority for ethics, security and trust. Floridi's recommendations [9] about an ethical framework for a good AI society, among others, should drive current and future AI-related developments through the 5 principles therein: beneficence, non-maleficence, autonomy, justice, and explicability. This approach has led to the definition of trustworthy AI, which is the focus of the European strategy on AI [10].

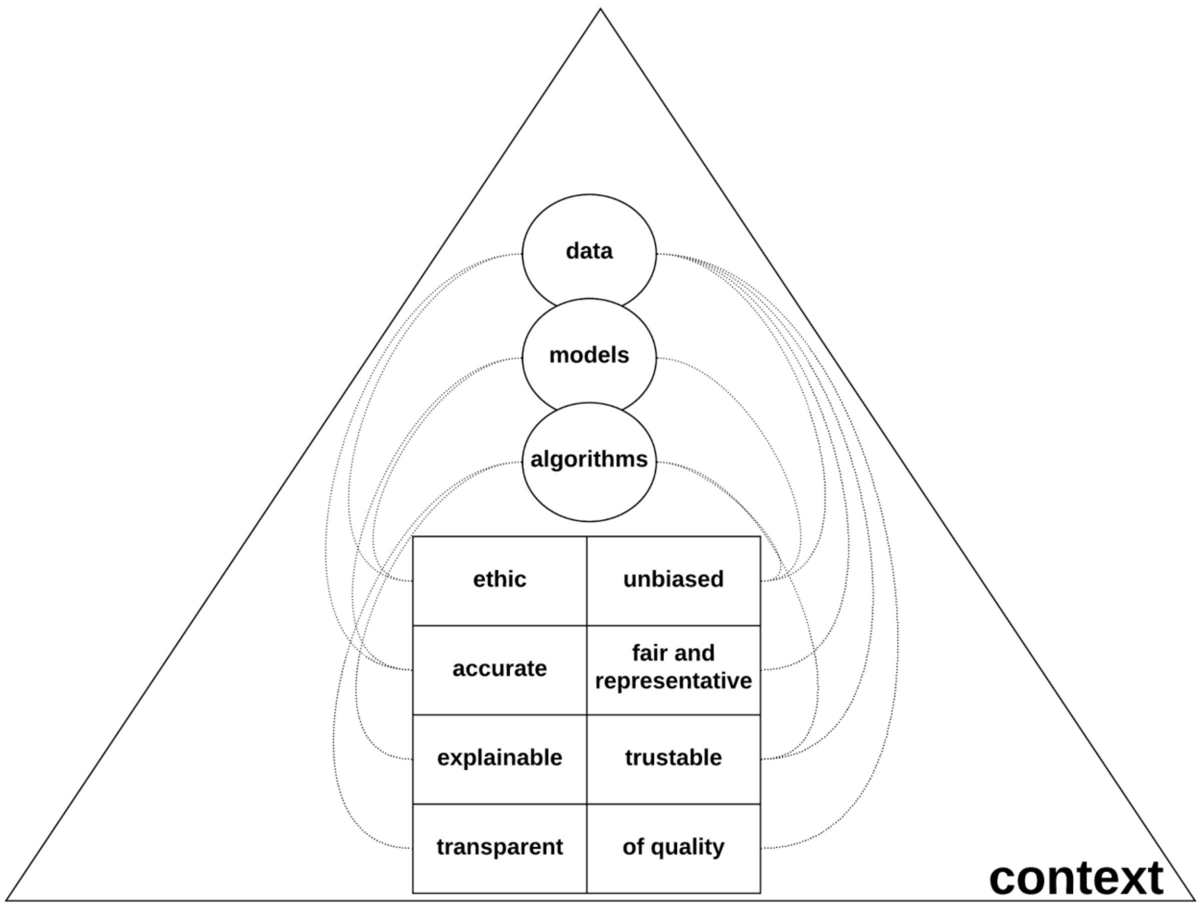

Figure 2: MAD in a given context exhibiting key desirable features.

The scope of the MAD WG is to investigate how models, algorithms, and data (MAD) can be leveraged to treat system-of-systems in an organic approach towards solving the societal grand challenges, and clearly taking into account key desirable features that MAD should exhibit in a given context, as logically depicted in Figure 2.

In this respect, we would like to debate the real breakthrough required to merge data-driven learning and model-based theories, so to create more informative and really supportive AI-based tools that can shade lights on and help to manage, possibly in a semi-automatic way (according to the above exoskeleton metaphor), the complexity of real-world challenges. The ultimate goal is the delivery of intelligent machines that are able to address the limitations of learning and inferential analyses, including knowledge gaps and convoluted dynamics that depend on:

a) context, thus being able to explore causality versus correlation;
b) method, thus being able to be integrated into the scientific process;
c) latency and/or influence, thus becoming able to guide complex scientific models and interpreting ML algorithms;
d) variations due to new or unexpected conditions.

Main challenges

At today, the main challenges come from the following aspects, in our opinion.

The **complexity of systems** is increasing at a fast pace, so new algorithms and huge computing power are also necessary to support intelligent machines. Whether orchestrated or self-organizing systems, unprecedented complexity is likely to be expected to support 'creativity', 'feeling', and 'thinking out of the box'. Those 'features' should be accompanied by explainability and capability to provide causal relationships, exploiting context awareness and techniques as few-shot learning.

The field of **data science** should probably evolve for defining its quantitative and qualitative role coherently across all major domains and spheres of application. New generation of data scientists should be trained in using data to its best advantage towards current and future needs, in order to maximize the potential to drive discovery and develop a principled view of data science that properly combines inferential thinking and computational thinking seems necessary.

High-quality data are necessary to properly feed data-driven approaches, and quality should be preferred to quantity in order to reduce the impact of biased data, and thus provide a fair and representative description of a specific phenomenon, carefully considering its context. Thus, data-driven approaches must be accompanied by methodologies and algorithms that integrate and validate efficiently and efficaciously those data; or complemented by model-driven approaches that overcome the aforementioned limitations, as identified so far.

REFERENCES

[1] Leibo J.Z., Hughes E., Lanctot M., Graepel T., Autocurricula and the Emergence of Innovation from Social Interaction: A Manifesto for Multi-Agent Intelligence Research

[2] Renda A. Artificial Intelligence. Ethics, governance and policy challenges. Feb 2019

[3] Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D., A survey of methods for explaining black box models. ACM computing surveys (CSUR). 2018 Aug 22;51(5):93.

[4] Szegedy et al. Intriguing properties of neural networks, ICLR, 2014

[5] Huang et al. Adversarial machine learning. 4th ACM workshop on security and artificial intelligence, 2011

[6] Martinez D.R. et al. Artificial Intelligence: Short History, Present Developments, and Future Outlook. MIT Lincoln Laboratory Report, 2019

[7] Kahneman D. Thinking fast and slow. Farrar, Straus and Giroux, 2011

[8] Falchi F. About Deep Learning, Intuition and Thinking. ERCIM NEWS. 2019 Jan 1(116):14

[9] Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B., AI4People—An ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds and Machines. 2018 Dec 1;28(4):689-707.

[10] AI HLEG. Ethics guidelines for trustworthy AI. Apr 2019

[11] Goodfellow et al. Explaining and harnessing adversarial examples, in: ICLR, 2015

[12] Wang et al. Generalizing from a Few Examples: A Survey on Few-Shot Learning. arXiv, 2019

[13] Sengupta A. et al. Going Deeper in Spiking Neural Networks: VGG and Residual Architectures. Frontiers, 2019

[14] Alyusin A. et al. Memristive Element Functional Model for Computer Based Analysis and Hardware Emulation of Pulsed Neurons Adaptive Networks

[15] Serb a. et al. Memristive synapses connect brain and silicon spiking neurons. Scientific Reports 10, 2020

Sara Colantonio, CNR-ISTI (sara.colantonio@isti.cnr.it)
F. Manlio Bacco, CNR-ISTI (manlio.bacco@isti.cnr.it)

Scientific committee (in alphabetical order):
Enrico Capobianco, Fabrizio Falchi, Paolo Ferragina, Michela Natilli, Salvatore Rinzivillo

[i] Traditionally, the term AI was referred to an umbrella of disciplines focused on reproducing the human-like capabilities of reasoning, task accomplishment and decision-making. Under this umbrella, one can find logic and deductive reasoning (i.e., inference), planning and optimization, knowledge representation and decision-support theory, autonomous agent theories and methodologies. Their underlying commonality is the use of coded rules and procedures to mimic human intelligence and behaviour. Expert systems or logic-based agents and games are the typical products of these disciplines. Historically, this approach is known as classic AI and stood opposite to another group of methods and techniques, usually referred to as Machine Learning and Computational Intelligence, whose key idea is "learning and adapting from experience". Artificial Neural Networks, Genetic Algorithms, Decision trees and statistical methods, such as Bayesian Networks, gather in this group. Their peculiarity is the ability to discover patterns and information from a set of data or examples. This ability enables them to autonomously learn how to recognise categories, perceive stimuli, predict trends, make decisions and apprehend behaviours. Simply speaking, ML methods are not specifically programmed to solve a task, but learn how to solve it based on the data available. ML is at the core of data science and big data analytics. In the last decade, boosted by the data deluge, ML has made significant steps forward. A particular type of Artificial Neural Network models, based on Deep Learning (DL), has demonstrated to perform unprecedentedly well, especially when solving *perception* tasks, such as vision, object recognition and natural language processing. The peculiar prerogative of DL methods are their capacity to autonomously learn what in the data is really salience to accomplish a given task. This is why DL is also called *representation learning*. The large success of DL, fueled by the Big Data era, is the main driver behind the recent renaissance and flourishment of AI.

[ii] Already a couple of years ago, one of the fathers of Deep Learning, Yann LeCun, has remarked the limitation of current learning models saying that "though perception 'really works', what is still missing is reasoning". Here, the approaches on how to meet this goal fall in two opposite lines of research. The first one stems from the idea of mixing together the different souls of AI into hybrid models able to combine data-driven, model-free methods with knowledge-based, model-driven ones, working to overcome the limitations of the two fields. The other line of research is convinced that the real breakthrough could come from advancing data-driven models to reproduce reasoning by modelling attention, consciousness priors and meta-learning. Also, the susceptibility to adversarial examples is being harnessed as a possibility to get further insights into the functioning of these models [11] and few-shot learning is being put forward as a way to cope with the lack of data and simulate human ability in this respect [12]. Reinforcement learning and continuous learning are also being further explored to let a learning model cope with changing environment. Other research avenues that take inspirations from the biology and morphology of human nervous system (by reshaping the concept of spiking neurons [13] or developing new neuromorphic technologies such as the memistor [14], also trying to link them to biological neurons [15]) or exploring the use of quantum computing.