*Steering Committee*

**International Conference**
**"Ethics and Artificial Intelligence"**
***(Venice, September 24-25, 2021)***

**Sotto l'Alto Patronato del Presidente della Repubblica**
*(Under the High Patronage of the President of the Italian Republic)*

*with the contribution of:*
TIM
Intesa Sanpaolo

*in collaboration with:*
Aspen Institute Germany
Institut Aspen France
Academy of Sciences of Bologna Institute

# *The Way to Venice*
# *Report*

# CONTENT



*Edited by Gyneth Sick*

# *The Way to Venice Report*

## Letter to Conference Participants

September 1, 2021

*Dear Participant,*

Thank you for confirming your attendance at the international conference on *Ethics and Artificial Intelligence*, on September 24-25, 2021. Prominent leaders are joining in this interdisciplinary debate, speaking on behalf of businesses, professional sectors, the academic and scientific worlds, institutions, the media and the general public.

The objective of the conference is to create an event with significant relevance in the growing debate over ethics and AI, and to build a base for future initiatives. Among the themes being discussed are the impact of artificial intelligence on citizens and businesses, how the scientific community can help companies design and deploy ethical AI, what norms and regulations need to be modified in different sectors, and the relative ethical considerations in an ever-changing approach involving the whole world.

This preparatory Report *"The Way to Venice"* was conceived as a tool, to enable participants to approach the themes of each session from a common starting point, anchoring discussions in an exploration of the conference's objectives. The Report serves as a springboard, challenging us all to re-think the question of ethics in artificial intelligence.

The Report was drafted by a *Steering Committee*: high-level figures with expertise in different sectors relevant to the subject at hand (the list of experts is provided in Annex 1). A *Business Users Advisory Board* also reviewed the Report, adding insights specific to corporate issues (the list of Advisory Board members is provided in Annex 2). A *Conference Honorary Committee* (the list of members is provided in Annex 3) is involved to promote a successful in-depth analysis and interdisciplinary debate. We, the undersigned, wish to express our gratitude to all of the above, for their valuable contributions.

By sending you this Report, we are welcoming you to join the spirit of this conference. "The Way to Venice" has been a collaborative journey, in the company of an accomplished, diverse group of peers. We hope the event itself will contribute to this community of learning, offering a meaningful touchstone for many years to come. We consider it an important step toward future initiatives aimed at encouraging an ethical approach to artificial intelligence.

Luigi Gubitosi
Chief Executive Officer
TIM

Giulio Tremonti
Chairman
Aspen Institute Italia

Carlo Messina
Chief Executive Officer
Intesa Sanpaolo

*The Way to Venice Report: the main issues*

# 1. AI and ethical challenges for businesses and the economy

Over the last two centuries, our societies have experienced an industrial revolution every fifty years or so. Very recently, however, technological revolutions have become much more frequent. Today, in particular, it is estimated that a revolution in the field occurs as often as every 3-7 years. The latest – and the one still in progress – is that of artificial intelligence.

This current revolution is having a very strong impact on our lives. While automation, for example, might primarily have affected the lives of factory workers, we all come into direct contact with artificial intelligence (AI) on a daily basis. We all use and are increasingly coming to depend upon devices that use AI. The services and improved performance they provide are changing our lives in an infinite number of ways. To cite just one small example: our mobile phone cameras are now capable of detecting people's faces and adjusting the focus to optimize quality.

Furthermore, the recent pandemic has accelerated digitalization processes. Public and private organizations have revolutionized the way they operate, changing their internal decisional and operational processes. Digital connections, remote work, cloud computing, 5G networks and so on have had a significant effect on businesses, supply chains, customer relations and the environment, decreasing the global footprint of companies as well as their economic costs.

While the Covid emergency has undoubtedly led to numerous and profound problems – for people's health, for our economies and on a social level – it has nevertheless also offered a fresh opportunity to respond to the big challenges of our times. From this perspective, as the European Commission President Ursula von der Leyen said, "It is up to us to build back better, together." The goal is to build a new development model for the European Union's post-pandemic world: a more digital, ecological and inclusive one.

*AI and ethics at a corporate level*

When discussing "Ethics and AI" at a corporate level it is necessary to make a stark distinction between two very different domains of AI applications. This is because it is not AI itself that calls for the introduction of a certain set of ethical principles. Whether AI systems abide by a common set of ethical principles or not depends first and foremost on the use people make of them.

The first field of applications relates to the use of artificial intelligence to improve the performances and the efficiency of a set of machines and/or equipment (i.e. a production process) or of internal processes. This domain operates "within the company's walls" and does not produce targeted effects on single individuals or firms (suppliers/clients). The use of such AI systems may be labelled as "internal" and/or "undifferentiated".

For example: common "customer care chatbot systems" (that rely on artificial intelligence) do interact with single individuals (also because the bots are typically trained with real customers' requests); however, as long as they do not provide different answers on the basis of the specific customer's characteristics (i.e. personal data) they are still to be considered "internal" and "undifferentiated". In other words, such AI systems are not susceptible to generating discriminatory outcomes.
A "bad" or "biased" functioning of internal and undifferentiated AI systems may affect customers, but would do so with respect to the company's overall customer base, not individually.

This is why it is not deemed necessary to subject internal and undifferentiated AI systems to any specific additional "ethical principle". The one exception to this rule is for the traditional "environmental, social and corporate" principles to which companies adhere. For this sort of application, AI systems are to be considered a "technological evolution of production management techniques". Examples of such systems are ERP (Enterprise Resource Planning) or SAP (Systems Applications and Products in Data Processing); that is to say, they are seen as tools that improve a company's efficiency and effectiveness and should be treated and evaluated as such.
If an AI system works properly, the company using it will gain efficiency, drive down costs and strengthen its competitive positioning. If it works less efficiently than expected, the firm might instead lose market share with respect to its competitors (and, as a consequence, it might decide to change its AI system). "Internal AI systems" are used to help companies to improve

their economic and technical performances; as such, it is thus correct to say that internal AI systems respond to market dynamics.

Conversely, the second domain of AI system applications has to do with individual and targeted responses or actions; these may produce relevant effects on single individuals or firms. Typical examples of this second domain of AI applications are approvals of mortgage loans, e-recruiting processes, on-line credit applications or tailored pricing practices.

Decisions directed to single individuals or firms and taken on the basis of specific personal information (that may also even prove to be incorrect) are not "neutral": they may cause unjust repercussions for the persons or individual firms targeted.

Moreover, decisions taken by automated systems (i.e. decisions that are based solely on data elaborated by a machine without human supervision) are not necessarily "fair". "Badly designed or trained" automated decision systems may lead to biased decisions. The fight against bias is extremely important for our values and our society. We cannot accept the risk that discrimination hide behind an algorithm. This is why it must always be possible to question a conclusion reached by an automated system. Furthermore, the decision to contest or overturn such a conclusion must be based on a factual and real understanding of how the machine has reached said conclusion.

For the above reasons, whenever automated decisions are based on an individual's personal data and whenever they may have important effects on people's lives, ethical principles shall be called into action.

*The underutilization of AI*

The impact of AI on corporate decisions, for now, has remained an unfinished revolution. According to estimates in the 2020 annual reports of the most important consulting firms, more than 80% of the C-Suite managers interviewed consider the use of artificial intelligence within their company to be strategic. This holds true regardless of the products or sectors to which they belong. And yet, according to the same reports, fewer than 40% of those managers had produced a clear AI implementation roadmap. This low adoption of technology within companies stimulates some further considerations.

First, the underuse of artificial intelligence risks slowing the growth of entire industries. Indeed, in addition to process efficiency and product innovation, AI allows companies to clearly distinguish users and customers, to greatly increase customer intimacy through individually tailored services, and to completely change their business models (for example, think of direct and indirect data monetization).
A great example of the underuse of technology is represented by digital payments. In Italy, the number of businesses that accept digital payments is far below the European average; the digital instrument is rarely preferred to cash as a means of payment. Nonetheless, digital technology is a tool to protect the consumer, who, thanks to artificial intelligence, can be accompanied in making financial choices, advised in managing budgets and risk, and protected from theft and fraud. Businesses can benefit as well, for example, by providing dynamic discounts on net commissions paid via digital payments. Such dynamic pricing can vary depending on the portion of digital payments accepted out of the total proceeds collected, with different magnitudes according to the size of the merchant (industry giants or SMEs). Furthermore, AI has the power to act proactively – not only retrospectively: by analyzing recent transaction volumes, for example, AI systems can provide businesses with insights and analysis on the most successful business hours, or make suggestions on marketing campaigns. Such activity can play a key role in addressing the needs of targeted clients.

Another significant downside, due to the underutilization of artificial intelligence, is the fact that those who move too late or make the wrong digital strategy choices risk being left out of the market. Market data from recent years show, in fact, that in the digital sector, those who succeed in quickly imposing their value propositions become leaders. This supports a principle that McKinsey defines as "winner takes all".

The business community generally agrees that digital transformations are more about strategy and people than about technology per se. Indeed, in the vast majority of cases, technology has historically improved the quality and safety of work, and yet it has almost always had to overcome great resistance and even fear to do so. To gain a hold in society, new technologies have often had to go through extremely complex transformations. For example, consider the Luddite movement: in the late 1700s, the Luddites sided with workers against automation and industrial machinery, in the belief that such developments could cause a crisis in wage labor.

*AI impact on employees*

Automation has certainly changed the way we work. In some cases, it has even reduced the need for human labor, as the Luddites feared. Today, however, no worker is likely to be willing to return to pre-industrial revolution working conditions.

In the future, it will be increasingly important to build the skills, or capabilities, to use analytics and artificial intelligence to support different types of jobs. Indeed, we already have examples of new companies that have made AI their trump card. Predictably, the net positive impact on employment from AI use is to increase (not replace) areas of employment, while also fostering the emergence of new economies and new professions. Even new disciplines and opportunities for academia arise from the adoption of AI systems. Within companies, the most obvious impact is that the organizational structure remains; the greatest transformation is seen more in the reorganization of activities. It follows that most jobs as we know them today will be transformed; new figures and professional roles will enter or evolve to fill skills gaps, as talent is developed. It is thus reasonable to assume that artificial intelligence will contribute to expanding employment.
In this context, it is clear that technical and managerial skills need to be continuously updated, and that employees need to be accompanied on a path of growth from a business perspective. It will be important to work on the diffusion of understanding artificial intelligence and on an AI culture. Furthermore, as AI requires human supervision, the emergence of new categories and skills is not limited to the technical sphere: updates will also be necessary in areas involving morality, ethics, and social intelligence. Many business sectors have already initiated a dialogue with major Italian universities with the aim of developing, through ad hoc training courses, the new competencies necessary to meet the challenges of tomorrow.

Today, artificial intelligence is being greeted with enthusiasm and high hopes, but also with fear and resistance. Resistance owes a great deal to a general lack of expertise. Not appreciating the potential nor understanding the limits of artificial intelligence generates a sense of powerlessness, which can result in rejection, as we saw with the Luddites.

In France, for instance, there is an institute – with a significant number of followers – that fights the use of artificial intelligence. The success of this

kind of association is at least partially due to a lack of information or its poor quality, even from highly competent journalists and scientists.

*The question of data privacy and protection*

A landmark study published in Amsterdam in July 2020 shows that 7 out of 10 people worldwide are concerned about sharing personal information, while two-thirds of the global population does not like the current privacy practices of most data collectors. The survey of more than 25,000 people in 40 countries was designed to understand people's attitudes to sharing personal data with businesses and was conducted by the Worldwide Independent Network for Market Research (WIN) and ESOMAR, the global trade association for the data and insight industry. The results show that across the world the vast majority of people entertain significant concerns about sharing their personal data. Furthermore, many have already been the victims of data misuse. Overall, the study showed a huge disconnect between businesses and the public in regard to data collection, its value, and its use.

The collapse in trust between the public and companies most likely has its roots in widespread data misuse (spam, phishing, email hacks, leaked personal details or bank account/credit card hacks). More than half of the world's population has been a victim of data misuse. In Europe, that figure reaches more than two thirds; in Latin America and Africa, around three in five people have been victims, while in North America a staggering 80% of the population has been the victim of some kind of data misuse.

Globally, more than two-thirds of people understand that their personal information is valuable for data collectors, and this is particularly pronounced in Europe (where 74% of people agree), and in the Asia Pacific (75%). However, less than half of the world's population considers sharing personal information a vital necessity in our connected world – a figure that rises to two-thirds of people in the Asia Pacific region. Less than a third of people in Latin America and the United States consider sharing data vital. This suggests that many consumers consider data collection to be a one-sided deal: it's valuable for businesses, but on the whole not necessary. In North America, Latin America, Europe, and MENA (Middle East and North Africa), fewer than half of all consumers are aware of what happens to their personal information after they have shared it with a data collector.

So one of the key challenges facing us today is how to raise awareness about the benefits of sharing personal data – as opposed to the risks. To cite a simple example: sharing medical records and early insights on life-threatening diseases clearly benefits patients, conceivably outperforming any risk associated with the potential loss of privacy. Enhancement in cybersecurity applied to data protection will be key in this respect.

*Adapting to the job market*

According to a study by AlphaBeta consulting, a large part of today's jobs will disappear or be heavily modified. The good news, however, is that each of us is one skill away from at least 17 of the jobs of the future. So how can we keep our profile attractive and up-to-date with the market?
One answer is to keep upskilling. Upskilling, however, is a complicated issue and if it is to be structural, it will necessarily change the organization of time and work. On this, Europe more than any other region leaves accountability for such reorganization to the initiative of the company. According to the World Economic Forum, in fact, companies in the European Union tend to prefer to provide their employees with new skills rather than replace them. When the company, however, does not believe in or cannot activate continuous learning projects, the responsibility must fall on the individual or on the State. For some years now, for example, Singapore has launched a mandatory lifelong learning program for all its citizens. This type of program goes hand in hand with a national strategy to meet the needs of a country's economic fabric in the short, medium and long term.

*Technology infrastructure as an AI-enabler*

In the global artificial intelligence market, Italy in particular runs a further risk. For the algorithmic part to develop and industrial applications to be allowed, there is not only a need for expertise, data and algorithms, but also for a great deal of implementation, infrastructures and trust.

The banking industry has certainly developed an important use of AI already, but a more extensive application should still be pursued. In fact, the kind of technology used most often is the so-called "weak AI". This term characterizes systems that are not equipped with strong AI but behave as if they were: SIRI, for example. To an outside observer, such systems appear to be equipped with functionality that is not actually implemented within them. For instance: some banking institutions have modernized and have begun managing customer relationships through personalized, hyper-

performing banking chatbot systems; these can understand customer needs and answer questions in a relevant way about specific and common topics. Such systems are, broadly speaking, problem-solving programs capable of replicating some human logical reasoning, making decisions, and so on. In contrast, strong AI – which has yet to be adopted by banks – characterizes those computer systems that can perceive, understand, and interpret their environment, learn from past behaviors, and apply this learning to future behaviors. So far, AI systems are mostly used in customer care, business planning and development, and operations. Overall, AI systems are increasingly in production, and various systems under study are being applied in the pilot phase. A case in point is in the area of AML (Anti Money Laundering) where, despite a low percentage of application (20%), all the cases reviewed are already in production.
Finally, the areas in which, to date, there have been fewer cases of AI in production are in fields that can be summed up under the headings "credit", "finance" and "know your customer". Support for the world of compliance – especially in terms of automatic control of the correctness and completeness of documentation – is improving. This is especially true of large banks (around 50% of the sample analyzed), while medium-sized banks are following suit.

In order to effectively leverage artificial intelligence as the real driving force behind the transformation of banking, we must also consider its challenges. Its implementation is undoubtedly complex. As reported in the ABI Lab Scenario, 40% of companies boast a high rate of AI implementation, while the other 52% have opted only for medium implementation. Another challenge on this front relates to improving internal skills (36% high – 56% medium) and, finally, reducing implementation times (16% high – 76% medium). Interestingly, the issue of corporate culture was considered critical by 52% of respondents, which highlights the importance of pursuing cultural transformation initiatives that help create fertile ground for AI-driven innovation opportunities.

*AI for sustainability*

AI is an increasing factor of sustainable economic growth. Many researchers have shown how the wide range of methods referable to AI and Data Science is significantly increasing the ability to observe, measure, model and predict complex socioeconomic phenomena, magnifying the impact of AI systems on our society.

A study conducted by McKinsey in 2018 mapped 160 cases of AI with a social impact. The analysis demonstrates how AI can actively contribute to the achievement of all 17 of the United Nations' Sustainable Development Goals. Specifically, thanks to the computational ability to process large amounts of data, AI plays an increasing role in areas such as access to health care (SDG 3), education (SDG 4), and environmental sustainability (SDG 12, SDG 13). Sustainable AI can work in the direction of improving AI development and applications by fostering change throughout the entire life cycle of AI products and towards greater ecological integrity and social justice. As such, sustainable AI is focused on more than mere applications; rather, it tackles the whole socio-technical system of artificial intelligence, so as to develop a technology that is compatible with sustaining environmental resources for current and future generations.

# 2. The regulation of AI: between freedom and dignity

The AI revolution has disclosed huge and unexpected opportunities for everyone. However, together with this potential for greater welfare, knowledge and wealth, the development of AI and deep learning-based technologies can also be a source of new challenges.

In 2019 the European Commission High-Level Expert Group on AI published the "Ethics Guidelines for Trustworthy AI". This document aims at encouraging AI technology that is lawful ("respecting all applicable laws and regulations"), ethical ("respecting ethical principles and values") and robust ("both from a technical perspective while taking into account social environment").

Legislation and standardization – at national, EU and international levels – play an important part in guaranteeing safe and reliable products, environments and services. They provide common rules for actors and define the minimum requests for activities and research results and innovation. However, legislation alone cannot guarantee that research results and the direct and indirect effects of innovation activities are ethically acceptable and sustainable. The laws are intended to establish common rules and procedures in order to prevent unethical behavior and foreseeable risks. This is why a reflection on ethics in the context of technological innovation is essential. It is a complex issue, which requires a multi-stakeholder approach (involving providers, companies, academic bodies and institutions), in which companies first and foremost have the responsibility of fueling an open and inclusive debate, sharing best practices, discussing dilemmas and reporting emerging issues relating to ethical concerns.

In a broader sense, legislation should guarantee safety and security for its users and stakeholders. The legal protection of consumers and workers depends on fundamental rights, indeed, and is paramount. When it comes to artificial intelligence specifically, safety risks could arise from autonomy (autonomous decisions that may deviate from original intentions), data dependency (faulty data and poor decisions), opacity (difficulty in predicting decisions), complexity of products, interaction with other devices, and the value chain. Moreover, such risks can be exacerbated in case of AI applications in robotics (autonomous vehicles, drones, care robots, medical robots, etc.). To minimize the risk of harm to users, workers and other

exposed persons, when developing and providing AI systems, a high level of protection is crucial. In this regard, in order to build trust on AI and develop it in a human-centric direction, AI systems should integrate safety and security-by-design mechanisms.

In AI systems, cybersecurity threats can also affect people's safety, as well as people's right to privacy and data protection. Harmful decision-making in AI, with possible consequences for personal safety, can come from a violation of availability, integrity and confidentiality of the information processed. An unauthorized third party accessing a car's control system, for example, could easily lead to an accident. Other fundamental rights and freedoms, such as the right not to be discriminated against or freedom of movement, are also at risk.

On this matter, the European Commission's "Proposal for a Regulation of the European Parliament and of the Council Laying down Harmonized Rules on Artificial Intelligence and Amending Certain Union Legislative Acts" (AI act), published on 21 April 2021, recognizes the strict relationship between security and safety. The Proposal stresses the crucial role played by cybersecurity in ensuring the resilience of AI systems against attempts to alter their functions. As the European debate shows, regulations on safety and security relating to AI is a crucial topic. Solutions will help build trust about AI and protect fundamental rights.

Many of the ethical issues raised in the following parts of this Report are likely to find a valuable solution through the implementation of the EU AI Act Proposal, an important milestone in a process that has already seen the adoption of the Ethics Guidelines (February 19, 2020) and the White Paper on Artificial Intelligence (April 8, 2019). The Machinery Product Regulation (which covers new risks stemming from the new digital technologies and focuses on the safe integration of AI systems into machinery) and the General Product Safety Directive and Radio Equipment Directive are also part of the ongoing development of ethical AI.

The AI Act succeeds in outlining a profitable way to AI regulation and to the construction of a European AI market. It aims to define a balanced and proportionate horizontal regulatory approach to AI that is limited to the minimum necessary requirements to address the risks and problems linked to AI. The Act should not unduly constrain or hinder technological development nor otherwise disproportionately increase the cost of placing AI solutions on the market.

The proposal sets a robust and flexible legal framework. On the one hand, it is comprehensive and future-proof in its fundamental regulatory choices, including the principle-based requirements with which AI systems should comply. On the other hand, it puts in place a proportionate regulatory

system centered on a well-defined risk-based regulatory approach that does not create unnecessary restrictions to trade, whereby legal intervention is tailored to those concrete situations where there is a justified cause for concern or where such concern can reasonably be anticipated in the near future. At the same time, the legal framework includes flexible mechanisms that enable it to be dynamically adapted as the technology evolves and new concerning situations emerge.

The European Commission's Proposal is characterized by a promising approach based on the following principles: the pursuit of AI systems that are respectful of fundamental rights, reliable and safe; the necessity of harmonized regulation; the application of European rules even to AI operators not established on EU territory; a coherent system to evaluate the different risks of impact on fundamental rights; the proportionality between level of risk and level of constraint; the choice to define AI systems rather than AI itself; and the encouragement of innovative research on AI through the so-called "sandbox mechanism" – a system of limited and controlled spaces aimed at mitigating the impact of failures and security problems. Finally, the document assigns special powers to the European Commission, to define the eventual risks of new AI systems.

*AI and fundamental rights*

Any regulations adopted in the area of AI must first take into account their impact on fundamental rights as enshrined in the Charter and the EU Treaties. The EU and Member States are obliged to rely on "robust evidence concerning AI's impact on fundamental rights" so as to ensure that any restrictions (say for consumer protection) respect the principles of necessity and proportionality. This is a primary safeguard in the field of law and technology, where the novel nature of technology means that there may not be any case studies assessing the necessity or proportionality of a specific limitation on fundamental rights and freedoms. Another challenge is how to define AI, which must constantly be updated in line with technological developments.

Within this framework, relevant safeguards must be put in place in order to ensure effective protection against arbitrary interference with fundamental rights. Legal certainty for both AI developers and users is also necessary. Voluntary schemes for observing and safeguarding fundamental rights in relation to the development and use of AI can help to mitigate any further violations. Therefore, the notion of legality is also firmly linked to a precise yet not too rigid definition of AI technologies. In line with the minimum requirements of legal clarity – as a basic principle of the rule of law and a

prerequisite for securing fundamental rights – lawmakers must exercise due care when defining the scope of any such AI law.

Outside the field of AI, the challenges of contact tracing, for example, have laid bare this dilemma. The dominant idea, according to which a trade-off must be made between the degree of precision of virus-mapping and the need to respect the quite demanding European data protection rules (GDPR), is entirely misleading. This essentially focuses only on the proportionality test and not (as Article 52 of the Charter, Article 23 of the GDPR and Article 15 of the e-Privacy Directive show) on the exceptional need to limit privacy. In other words, it is taken for granted that contact tracing applications will be effective in combating the virus, and it is consequently taken for granted that it will be necessary. This is a typical expression of "technology solutionism", to use Morozovian terms, according to which every problem must find an almost immediate technological solution. In this case, the solution should be the digital contact tracing system.

The protection of fundamental rights in the age of AI also requires a focus on which public and private instruments can be used by which actors to assess the impact of these technologies. In recent years, increasing attention has been placed on the development and implementation of AI technologies by the private sector. In particular – in line with existing international standards (notably the United National Guiding Principles on Business and Human Rights, or UNGPs) – businesses should put in place "a human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights" (Principles 15 and 17). This applies irrespective of size or sector and encompasses all businesses that work with artificial intelligence.

Impact assessments are an important tool not only for businesses but also for public administrations. They help mitigate the potential adverse impacts of any given activity on fundamental rights. Indeed, EU law requires particular forms of impact assessments in certain sectors, such as data protection impact assessments (DPIA) under the GDPR, or risk assessments under the Machinery Directive (these determine the health and safety requirements that apply to machinery).

However, although a large number of DPIA have already been carried out, they tend to evaluate the security levels of data more than the impact of any leakage. The assessments carried out so far have focused mainly on technical aspects, without addressing potential impacts on fundamental rights. According to some interviewees, such assessments – that is, of the impact on fundamental rights – are not necessary when an AI system does not appear

to have any negative effect. Clearly, an emphasis needs to be placed on the role of effective impact assessments in preventing adverse consequences. In the near future, this might be addressed by some "risk board" in organizations with high-risk AI systems. Such a tool is critical in ensuring that EU lawmakers give consideration to all fundamental rights. Furthermore, impact assessments should cover both private and public sectors and should be carried out before any AI-system is deployed. They should take various characteristics into account, such as the level of automation and complexity, as well as any potential harm.

In truth, assessing adverse impacts is not enough. In order to guarantee accountability, it is also necessary to ensure effective oversight (as required by the AI act) as well as enforcement mechanisms. This is necessary in light of the evolution of AI technologies, combined with their increasing rollout across various sectors. A variety of bodies are potential candidates for providing AI oversight from a fundamental rights perspective. These include specialist bodies established in specific sectors (for example, banking and data protection supervisory authorities). However, many of those interviewed from the private and public sectors are uncertain about just what responsibilities should be accorded to AI oversight bodies.

One option could be for national human rights institutions (NHRI) to play a primary role. These bodies are increasingly playing a leading role in monitoring and ensuring the effective implementation of international human rights standards at national levels. The non-judicial status and functions of NHRIs mirror their essential role in promoting and protecting human rights. These institutions promote awareness campaigns and can play an important role in preventing violations of rights. At the same time, they are also able to complement the remedies available through the courts, as the courts are not necessarily able to restore the status quo in every case; indeed, in some cases they are not easily accessed by minorities or vulnerable sectors of society.

Some scholars have focused attention on the possible discriminatory outcome of artificial intelligence. AI technologies are used to profile individuals, and to create clusters for classifying behaviors, relationships or other characteristics. This process is heavily influenced by biases introduced by AI developers as well as by data sources. Moreover, big data analysis raises the problem of inference between new data and information analysis. It is not always possible to verify the quality of the data itself, having been generated through correlations hidden within the logic of the AI system.

The obligation to respect the principle of non-discrimination reflects a broader need to ensure protection for human dignity as enshrined in Article 2 of the Maastricht Treaty (TEU), Article 10 of the Treaty of Rome (TFEU),

and Articles 1, 20 and 21 of the Charter. The TFEU, in particular, requires the Union to combat discrimination on a number of grounds, and the Charter demands equality before the law and non-discrimination. These sentiments and principles are also enshrined in various specific and detailed provisions contained in a number of EU directives.

Nonetheless, AI can also be used as a prominent tool for remedying discrimination. AI systems can be used to test for and detect discriminatory behavior, which can be encoded within datasets. However, according to a report by the European Union's Fundamental Rights Agency, very few interviewees mentioned the possibility of collecting such information concerning disadvantaged groups in order to detect potential discrimination. Given the lack of any in-depth analysis of potential discrimination within the actual use of AI systems, there has also been almost no analysis of the potential positive effects of using algorithms to make decisions fairer.

*The right balance between freedom and dignity*

Digitalized societies have to cope with a twofold ethical question. On the one hand, they should promote the widest use of artificial intelligence – through freedom of entrepreneurship, research, education – in order to deliver the concrete advantages of this technology to citizens, without discrimination; on the other hand, they should make sure that AI and deep learning-based technology remain respectful of human dignity and other individual rights, such as privacy, transparency and the fair use of data. Moreover, AI can play an important role in the protection of so-called "diffused rights", such as respect for the environment or the promotion of sustainable economic growth.

Laws and regulations are not independent from any given society's political goals concerning social environment and shared ethical values. In particular, it has been noted that the central goal of regulations and policies in the field of AI is to provide the best balance between ethical values, research, enterprise, and any number of practical advantages granted by AI. Clearly, circumstances may arise that create conflict between ethics and business, for example, while evident improvements in healthcare, living costs, customer experience, public administration efficiency and so on still must not undermine privacy, transparency, equality, non-discrimination, and – more generally – human dignity.

In order to reach a satisfying balance, it seems advisable to pay attention to the problem of over-regulation and complex regulation. In fact, policies based on over-regulation can, unintentionally, increase costs for companies

and hinder customer assistance. Simplification and – in certain cases – deregulation can actually play a substantial role in reducing the digital divide. This is a particularly important point for Italy: if Italy's level of digitalization and computational power is compared with those of other European countries, it is clear that the impact and efficacy of laws is crucial.

*AI and privacy*

In the field of privacy, many companies do not have the necessary dimension for a correct application of the GDPR. The complex regulations have produced the side effect of giving big players an advantage and of ruling out minor corporations from market competition. Furthermore, over-regulation may still fail to govern all the different specific sectors to which it applies and make it particularly difficult to remain updated, as opportunities naturally evolve, thanks also to AI development. Unjustified requirements produce only a formal and not substantial compliance; the agreement or consent given by customers is not really understood if the text is too long or difficult and therefore it does not represent real consent. Certain rules on the elimination of data are sometimes very hard to control, given the pervasiveness of AI and digital systems, and an excessive number of requirements may only distract authorities from the prosecution of major offences. Too many rules end up undermining the efficacy and certainty of the law and discouraging entrepreneurs, investors, and scientists. They do not make Europe the best environment for new initiatives and projects.

Many AI applications depend on the availability of high-quality data or use machine-learning or deep-learning techniques. One of the major challenges faced by banks and insurers when developing AI systems, for example, is the restricted or limited access to data that could improve such systems and create a better service for customers. This raises questions over how access to this data should be governed, as well as technical issues of interoperability and the standardization of data. In order to promote and support the development and uptake of AI, access to essential data must be facilitated.

The development of a common EU data strategy — one of the key focus areas of the Commission — will hopefully provide an opportunity for the EU to adopt a future-proof, innovation-friendly framework that supports data-driven business. It should enable the digital transformation of society while also ensuring appropriate protection for consumers. Greater availability of data could help insurers to improve risk monitoring and assessment, say, so as to offer a better customer experience and to improve fraud detection. Fraud, in insurance, is not just a problem for the relevant company but for all clients, because insurance is based on mutuality. The more data that is

available for the common good, the better the digital solutions and analytical models. Individuals should be able to allow access to their personal data to a much greater extent than is possible today. There should, for example, be practical solutions that would allow individuals to exercise control over their own data, with appropriate consideration given to the security of anything sensitive. Individuals should be able to grant other parties access to the data they generate.

Enhancing legislation on accessing, processing and sharing data will be important to promote innovation and competition. To foster AI-driven innovation, a first move in this direction is being made by the EU that is working on a Data Governance Act and on an Open Data Directive to establish trusted mechanisms and services for the re-use, sharing and pooling of data, necessary for AI high-quality systems. The insurance industry supports efforts to facilitate appropriate data-sharing based on voluntary agreements and on a true level playing field between different insurance providers. However, much will depend on the specific model and implementation chosen for any data-sharing framework. As for financial activities, it is necessary to ensure that the relative regulatory framework is digital-friendly, technologically neutral and sufficiently "future-proof" to be fit for the digital age. New technological opportunities and customer behavior inevitably give rise to new service concepts, and new service providers regularly enter the market. For such providers, regulatory requirements can often be less strict than those faced by the traditional financial services industry.

In general, a related key issue concerning the right to privacy is that of the problem of harmonization. Regulations in different areas of the industrialized world are significantly different. It is therefore crucial to respect the principle of "same activities, same risks, same rules" and strive for a true level playing field between all market players. It is clear that global discipline is lacking and that approaches to AI around the world differ enormously. While the EU took action in 2016, enacting the General Data Protection Regulation, China does not have the same concern; in the United States only California (with its California Consumer Privacy Act) has enacted regulations about personal data comparable to those of Europe. The right balance in the protection of both freedom and dignity (in other words, the right evaluation between opportunities and risks) is essential for scientific research and its application.
Take the healthcare sector for example: rules on AI must be simple and clear in how and when to apply them; they must also be sufficient in their scope

so as to regulate current and future potential uses without restricting innovation. Given the speed of change in technological capabilities, new approaches to existing guidance/regulations are needed to ensure that they remain robust as technology advances. In this sense, it could be useful to use risk-based approaches; sectorial Codes of Conducts on AI can help guide organizations where the law does not. In any case, any framework for AI use in healthcare must be advanced globally across jurisdictions so as to ensure the adoption of similar rules. AI has the potential to make a significant difference to healthcare. A broad range of techniques can be used to create AI solutions to carry out or augment tasks that have until now been completed by humans or that have not even been possible. AI can equip healthcare professionals (HCPs) with new tools to gain detailed data analyses delivered through the generation of novel scientific insights; it can allow HCPs to dedicate more time to irreplaceable human interactions. AI platforms also hold the potential to design optimal drug combinations that are effective and based on real experimental data rather than mechanistic assumptions or best guesses. AI platforms can perform predictive modelling using data mining and probability to forecast or estimate more granular, specific scenarios or outcomes, thereby predicting the benefits of combining specific drugs and therapies.

*The impact of AI on sanctions and judicial decisions*

The choice of the right balance between different interests and values has an impact also on liability, eventual sanctions, and the procedural side of the law.

If we consider, for example, the case of self-driven cars, the lawgiver has to decide between a system of only civil liability or a system of both civil and criminal liability. In the first hypothesis, it might be objected that a fundamental interest like people's safety demands a criminal law regulation; in the second hypothesis, it might be held that producers and retailers would be too disadvantaged. It seems that there is no technically right answer to the question, since any choice has to mirror the value balance, the political view, and the very concept of socioeconomic development that policy makers want to pursue.

However, at a more technical than ethical level, jurists suggest that it should always be possible to determine where blame lies (as is usually the case with military technology), so as to avoid an uncertain and dangerous diffused liability. A lack of certainty would undermine the solidity of entrepreneurship and business activities. "Diffused liability" in medical teams, for instance, has already proved to be one of the causes of defensive

medicine, with all the related harm to patients. In addition, relative liability may cover not only the intended use of a product, but also the foreseeable use and in some cases even the reasonably foreseeable misuse (Machinery Directive). Therefore, risk assessments on safety and security of AI products is paramount.

Moving to the procedural side of law, AI can certainly help implementation in the field of justice and, in particular, in the field of judicial decisions. In fact, technology is already changing the practice of law and reshaping judicial systems.

There are three main ways in which this is happening. First, artificial intelligence is helping lawyers in the analysis and drafting of legal documentation: for example, hundreds of law firms and courts around the world use machine-learning programs, like Luminance, to analyze documents; this vastly reduces the amount of time spent on document review. Second, artificial intelligence is replacing humans in carrying out some functions and activities. Third, courts are being receptive to the use of predictive algorithms in sentencing. For instance, judges utilize AI machines in order to understand whether defendants present a "risk" to the community that is high enough to render them ineligible for probation. In 2016, for example, the Wisconsin Supreme Court became one of the first courts in the United States to uphold the use of a predictive algorithm in sentencing: the judge had used COMPAS, a risk assessment algorithm, to conclude that the defendant was ineligible for probation.

As clearly shown by the last example, the practical use of artificial intelligence in the field of judicial decisions entails ethical challenges: do we really want to rely on automated processes in order to determine a person's detention? And also, how do we protect the privacy of people who are concerned by the decision? At the end of the day, the real question is the following: how do we balance human freedom and dignity with other rights and interests involved in the judicial system (e.g. right to trial within a reasonable amount of time, companies' trade secrets, etc.) when it comes to the use of AI programs?

There are no easy answers. From a methodological point of view, however, any assessment on this topic should take into consideration the risks related to the overuse of automated and/or standard decisions that leave out any human intervention.

Scholars have identified three main risk categories here. First, algorithms themselves may be flawed in some fundamental ways. An algorithm's objective may not be clear; and even if seemingly clear, an algorithm will

reflect the perspective and biases of its creators. Second, algorithms that produce life-altering judgments on everything from parole to credit are often "black boxes". This is true in two senses: 1) the calculations used to make a decision are usually inscrutable to the person affected by that decision; 2) companies often refuse to divulge information about them (because of trade secrets). As a consequence, people affected by these decisions are not able to defend themselves: they do not know the intimate basis of the decision. Third, algorithmic decision-making has been subjected to the "garbage in – garbage out" critique: any decision is as good or as bad as the data relied upon by the system. For example, when algorithms in the criminal justice system rely upon data that contains racial bias, the decision will inevitably reflect that racial bias.

In order to address all these challenges, it is time to implement what has been called the "technological due process": procedures that ensure fairness, accuracy and respect for human freedom and dignity by AI machines. These procedures should apply not only to the scoring algorithms themselves ("technology-driven rule making") but also to individual human decisions based on algorithmic predictions ("technology-driven adjudication"). These procedures might include ensuring people a set of rights: for example, a "right to explanation". Some scholars have argued that the European General Data Protection Regulation (articles 13, 14, 15 and 22.1) already recognizes a "right to explanation": a person affected by an algorithmic decision can ask for an explanation of it. This does not necessarily jeopardize the programmers' concern that the technical aspects of the system could not be kept secret: indeed, they may still be covered by a protected order requiring their confidentiality. This "transparent approach" also requires a "meaningful notice" system. In other words: individuals must be informed when a decision is driven by an algorithmic decision-maker.

A second step would be to require a system's software testing. Testing protocols should be run before a system's launch, during implementation and every time policies change. This testing would help identify and eliminate possible biases and software bugs.

Third, lawmakers should explore ways to allow the public to participate in the building of automated-decision systems: for example, they could establish information technology review boards that would provide opportunities for stakeholders and the public at large to comment on a system's design and testing.

Last, courts might consider refraining from relying completely on "automated justice", since programmers will probably never be able to include complex (and inexplicable) legal concepts like equity and fairness in any artificial intelligence machine.

The legislative proposal released by European institutions on April 21, 2021[1] defines High-Risk AI systems (i.e. those that have to be subject to specific safeguards as well as to binding ex-ante conformity assessments) on the basis of the intended purpose of AI systems. The EU legislative proposal stresses the importance not only of security and accuracy but of human oversight, as well as of transparency. Documentation, records and access to correct information must be available for users. These criteria serve the purpose of making AI systems accountable for the decisions they make (and for the underlying mechanisms upon which they rely). The very first requirement needed to make people and/or machines accountable is to demand that they provide transparent and verifiable explanations of the logic they follow to reach any given conclusion.

The most promising way to facilitate the gradual adaptation to AI use in the field of justice would be to embrace uncertainty through a policy of experimentation. However, it is difficult to experiment in a space where human liberties are at stake while simultaneously preserving the stability of the legal system and minimizing risks of disillusionment.

*The role of Europe for a value-based AI revolution*

The ongoing adoption of AI-driven technologies could also have an impact on the relation between citizenship and "consumership". Business models are not democratic models. The concerns raised by such issues are becoming ever more worrying as the high costs of AI-driven technologies become apparent. The new monopolies that have emerged ("big tech", or GAFAM – Google, Amazon, Facebook, Apple, Microsoft) have business models that are based on AI-driven technologies, and they serve to illustrate a shift in power relationships. Big tech has entire countries or supranational entities as interlocutors and yet GAFAM are built on corporate governance models that are not based on the rule of law. Their economic power is a threat to "democratic globalization". They turn citizens into consumers, and their regulation is quite a challenge.

GAFAM are powerful enough to resist rules, so the true challenge is an ethical one. The big players in artificial intelligence must be coerced into acting according to a strong ethical approach; their governance models need to be ethical by design, and they must abide by international and local laws. The internal rules governing big tech must somehow resist any risk of

[1] Regulation of the European Parliament and of the Council laying down harmonized AI rules (Artificial Intelligence Act) and amending certain EU legislative acts.

takeover by rogue states. Any violation of our rights facilitated by these AI superpowers is dangerous not only for commercial/business reasons, but also for how they can be manipulated to political ends.

Europe, on the basis of its economic role, its legal culture and its philosophical tradition, could design and interpret a new way to discipline AI. Europeans could create a model able to grant the protection of human dignity and democracy while encouraging innovation and economic freedom. Policies and discipline will be consequent to these ethical choices. To this end, a "Chart of Values" concerning the relationship between artificial intelligence, business and society could represent an ideal ethical source of inspiration for European lawyers, judges and policy makers.
The scientific community can help companies speed up the adoption of AI solutions by defining a comprehensive framework or methodology able to bridge the gap between the theory and principles of the Ethical Guidelines and the actions or processes that business can actually put in place in order to develop and deploy ethical AI solutions. If Europe succeeds in designing an attractive legal discipline (both value-oriented and technically valid) for trustworthy AI, it might also prove able to lead a process of harmonization among different normative approaches. Europe could encourage the adoption of a number of common principles – by the EU, the United States and China – which would then be followed everywhere AI and the digital revolution puts down roots. In particular, a common transatlantic organic base would help; to this end, perhaps it would be better to start from the due data process in procedural rights rather than substantive rights.

*The irreplaceability of human intelligence*

In conclusion, further thought must be given to the relationship between human dignity and artificial intelligence. One of the main tasks of anyone investigating the ethics of AI, today or tomorrow, is to recognize that human dignity consists (also) in the irreplaceability of the human being with machines. An ethical reflection about AI should begin with a consideration of the need to reclaim the structural difference between humans and machines. Underlying this difference, paradoxically, is intelligence.
Intelligence is a specific feature of the human being. This point has been defended in the contemporary ethical debate on AI with at least three arguments. The first argument claims that humans have intentionality – namely, the ability to represent something, to interiorize both physical objects and abstract objects (such as duties, future events, or values). On the contrary – for the time being at least – deep learning machines have no such

intentionality. The second argument claims that despite what the structure of the name "Artificial Intelligence" seems to suggest, this amazing and sophisticated technology is not a marriage between an engineering system and intelligence; it is rather a separation of the ability to solve a problem and the need to be intelligent in doing it or, in other words, a divorce between agency and intelligence. "Artificial intelligence" has detached the ability to perform a task from the need to be intelligent. In fact, even at its birth, AI was defined by Marvin Minsky as the "science of making machines do things that would require intelligence if done by men." The third argument points out that even when an algorithm-based system fulfills a particular task as well as (or better) than a person, the suggestion should not be made that this system is "intelligent". Though a deep learning system was able to beat the world champion of GO, for example, that system cannot competently deal with other tasks – tasks that prove easy for the man that was beaten. Technically, it has been said that it lacks "general intelligence": the system does not have the ability to solve problems or fulfill tasks for which it was not programmed.

The above considerations are relevant to those situations in which the object of a certain procedure is a single human individual. In such a case, AI cannot be treated as if it were simply an "intelligent machine", the nature and functioning of which can be understood in terms of digital algorithms and technologies. An important example of this, as seen above, is in criminal judgment: when a human being is judged, with all the complexities inherent in his personality and in his cultural and social context, human intelligence is called for. It is legitimate for a judge to rely on an algorithm that exploits a database, but that judge must nevertheless take into account a great number of other considerations. An algorithm's conclusion has limited value within a context of other human actions and life experiences and questions of conscience. Errors are always possible, but those who place too much faith in algorithms ignore something important: when algorithms are applied to complex problems, they are characterized by non-linearity; this, in fact, excludes the certainty of the prediction of any single case.

A similar argument can be applied to the medical treatment of a patient: within medical ethics it has already been recognized that the most correct choice must emerge from a global judgment, in which expert systems undoubtedly offer useful information, but do not provide an exclusive basis for the treatment of a specific person.

Often, technology is indeed meant to replace a person in certain activities. It can seem beneficial to rely on machines for certain activities of the mind, such as intelligence, reasoning, calculation, choice, decision-making. This

tendency can gradually lead to a perception of the human being as a particularly complex machine, profoundly affecting the concepts of human dignity, moral responsibility, human rights. In truth, these concepts are not applicable to machines; it is essential to show that, in a person, thoughts and actions occur on the basis of many characteristics – characteristics that technology does not contemplate and that it is unable to explain. In other words, emulation – whereby a machine can do a certain thing better than man – must not be confused with simulation, which suggests a mechanical and materialistic interpretation of human nature. The identity of result (the fulfilment of a particular task, done by a person or a machine) must not be confused with the identity of process.

It can be said that it is not artificial intelligence that presents a threat to human dignity, but rather the human understanding of AI.

# 3. Algorethics: technical and ethical limits of AI in corporate applications

The ethical issues that relate to artificial intelligence can be structured following a framework that divides them according to what is involved: physiology, psychology, sociology, philosophy, and science. Regardless of the field engaged when using artificial intelligence, there are inevitably general principles and technical limitations that must be considered when making decisions.

In 2019, the European Union defined four principles and eight requirements, which form the basis for the "Ethics guidelines for trustworthy AI". The four principles are: respect for human autonomy, prevention of harm, fairness, and explicability. The eight requirements are: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, accountability.

In the present discussion of "algorethics", we will use the taxonomy defined by the European Union, starting with the "respect for human autonomy" principle. According to this principle, "AI systems should not unjustifiably subordinate, coerce, deceive, manipulate, condition or herd humans. Instead, they should be designed to augment, complement and empower human cognitive, social and cultural skills." Like all technologies, artificial intelligence can in fact be used to protect fundamental human rights or to threaten them.

Take, for example, the issue of personal data. On the one hand, artificial intelligence offers a key tool with which to track down one's own data and to establish whether it is being used correctly. On the other hand, the same technology can contribute to the misuse of one's data and can represent a threat to one's privacy.

Furthermore, the concept of human oversight is absolutely crucial in the effort to ensure human autonomy. Human oversight helps to ensure that an AI system does not compromise human autonomy nor otherwise cause any negative effects. Supervision can be achieved through governance mechanisms such as a "human-in-the-loop" or "human-in-command" approach.

The second principle proposed by the European Union is the prevention of harm. This concept encompasses the protection of human dignity and mental and physical integrity. When we refer to this topic, we typically speak of the information asymmetry between government and citizen or between employer and employee, which generates an imbalance in power and possibilities. While surely important, this is not the only relevant phenomenon: the growing diffusion of connectivity and social networks, for example, enables the mass sharing of information. It is not uncommon for fake news to be launched for the sole purpose of damaging someone's reputation or in hopes of receiving some advantage to the detriment of someone else. This type of false information, moreover, tends to target an audience that is already polarized on certain positions. In the end, divisions only increase, and positions become even more polarized. Deep-fake videos and deep-fake photos are some of the worst sorts of fake news. These often have the sole objective of constructing false multimedia content with a sexual background, therefore causing psychological damage to the victim.

The third principle proposed by the European Union is in support of fairness and to block bias. The very definition of this principle is complex, and an attempt has been made to approach it in two dimensions. The substantive dimension, which implies a commitment to ensuring the equal and just distribution of both benefits and costs, aims to protect individuals and groups from unfair bias, discrimination and stigmatization. The procedural dimension focuses on the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them.
From a practical point of view, however, these definitions and dimensions are not sufficient. Let us start with a consideration of "fairness". When the debate over autonomous machines took hold, the public was invited to express an opinion. Generally, when people were asked if they were happy for autonomous cars to roam the streets, the answer was positive. When asked who should define the ethical paradigms underlying autonomous car decisions, however, neither the government, nor car manufacturers, nor universities were deemed reliable decision-makers. The Massachusetts Institute of Technology (MIT) therefore decided to propose a system – The Moral Machine – that would confront the population with ethical dilemmas and ask them to express a preference on each. The result was surprising: different cultures and geographies entertained very different ideas about what was good and what was fair.

It is clear, therefore, that fairness is difficult to define. Indeed, it often depends on the given context, and this has practical implications. Artificial

intelligence does not tend to work well or behave correctly in all contexts; it works best in narrower contexts (AI is often highly specialized). So how to make autonomous cars viable?
When it comes to AI, it is the goal of the machine and the objective of the algorithm that must be built correctly. The goal of flexibility – the system's ability to change and adapt – can in fact collide with the goal of becoming more and more efficient in performing a specific task.

Bias is another difficult but important question. From a practical point of view, bias can be translated into a prejudice in favor or against a person, an object or a position that somehow affects the outcome of an algorithm, rendering it unfair. Artificial intelligence systems can suffer from the inclusion of unintentional historical biases, incompleteness, and mismanagement patterns that propagate such biases and can lead to the exacerbation of prejudice and marginalization. Bias removal is an open issue to which there is no general technical solution. Indeed, for particularly large AI-based "black boxes" (artificial intelligence systems whose operations are not visible to users), detecting and removing bias has been shown to be nearly impossible.
Nevertheless, if we want companies to benefit from AI solutions, an effort should be made to resolve this issue. Automated systems must be developed that – with human supervision – analyze the huge datasets used to train AI systems and highlight any possible biases, based on the asymmetry of the set's various features. This could become a standard procedure: not only would it help detect ethical bias, but it would create much more robust and reliable algorithms.

The fourth and last principle proposed by the European Union is the issue of explicability. This basically tasks intelligent systems with being transparent: any outcomes and choices made by AI must be verifiable and justifiable.

Regardless of the shape AI ethical principles may take (different countries or regions may set different principles), in order to be able to monitor and check the actual compliance of AI systems, it should be possible, at least in principle, to attribute any single decision to its root cause. This does not mean that every single decision will have to be traced back, but, if deemed necessary, it should be possible to do so.

The resilience and the strength of any agreed set of AI ethical principles and their effective chances to be fulfilled depends precisely on this possibility. If

the compliance of automated decisions cannot be verified, the AI ethical debate will become void and meaningless. A decision that can be retraced to its root cause is a decision that can be explained. And an automated decision that can be explained is also a decision that can be contested or upheld on the basis of logic or of the personal data used.

Sadly, not all AI systems provide explainable decisions. While the idea is that "explainability" cannot and shall not be taken for granted, rather the contrary is true: the working mechanisms (the algorithms) of AI systems are often created directly from the set of data used to train the system. This means that – most often – the algorithms that underlie the AI systems' functioning are not known ex ante; most of the time, they are not even known ex post.

As the amount of data processed by an AI system increases, so does the logic with which it reaches conclusions. This is particularly true for AI systems which continue to learn after being put into service. Because of this "dynamic learning" (or "machine learning"), AI mechanisms often turn into black boxes that cannot be interpreted or understood, even by experts, or by the very engineers that created them. To tackle this issue, different potential measures could be considered at the policy level. Among these, for example: 1) publishing the algorithm (here there would be interplay with issues of commercial secrets); 2) utilizing decomposable systems only (i.e. structuring the algorithm in stages, so that results are interpretable); 3) adopting a proxy model (i.e. using a second interpretable model, which approximates the complex black box system); 4) visualization mapping (illustrating how strongly different input features affect output); 5) applying counterfactuals (i.e. identifying which input data, when changed, produces a different output).

If algorithms are to play an ever-increasing role in our daily life, they need to be trustworthy. People do not trust what they do not understand, especially what even experts and scientists cannot explain. Trust can only be gained through openness, transparency, understanding and the possibility of changing whenever something has not worked as expected. Transparency and openness are crucial elements in settling potential disputes; they are key to remedying any "unfair" decisions. Indeed, these represent fundamental principles for the rule of law in democratic states.

Opaqueness (or the creation of those "black boxes") is one of the weakest points of AI systems. Indeed, all automated decision systems suffer from this

and the problem needs to be coped with quickly and effectively. Some progress here is evident in the emergence of the so-called "explainable AI" field. The more AI systems gain in importance, the more crucial does it become that their working mechanisms can be explained.
The "explainable AI" field represents a fundamental stream of development for AI systems. It is increasingly clear that algorithms must be rendered more amenable to ex ante and ex post inspection and to human oversight. There is a need for human understanding and interpretability in algorithmic design. Indeed, this need of "human" understanding is not that different from what regulators (i.e. privacy and data protection authorities) and antitrust bodies advocate for search engines or for targeted commercial advertising practices. Algorithms may be considered a "business secret" but the logic they follow to get to a certain result has to be known and transparent.
In AI systems, human understanding is needed and necessary, just as it is in any other technological field or social matter. When human understanding is lacking, human beings risk losing control. And it is crucial that humans maintain control over the consequences of any machine's actions. That means, indeed, maintaining control over the machines themselves.
AI must not be developed in terms of autonomous evolutionary systems; AI must be developed rather as a tool – a tool that aids the human being, increasing his or her cognitive abilities. This means making the algorithmic criteria underlying the machine and its functioning (including self-learning) transparent and accessible to human cognition.

The smartest way to deal with this issue today is to approach it with a practical scenario-based method called "counterfactual analysis". By way of example: suppose that there is an intelligent system capable of deciding whether a person deserves to receive a bank loan or not. To the extent that the algorithm expresses itself negatively, understanding why may be extremely difficult; but it is possible to interrogate the algorithm by manually modifying certain conditions (say the person's income) to see how the outcome changes.
The development of a standardized or certified process might provide some level of clarity. Such an automated heuristic counterfactual analysis would shed light on those black boxes: algorithms could undergo the analysis to be "certified" before release and/or the analysis could be used to check on the algorithm throughout its lifetime.
The issue of explicability is not only technical: it also has a cultural element that touches on the relationship between basic and applied sciences. From a practical point of view, in fact, approaching some machine learning problems has become extremely simple and requires only basic scripting

skills. But fully understanding the processes and principles underlying the pure practice remains extremely complicated. It is therefore appropriate to differentiate the "data practitioner" from the "data scientist": the first figure is able to use the tools and algorithms of machine learning to solve problems; the second is able to diagnose the reasons underlying a mismatch between expectations and outcomes and understands the science behind the algorithms. Let us be clear: the existence of data practitioners is extremely valuable. Indeed, the big questions facing scientists today are sometimes posed by practical problems. Often in AI, it is the general principles that are abstracted from applications, rather than vice versa (see Fourier analysis or cryptography, for example).

Nevertheless, it is also clear that an over-regulation of AI algorithms would pose serious risks, perhaps hampering the full exploitation of its benefits, especially within those economic contexts at an early stage of AI adoption. This risk can be particularly high in those countries or sectors where companies are not big enough to manage complex regulations. (This is the case in Italy, for instance.) In this sense, an evaluation of the GDPR adoption experience by country and by market sector – with specific reference to the economic impact of the law and to related difficulties – could help to define rules that foster the ethical development and adoption of a technology, while preventing its biased misuse. Europe – and Italy in particular – is still at an early stage of expansion in this area, so clear rules would be particularly helpful.

Furthermore, in order to avoid the abovementioned risks of AI over-regulation, and given the close relationship between data and algorithms, a progressive regulatory approach is desirable. This could be developed starting from the current GDPR in terms of consumer protection. Such an approach would allow companies to move in continuity with what they have already implemented. They could thus avoid the introduction of complex new regulatory elements, which are more suited to contexts where the level of AI adoption is high.

# 4. AI: education and information for digital citizenship

The technological and scientific culture of industrialized countries is fertile ground for welfare and economic growth. AI has a prominent and increasingly important role to play here, even if technological development is not sufficiently widespread nor encouraged by institutions and the leading classes. Moreover, AI has generally improved the quality of life of citizens around the world. However, AI has also generated – even if not intentionally – new forms of discrimination, because it is not easy to keep up with this kind of technology. Older people in particular are often unable to use AI and are therefore marginalized.

The fact that such marginalization is not intentional does not free us from considering the problem from an ethical point of view.

Artificial intelligence is not only too complicated for much of the world's population, but it also confounds many CEOs, politicians, academics, judges, and so on. The small elite who actually understand AI well can exploit this widespread ignorance – even on the part of the ruling class – to misuse the power AI provides. The so-called "AI comprehension tools" (that is, the ability to understand AI and its implications) must therefore be radically improved in the near future. Indeed, general education on artificial intelligence is called for, and comprehensible information on AI should be made available, far and wide. Such an AI education may follow at least three paths.

*The education system*

The first path is the enhancement of computer science and data science courses in schools and universities for the new generations. At the university level it is important to create AI scientists, not simply big data practitioners. In various European academic institutes some math and computer science courses have also been reformed in order to include data science and AI. The connection between math and AI is fundamental, of course, and theoretical research in mathematics has produced important discoveries that later have become tools for artificial intelligence and its applications (see cryptography for instance).

Since teaching is valuable and up-to-date only if well-grounded in academic research, education is strictly related to the in-depth study of such areas as informatics, business applications, and even ethical issues. Research inevitably provides the nourishment for teaching.

Education systems the world over struggle to provide enough AI researchers and experts. This talent shortage in Italy, in particular, is worsened by the aggressive hiring policies of tech giants and large investors, especially in the US. The appetite for AI experts is very high but the market is very small. Given the salaries offered and the cutting-edge nature of the environment, it is only natural that AI scientists, analysts and researchers prefer to be hired by Silicon Valley companies. This may create a worrying imbalance in the knowledge about and assessment of this transformative discipline. Therefore, in the near future, Europeans and Italians in particular need to face the risk of low-quality academic research: it is worth underwriting educational policies at an international level to combat such imbalances.

*Lifelong education*

A second path consists in specialized courses on AI and its implications aimed at business leaders, public administrators, and employees in a variety of sectors: in fact, while cultivated adults are likely to be acquainted with poetry or history, say, they often lack a scientific and technological culture. Indeed, a certain resistance to the language of science and mathematics is widespread.
Sometimes, the underuse of AI is not caused by legal constraints, but by prejudice or fear, or by what we might call an "allergy" to AI applications: the outcome is an unjustified self-constraint. For instance, an algorithm-based machine that, processing a set of data, produces useful information for oncological diagnoses can be extremely valuable for medical practitioners, so its underuse is ethically regrettable. Naturally, the employment of such a tool should only be allowed in medicine under the supervision of professionals, which also means that doctors and technicians need to fully understand the technology.

Even professionals who have studied AI often lack the multidisciplinary competences and profiles requested by the most innovative corporations and public administrations: in fact, the further revolutionary development of AI in the near future is expected to be driven primarily by the robust growth of interrelations between technologies (even more than the birth of new technologies). This new expansion will likely be based on cloud

infrastructure. Therefore, it is crucial that professionals be equipped with these new skills and understanding, with a multidisciplinary education and with ethical awareness.

*Information and popularization initiatives*

Since our life is dominated by the presence of new technologies, there is also a need for a third path. This should consist in general and continuous information about technology. By popularizing AI, almost everyone can be reached, leaving nobody behind. The popularization of technology and, more in general, of science must be accurate and accessible. Accessibility must not preclude accuracy, so pseudo-science is to be avoided at all costs. Scientific knowledge depends on specific standards of method: in accordance with Popper's approach, for example, pseudo-scientific theories cannot be classified as science because they cannot be falsified.

Within the popularization of science and technology, contemporary societies also have the specific task of creating an AI culture. The contemporary ethical debate on artificial intelligence, in fact, highlights the importance of overcoming the digital divide. AI literacy can even quash related forms of discrimination, and contribute to the development of a newly unified citizenship. This "literacy 2.0" could be similar in its aims to the process that took place in Italy over the last century, whereby people were educated to a correct use of the Italian language. Through the synergic action of policy makers, business companies, universities and cultural institutions – as well as mass media – the general public can gain significant scientific understanding.
In such a scenario, the role of companies is more crucial than ever. Firms should continue to promote experimentation in different fields and boost public debate on AI, promoting an internal as well as external paradigm shift. To start with, they should explain what is under the surface of AI technology. Such an explanation would go a long way towards promoting artificial intelligence and its ethical use, since it is only through profound understanding that widespread feelings of fear can be eased. As said above, it is ignorance and uncertainty that stimulate fear.

In a big data and digitalization society, people should be aware that formulating a question correctly is key to getting the best response from a machine. Paradoxically enough, AI systems are only as reliable as the data they receive: to look for the relevant information and interpret the correct answer, a machine needs a human to insert the right data and ask the right

questions. Many ethicists note that while we certainly need to promote the ability to live in a digitalized world, we must nevertheless take care not to lose our own ability to reason, to remember, and to solve problems. Humans must not come to depend too much on technology, as it may weaken our thinking and our self-determination.
Teaching the general public how to interface with AI and how to understand the power of data is a necessary condition for a competent and informed citizenship. A greater awareness among people will help to slow down the circulation of "deep fakes" and other forms of manufactured or manipulated content, thereby limiting their negative impact.

In this endeavor, AI itself can be leveraged to design and develop new and more accessible interfaces. These could enable more people to access systems and services, rendering them more autonomous and even perhaps improving their quality of life. Research in this direction should be encouraged, also in consideration of the European Commission's "Twin Transition", which invests in digital activities partnered with sustainability. The idea here is to use artificial intelligence to improve access to services of all kinds, regardless of an individual's physical abilities, age, technical skills, education, or any other characteristics. This represents an important step towards the democratization of resources.

*AI and the post-Covid world*

At last, it is worth noting that a reflection on the use of artificial intelligence during a pandemic – given not only the health crisis but the crisis in education suffered around the world – reveals something about our hopefully soon-to-be post-Covid world. During the pandemic, different countries combated the virus in different ways, highlighting, among other things, different approaches to school systems and to higher education. To get out of this global crisis, the integration of different national educational systems might help. A collaborative and coordinated approach could create a new global community.
In this perspective, the diffusion of AI and big data technologies can provide the glue for a new process of global integration, if ethical values are also linked to the democratization of global educational processes. From primary school to adult education, efforts should be made to avoid the creation of barriers between generations. Teacher training is crucial, of course, as teachers themselves must overcome any fear of AI before convincing their pupils of its benefits. Programs about the appropriate use of AI and big data – programs that stress ethical issues and warns against new disparities or

social fragmentation – should be shared among countries, to train not only science teachers but also those who work in the humanities.

Improving the world's education systems, and enlightening the general public about artificial intelligence in particular, can become the tool we need to spread new values over the global social structure. With a wide consensus on the founding values of the new post-Covid society, progress can be made towards an ethical AI for all.

# 5. The paradigm shift of business models enabled by digital technologies and AI applications

Artificial intelligence – of the many new technologies available today – is perhaps the one with the greatest impact on corporate digital transformations. When we think of a transformation within a company, we too often think only of the efficiency of processes or of the technical evolution of products and services. But by leveraging the opportunities given by technology, it is instead possible to change the way a company is positioned within its reference market and within its production chain. The very way a company generates revenue and delivers value can be transformed, thereby revolutionizing an entire industry.

One good example of how an industry underwent a significant change leveraged by technology is provided by the banking industry: after the introduction of the second payment services directive (PSD2), AI-based products and services were adopted widely. Open Banking has indeed acted as an enabler, allowing for the integration of services provided by Third Party Providers with those traditionally offered by banks.

Traditionally, companies had to define their position in three strategic dimensions: 1) customer intimacy (who); 2) operational excellence (how), and 3) product leadership (what). By way of example, when it comes to mass production, automation and standardization become core values, but customization and personalized customer experiences are left behind. Similarly, when it comes to building customer-centric products or services, operations tend to scale poorly.

Today, this principle has changed radically. Thanks to major technological paradigm shifts, the very approach to business is changing: thanks to AI-driven approaches, it is now possible to excel in all three dimensions at the same time.

The dimension that has perhaps benefited most from data-driven approaches is customer intimacy. In fact, the AI paradigm is rapidly changing the fundamental processes and frameworks used in data processing, data usage and data storage. Previously unforeseen applications are being enabled. Advanced AI-based analytics have drastically improved the process of transforming data into actionable information for businesses and industries around the world. Today, companies are able to learn more,

better and faster, triggering more personalized products and services, and thereby more quickly and efficiently meeting their customers' needs and desires.
Without AI technology, it would not be possible to classify your customers by behavior, identify micro-clusters in your customer base, or build mass-customization schemes. Thanks to this technology, it is no longer the customer who must adapt to the product, but rather the product that can automatically modify itself to improve the customer's experience.

From an operations point of view, artificial intelligence has always been a highly desirable tool. Initially, very few companies could afford to make the huge investments needed to design AI that can perform expensive tasks based on state-of-the-art machine-learning algorithms fed by enormous amounts data. This has historically been the greatest barrier to entry: only a few players could ever afford the significant IT infrastructure investments, not to mention the costs of maintaining a skilled workforce capable of managing this complex infrastructure.
Great opportunities are arising today from infrastructure-as-a-service providers and cloud providers in general: when it comes to computation, there are a number of newly available tools that can effectively convert a CAPEX-intensive prerequisite (that is, fixed costs) into pure OPEX (or variable costs), thus democratizing AI market access. For example, today, when purchasing cloud computing power – or pay-per-second virtual machine instances – costs are scalable, and only actually paid for when consumed. AI is thus rapidly becoming affordable for any company, and the growth of the data economy will only accelerate this trend, further enabling competition among digital business models based on customer data.
In addition, even the smallest players are aided by high quality connectivity and network technologies (optic fiber, 5G and beyond), whereas edge-based computational schemes (that work as closely as possible to a given application) help provide higher standards of security and privacy.
As far as products are concerned, the fact that digital representations of products and services can be created opens up the possibility of accessing global markets and differentiating any given offer according to the audience. There are clear advantages to being able to deal with a constantly growing number of intangible resources (that is, data).

Intelligent data processing can connect demand and offer a range of meaningful services such as transportation, accommodation, or food, thereby creating powerful digital markets. AI increases a platform's capabilities in matching demand and supply, learning from its customer's

needs as well as its supplier's offerings, to create perfect and timely matches. Digital marketplaces even have the power to shift wealth and GDP: if a Parisian calls an Uber, for example, he or she avoids any need for local goods and actually contributes to America's GDP.

Finally, from the point of view of business risk, digitalization is making an enormous contribution. The previous reaction-based approach has been replaced by a more competitive, proactive one: it is now often possible to foresee the course of events and thus to prevent problems before they happen. This allows for great advantages in terms of efficiency, and represents a real paradigm shift.

In manufacturing, for instance, AI has a remarkable track record in preventing problems with machinery. Its effect on the managing of routine activities such as process control (formerly assigned to humans) has also been impressive. The adoption of AI applications in manufacturing allows workers to focus on activities where their value added is greater.
In the oil and gas sector, to consider another example, the reduction of business risk is achieved through the application of predictive models. The extensive use of AI technologies reduces mining uncertainties, minimizes exploration risk, simplifies and dematerializes internal processes, improves the integrity of assets, and optimizes production and energy efficiency. It improves the safety of plants and people by evaluating potentially dangerous situations and identifying potential risks.

The construction of a roadmap that points to the inclusion of technology within a company is complex. It requires strategic direction and vertical expertise on a great number of components. In building a coherent implementation strategy, therefore, it is important to consider the strategic paradoxes involved.

In particular, when considering artificial intelligence as a competitive factor, it is necessary to take into account four different dilemmas:
First: productivity vs strategy. If AI is considered a tool capable of enhancing the business model of a company – and not only of making internal processes more efficient – then the net result on employment may actually be positive.
Second: automation vs augmentation. Technology is typically a revolution to be embraced, and the use to which it is put can well augment the decision-making ability and performance of workers.

Third: artificial vs human intelligence. To date, human cognitive functions remain more complex than those that computers can simulate. Creativity and implicit intuition are examples of this.
Fourth: data vs algorithm. While the algorithm evolves, data remains a constant. Data can assume increasing importance over time and its governance cannot remain in the hands of a few.

## *Steering Committee*

**International Conference "Ethics and Artificial Intelligence"**
***(Venice, September 24-25, 2021)***

**Members:**

- **Barbara Caputo**, *Chairwoman*
  Professor of Informatics Engineering, Politecnico di Torino; Coordinator, ELLIS - European Laboratory on Learning and Intelligent Systems, Turin

- **Evandro Agazzi**, Director, Interdisciplinary Center for Bioethics, Universidad Panamericana, Mexico City
- **Carlo Bagnoli**, Professor of Strategy Innovation, Ca' Foscari University, Venice
- **Paolo Benanti**, Professor of Moral Theology, Bioethics and Neuroethics, The Pontifical Gregorian University, Rome
- **Patrizio Bianchi**, Minister of Education; Scientific Director, IFDA - International Foundation for Big Data and Artificial Intelligence for Human Development; Professor of Applied Economics, University of Ferrara; Member, Bologna Institute Academy of Sciences, Bologna
- **Stefano Canestrari**, Professor of Criminal Law, Alma Mater Studiorum, University of Bologna; Member, Bologna Institute Academy of Sciences, Bologna
- **Pierluigi Contucci**, Professor of Mathematical Physics, Alma Mater Studiorum, University of Bologna; Member, Bologna Institute Academy of Sciences, Bologna
- **Carlo D'Asaro Biondo**, Executive Vice President, Division Partnership, Alliances & TIM Cloud Project, TIM, Rome
- **Alessio Figalli**, Fields Medal 2018, Professor of Mathematics and FIM Director, ETH Swiss Federal Institute of Technology, Zurich
- **Daniel Innerarity**, Professor of Political Philosophy, University of the Basque Country; Chair, AI & Democracy, European University Institute, Florence

- **Laura Li Puma**, Head of Artificial Intelligence Lab, Intesa Sanpaolo Innovation Center, Turin
- **Nikolaus von und zu Liechtenstein**, Chairman, Institute D'Etudes Politiques, Bendern
- **André Loesekrug-Pietri**, Chairman, JEDI – Joint European Disruptive Initiative, Berlin, Brussels, Paris, Rome, Tallinn, Warsaw
- **Stefano Lucchini**, Chief Institutional Affairs and External Communication Officer, Intesa Sanpaolo, Milan
- **Benigno Pendás García**, Professor of Political Science; Vice President, Real Academia de Ciencias Morales y Políticas, Madrid
- **Angelo Maria Petroni**, Secretary General, Aspen Institute Italia; Professor of Logic and Philosophy of Science, Sapienza Università di Roma; Member, Bologna Institute Academy of Sciences, Rome
- **Elise Poillot**, Professor of Civil Law, University of Luxembourg
- **Oreste Pollicino**, Professor of Constitutional Law, Bocconi University; Member of the Executive Board, European Agency for Fundamental Rights, Co-founder and Partner, DigitalMediaLaws Counseling, Milan
- **Massimo Proto**, Professor of Private Law, Link Campus University, Rome
- **Andrea Rigoni**, Partner, Global Government and Public Services Industry Cyber Leader, Deloitte Risk Advisory, Milan
- **Eric Salobir**, Executive Committee Chair, Human Technology Foundation, Paris
- **Pedro Schwartz Girón,** Fundación Rafael del Pino Professor, Department of Economics, Camilo José Cela University, Madrid
- **Mariarosaria Taddeo**, Associate Professor and Senior Research Fellow, Deputy Director, Digital Ethics Lab, Oxford Internet Institute, University of Oxford; DSTL Ethics Fellow, Alan Turing Institute, London; President, Noovle-TIM Group, Milan
- **Tomasz Trafny**, Head of Science and Faith Development, Pontifical Council of Culture, Vatican City

*with the assistance of:*

- **Silvia Castagna**, Head of Institutional Relations and Large Customers, BVA Doxa Istituto Ricerche Statistiche e Analisi Opinione Pubblica, Rome
- **Stefano Colloca**, Associate Professor, Philosophy of Law, University of Pavia
- **Lorenzo Iannarilli**, Honorary Fellow, Marketing & Digital Communication, LUMSA University, Rome
- **Giuseppe Roberto Marseglia**, Aspen Junior Fellow, Aspen Institute Italia; Adjunct Professor, Business Modeling, University Ca' Foscari, Venice
- **Michele Palermo**, Head of Industry and Macroeconomic Scenarios, TIM, Rome

**Coordinator of the Steering Committee**:

- **Francesco Leopardi Dittaiuti,** Senior Advisor, Aspen Junior Fellows, Advanced Education and Research, Aspen Institute Italia, Rome

# Annex 2

## *Business Users Advisory Board*

**International Conference "Ethics and Artificial Intelligence"**
***(Venice, September 24-25, 2021)***

**Chairman:**

- **Angelo Maria Petroni**, Secretary General, Aspen Institute Italia; Professor of Logic and Philosophy of Science, Sapienza Università di Roma; Member, Bologna Institute Academy of Sciences, Rome

**Co-Chairs:**

- **Carlo Nardello**, Executive Vice President, Chief Strategy, Business Development & Transformation Office, TIM, Rome
- **Alessandro Picardi,** Executive Vice President, Chief Public Affairs Officer, TIM, Rome
- **Mariarosaria Taddeo**, Associate Professor and Senior Research Fellow, Deputy Director, Digital Ethics Lab, Oxford Internet Institute, University of Oxford; DSTL Ethics Fellow, Alan Turing Institute, London; President, Noovle-TIM Group, Milan

**Members:**

**A2A**

**ABI - Associazione Bancaria Italiana**

**Acea**

**Assicurazioni Generali**

**Compagnia di San Paolo**

**Confagricoltura**

**Eni**

**Farmindustria - Associazione delle Imprese del Farmaco**

**Hera**

**Leonardo**

**Nexi**

**Pirelli**

**RAI**

**TIM**

**Unipol Gruppo Finanziario**

**with the assistance of:**

- **Silvia Castagna**, Head of Institutional Relations and Large Customers, BVA Doxa Istituto Ricerche Statistiche e Analisi Opinione Pubblica, Rome
- **Stefano Colloca**, Associate Professor, Philosophy of Law, University of Pavia
- **Lorenzo Iannarilli**, Honorary Fellow, Marketing & Digital Communication, LUMSA University, Rome
- **Giuseppe Roberto Marseglia**, Aspen Junior Fellow, Aspen Institute Italia; Adjunct Professor, Business Modeling, University Ca' Foscari, Venice

**Coordinators of the Advisory Board:**

- **Francesco Leopardi Dittaiuti,** Senior Advisor, Aspen Junior Fellows, Advanced Education and Research, Aspen Institute Italia, Rome
- **Michele Palermo**, Head, Industry and Macroeconomic Scenarios, TIM, Rome
- **Guido Giacomo Ponte**, Head, Economic Studies, TIM, Rome

# Annex 3

## *Conference Honorary Committee*

**International Conference "Ethics and Artificial Intelligence"**
***(Venice, September 24-25, 2021)***

**Chairwoman:**

- **Barbara Caputo**, Professor of Informatics Engineering, Politecnico di Torino; Coordinator, Ellis - European Laboratory on Learning and Intelligent Systems, Turin

**Members:**

- **Jean-Luc Allavena**, Chairman, Institut Aspen France, Paris
- **Carmela Decaro Bonella**, Former Professor of Public Law, Rome
- **Giusella Finocchiaro**, President, Fondazione del Monte di Bologna e Ravenna; Professor of Private Law and Internet Law, University of Bologna
- **Stormy-Annika Mildner**, Chairwoman of the Management Board and Executive Director, Aspen Institute Germany, Berlin