A EUROPEAN APPROACH TO ARTIFICIAL INTELLIGENCE

A POLICY PERSPECTIVE

CONTENT

EXECUTIVE SUMMARY

This report features in the EIT Digital policy perspective report series and is the result of a combined effort from five EIT KICs (EIT Manufacturing, EIT Urban Mobility, EIT Health, EIT Climate-KIC, and EIT Digital as coordinator). The EIT Knowledge and Innovation Communities represent the largest innovation ecosystem in Europe with more than 1,500 organisations from business, research, innovation, and higher education from all across Europe. By tapping into the vast innovation and application knowledge, the report identifies both general and sector specific concerns and opportunities for the further deployment of AI in Europe.

The report provides business and policy decision makers with a scenario-based impact assessment instrument. The scenarios off-set generic across the board application of policies and regulation against domain specific policies and regulation; and on the other hand, the level of regulation varying from soft to firm.

Next to the generic concerns and opportunities shared by all application domains, the report also identifies the key concerns and opportunities regarding AI in the domains of manufacturing, urban mobility, health, and climate and energy. This together with the core policy levels available to mitigate the concerns and embark on the opportunities.

Next to providing a pragmatic instrument for AI policy development, the main recommendations are:

- To ensure effective policy in the area of AI it is necessary to take context (sectors of application) into account.

- Policies regarding application of AI on personal data should be allowed to differ from policies regarding application of AI on machine data, especially in certain application sectors.

- General regulation or policy measures can be considered in relation to algorithm transparency and explainability.

- Regulation should be adaptable and flexible, whilst minimising and mitigating risks and ensuring human rights and European values.

INTRODUCTION

While a clear cut definition of Artificial Intelligence (AI) would be the building block for its regulatory and governance framework, there is not yet a widely accepted definition of what AI is (Buiten, 2019; Scherer, 2016). Definitions focussing on intelligence are often circular in that defining what level of intelligence is needed to qualify as 'artificial intelligence' remains subjective and situational[1]. Pragmatic ostensive definitions simply group under the AI labels a wide array of technologies, applications, and uses (i.e., robots, cobots, machine learning and deep learning, computer visions, speech recognition, diagnostic systems, autonomous manufacturing machines, self-driving cars, and dozens more). The European Commission, for instance, in its 2018 Communication (European Commission, 2018, p. 2) defines AI as referring to: "systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)". Broad AI definitions cover several technologies, including machine learning (algorithms whose performance improve as they are exposed to more data over time), deep learning, predictive analytics, computer vision and natural language processing (see for instance Annoni et al., 2018). Following Buiten (2019), in this report we focus on algorithms and the data they use, for two reasons. First, algorithms and the underlying data is what most, if not all, the technologies and applications grouped under AI have in common. Second, they are at the same time the source of the new opportunities and risks that are currently the object of ethical and policy debates.

The European Commission expects AI to significantly improve the lives of EU citizens and bring major benefits to society and economy through better healthcare, more efficient public administration, safer transport, a more competitive industry, and sustainable farming (European Commission, 2018b). Various AI applications are increasingly utilised in a diverse ranges of domains, including monitoring traffic congestion, employee hiring, metering smart energy grids, preforming credit checks, assessing recidivism risks when deciding sentences, and many more (Teich, 2019; Teich & Tirias Research, 2018). They can guide decisions that up until recently only extensively trained humans could perform. This has allowed for remarkable improvements such as the ability to analyse medical images in radiology (Dreyer& Allen 2018) and to detect potentially cancerous cells (Al-shamasneh & Obaidellah, 2017) or predict where and when the next big earthquake may strike (Fuller & Metz, 2018). The potential for AI to overcome limitations of humans when dealing with computationally intensive tasks and augment intellectual and perhaps even creative capabilities, opens up new application domains, with impacts on productivity and performance (Dwivedi et al, 2019). While the applications are already in place, a recent EIT Health Report envisages more in healthcare in the near future, such as remote monitoring, AI-powered alerting systems or virtual assistants, giving patients increasing ownership of their care (EIT Health & McKinsey, 2020). It also expects to see more AI solutions in clinical practice based on evidence from clinical trials, with increasing focus on improved and scaled clinical decision-support (CDS). Whereas in another field, as manufacturing, AI-powered predictive maintenance allows for drastic reductions in costly unplanned downtime, as well as for extending the Remaining Useful Life (RUL) of production machines and equipment.

Alongside opportunities, AI brings also risks and challenges, as always occurs with new waves of technological innovation. It presents new and unknown risks that may or may not require ad hoc laws and regulation. Some negative examples have been given wide attention in the media: a fatal accident involving an autonomous vehicle[2]; Microsoft's chatting bot Tay being shut down after 16 hours because it became racist, sexist, and denied the Holocaust[3]; racially biased decisions with credit checks and recidivism (Teich & Tirias Research, 2018). Such examples are fuelling a variety of concerns about accountability, fairness, bias, autonomy, and due process of AI systems (Pasquale, 2015; Ziewitz, 2015). Beyond these anecdotal instances, AI presents several challenges (Dwivedi et al., 2019), which are economic (need of funds, impact on employment and performances) and organizational (changing working practices, cultural barriers, need of new skills, data integration, etc.) issues to be tackled. At societal level AI may challenge cultural norms and face resistance (Hu et al, 2019). In Europe there is an ongoing discussion on the legal and ethical challenges posed by a greater use of AI. One key point is transparency, or lack thereof, of algorithms on which AI applications rely. There is a need to study and understand where algorithms may go wrong as to adopt adequate and proportional remedial and mitigation measures. Algorithmic rules may imply moral judgements, such as for driverless cars deciding which lives to save in the event of a serious accident (Nyholm, & Smids, 2016).

The European Commission has launched a series of policy initiatives with the aim to boost the development of sustainable AI in Europe, including the communication 'Artificial Intelligence for Europe' (European Commission, 2018a), the declaration of cooperation on AI (European Commission, 2018c), and coordinated action plan on the development of AI in the EU (European Commission, 2018d), among others. The European strategy aims to place people at the centre of the development of AI, what has been called 'human-centric AI'. It is a three-pronged approach to support the EU's technological and industrial capacity and AI uptake across the economy, prepare for socio-economic changes, and ensure an appropriate ethical and legal framework. The Commission has set up a High-Level Expert Group on AI representing a wide range of stakeholders and has tasked it with drafting AI ethics guidelines as well as preparing a set of recommendations for broader AI policy. The Group drafted AI Ethical Guidelines[4], which postulate that in order to achieve 'trustworthy AI', three components are necessary: (1) it should comply with the law, (2) it should fulfil ethical principles and (3) it should be robust. Based on these three components and the European values, the guidelines identify seven key requirements that AI applications should respect to be considered trustworthy[5]. These policies culminated in the White Paper on AI – A European Approach to Excellence and Trust (European Commission, 2020a) and a Communication on 'A European Strategy for Data' (European Commission, 2020b). The strategy set out in the Paper is built on two main blocks. On the one hand, it aims to create an 'ecosystem of excellence', by boosting the development of AI, partnering with private sector, focusing on R&D, skills and SMEs in particular. On the other hand, it aims to create an 'ecosystem of trust' within an EU regulatory framework. The strategy set out in the White Paper is to build and retain trust in AI. This needs a multi-layered approach that includes critical engagement of civil society to discuss the values guiding and being embedded into AI; public debates to translate these values into strategies and guidelines; and responsible design practices that encode these values and guidelines into AI systems making these 'ethical by design'. In line with this we have the European data strategy, adopted in February 2020, aiming to establish a path for the creation of European data spaces whereby more data becomes available for use in the economy and society but under firm control of European companies and individuals. As noted in a recent parliamentary brief (European Parliament, 2020), the objective of creating European data spaces is related to the ongoing discourse on Europe digital sovereignty (EPSC, 2019[6]) and the concern that, while Europe is at the frontier in terms of research and on a par with its global competitors, it nonetheless lags behind the US and China when it comes to private investment (European Commission, 2018a). The level of adoption of AI technologies by companies and by the general public appears comparatively low compared to the US (Probst et al., 2018). This leads to the concern that citizens, businesses and Member States of the EU are gradually losing control over their data, their capacity for innovation, and their ability to shape and enforce legislation in the digital environment. To address these concerns the data strategy proposes the construction of an EU data framework that would favour and support the sharing of data for innovators, particularly in the business-to-business (B2B) or government-to-citizens (G2C) domains: e.g. by open access to government data in sectors such as transportation and health-care (Burghin et al., 2019), privacy-preserving data marketplaces for companies to share data (de Streel et al., 2019). The genuine concern for innovators access to data is shown by the city of Barcelona where 'data sovereignty' clauses were introduced in public procurement contracts requesting its partners give back the data they gather to deliver services to the city in machine-readable format[7]. According to the cited European Parliament Brief (2020), if these clauses prove to be effective, they could be streamlined as best practices for the EU.

This report features in the EIT Digital policy perspective report series and is the result of a combined effort from five EIT KICs (EIT Manufacturing, EIT Urban Mobility, EIT Health, EIT Climate-KIC, and EIT Digital as coordinator). It identifies both general and sector specific concerns and opportunities for the further deployment of AI in Europe. Starting from the background and policy context outlined in this introduction, some critical aspects of AI are further discussed in Section 2. Next, in Section 3 four scenarios are proposed and assessed, from which a set of possible policy levers to address the concerns and exploit the opportunities are presented (Section 4). The paper concludes with a metaphor for the future development of sustainable AI in Europe (Section 5). In Annex (section 6) the main inputs from experts are reported. Al-shamasneh, A., & Obaidellah, U. (2017). Artificial Intelligence Techniques for Cancer Detection and Classification: Review Study. European Scientific Journal, 13(3), 342-370.

1 What we perceived as an intelligent and autonomous functioning of a machine 30 years ago, today may appear as nothing special. Labelling AI as intelligent or autonomous it means that AI is what we decide to call AI.

2 https://www.scientificamerican.com/article/what-the-first-driverless-car-fatality-means-for-self-driving-tech/

3 https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/

4 https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

5 The seven key 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.

6 The President of the European Commission, Ursula von der Leyen, in her 'Agenda for Europe' (von der Leyen, 2020) pledged to pursue the goal of achieving European 'technological sovereignty' in critical areas. A high-level political report by the Commission (Klossa, 2019), claims that that competition from global tech-driven players which do not always obey European rules and fundamental values, and which put data appropriation and valuation at the heart of their strategy, constitutes a major policy challenge for Europe.

7 https://medium.com/iipp-blog/putting-tech-and-innovation-at-the-service-of-people-and-the-green-transition-2e039ab8e083

CONSISTENCY, RELIABILITY, AND TRANSPARENCY: ALL CONTEXT-DEPENDENT

Many AI innovations are still emerging and in experimental phase, and as such they still have to prove their consistency and reliability. AI is implemented by algorithms that instruct a computer to execute tasks such as ordering possible choices (prioritisation), categorising items (classification), finding links between items (association), removing irrelevant information (filtering), or combinations of these tasks. Machine Learning (ML) and Deep Learning (DL) algorithms have the capacity to learn from data. ML algorithms can be used for web search, spam filters, recommender systems, ad placement, credit scoring, fraud detection, stock trading, drug design, and many other (Domingos, 2012). DL algorithms use artificial neural networks that can be applied to a wider typology of data (i.e., voice). The more training data the neural network processes, the more accurately the neural network can begin to process new, unseen inputs and successfully return the right results (Hof, 2013). ML and DL algorithms can go wrong because of data input quality or because of decision models used; at each step of the process depicted below one can trace the possible biases of AI algorithms.

The data may come from self-selected samples which may not always satisfy sound statistical rules, e.g. being representative for the real-world environment in which the system will perform. The samples may also reflect discriminatory biases already existing in society as for instance in the case when differences in arrest rates across racial groups may be replicated by an algorithm calculating recidivism risk (Chouldechova, 2017[8]).

The latter biases simply reflect societal process and are not created by AI algorithms, as these simply take the biased data for their decisions and hence in turn reflects inequalities and discrimination existing in society. An algorithm may produce unexpected and undesirable results if it encounters a situation that is markedly and rightly different from the data it was trained on. Statistical outliers (for instance disabled people) cannot easily be handled and are often just ignored. As algorithms become more sophisticated, their decision-making process may become less tractable[9]. The chosen decision model may also turn out to be unsuitable if the real-world environment behaves differently from what was expected. While more and better data be used for training can help improving prediction, it will never be perfect or include all justifiable outliers. On the other hand, as technology advances more instruments may become available to quantify the degree of influence of input variables on algorithm outputs (Datta et al., 2016). Research is also underway in pursuit of rendering algorithms more amenable to ex post and ex ante inspection (Jia & Liang, 2016). Nonetheless,

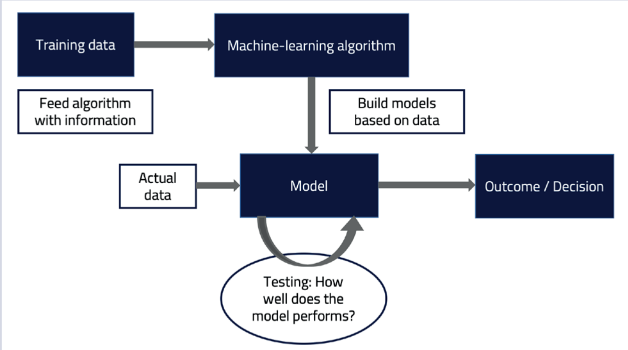

Figure 1: Where algorithmic biases can emerge,
Source: adapted from Buiten (2019).

in order to avoid biased decision and ensure fairness, we need to achieve transparency, explainability for accountability and liability for use of AI algorithms.

Algorithm explainability and transparency would mean that decisions/outcomes can be usefully and feasibly understood in such a way that it provides sufficient information to plaintiffs, defendants and courts in legal cases about contested algorithmic decisions or actions. In practice this require knowledge on what factors shaped an outcome (Doshi-Velez & Kortz, 2017, pp. 8-9) and could be achieved in three ways: a) transparency of the input (the training, testing and operational data) to check for biases (neutrality equal accuracy of data); b) transparency of the process (what inputs shape a prediction or decision), which is very challenging when decision models becomes more complex as in the case of for adaptive, self-learning systems (Ananny & Crawford, 2016, p. 982); c) transparency on decision/outcomes (does it cause harm?). Generating explanations of the functionality of complex algorithm is an extremely difficult engineering task, in particular when ML or DL is used. A legally binding right to explanation does not yet exist in the GDPR. Moreover, data controllers have an interest in not sharing details of their algorithms to avoid disclosing trade secrets, violating the rights and freedoms of others (e.g. privacy), and allowing data subjects to game or manipulate the decision-making system (Wachter et al., 2018). Because the ethical and social values of explanation remain despite difficulties, a counterfactual approach is suggested (Wachter et al., 2018). It could open the black box and avoid several barriers with the same results in a less time-consuming and more intuitive way: The example provided by Watcher et al (2018, p. 844) is as follows: You were denied a loan because your annual income was £30,000. If your income had been £45,000, you would have been offered a loan. The aim of such an approach would be to help data subjects in three ways: inform and help them understand why a particular decision was reached; provide them with grounds to contest adverse decisions; and help them understand how to achieve an outcome given current decision-making model.

Current policy challenges and debates, besides geopolitical discussions of digital sovereignty, all originate in the need of remedial and mitigation actions for risks that are unknown and difficult to envisage and control. The only possible pragmatic solution, avoiding unreasonable precautionary prohibition and requirements, is to understand and study such risks in specific situations, making them concrete and operational. Lifting the veils on specific algorithms does not necessarily need new laws and regulations. Existing law may already address problems effectively. Transparency and explainability as a way to control and mitigate possible biases producing discriminatory and/or harmful decisions/outcomes is very context-specific and should be considered in concrete and operational terms first, rather than being embedded into generalised and generic legal and regulatory prescriptions. Furthermore, the usefulness of transparency may depend on the risk associated with the decision and regulatory requirements should be based on risks to safety, fairness, and privacy in the particular context (Wachter et al., 2017).

8 It is not true that big data ensures validity and accuracy, also the quality of data matters (Domingos, 2012). If key data is withheld by design or chance, the algorithm's performance might become very poor (Olhede & Wolfe, 2017). The often-used implicit assumption that once we collect enough data, algorithms will not be biased, is not justified. (Barocas & Selbst, 2016). Bias can arise in algorithms in several ways. First, the data we have collected may have been preferentially sampled, and therefore the data sample itself is biased (Olhede & Wolfe, 2018). Second, bias can arise because the collected data reflects existing societal bias (Caliskan et al., 2017). To the extent that society contains inequality, exclusion or other traces of discrimination, so too will the data (Goodman & Flaxman, 2017).

9 Avoiding biased results rooted in social inequalities is difficult if sensitive information, such as ethnicity, is correlated with seemingly neutral variables, such as home address. In such cases, removing the sensitive variable will not prevent the biased result. Sophisticated algorithms may be able to reconstruct sensitive information from other inputs, even if they are not given this information.(Doshi-Velez & Kortz, 2017). With sufficiently large data sets, the task of exhaustively identifying and excluding data features correlated with 'sensitive categories' a priori may be impossible. If we are not aware of correlations between variables, these hidden relationships may obscure the rationale for how predictions are being made (Olhede & Wolfe, 2018, p. 4; Goodman & Flaxman, 2017, p. 4).

AI GOVERNANCE REGIMES: SCENARIOS AND THEIR ASSESSMENT

As a complement to the discussion our analysis confirms that AI applications share some common features but also have many differences. First and foremost, the sectors in which they are deployed have different drivers, objectives, characteristics, technologies, organisational structures and needs. Taking into account these aspects for each sector, as well as the level of deployment of AI applications, there are very different policy implications and challenges that policy makers will face. It emerges starkly that there are many different types of AI technologies and a wide range of domains, situations, and requirements. So, the challenge for Europe now is how to operationalise the principles of the White Paper across the diversity of AI technologies, domains of application, and deployment scenarios.

Just to give an obvious example there is a stark difference between what is needed for B2C deployment targeting end users and leveraging personal data and the deployment in manufacturing mostly relying on machine data. Policy makers must balance between the two different poles of introducing new regulation or just relying on the general provision of GDPR and general principles of the White Paper and leaving the operationalisation to various forms of co-regulation and self-regulation. The choice of further regulation could be based on the application of the precautionary principle in the face of uncertain but possible risks. To some extent one could see this approach when the EC AI White Paper uses general expressions such as 'exceptional circumstances' or 'immaterial damages'. They seem to suggest a general assumption about uncertain and undefined dangers and, so, a worst-case scenarios across-the-board precautionary approach. On the opposite end, a precautionary approach specific to certain domains would favour that specific risks are mapped and considered as input to a

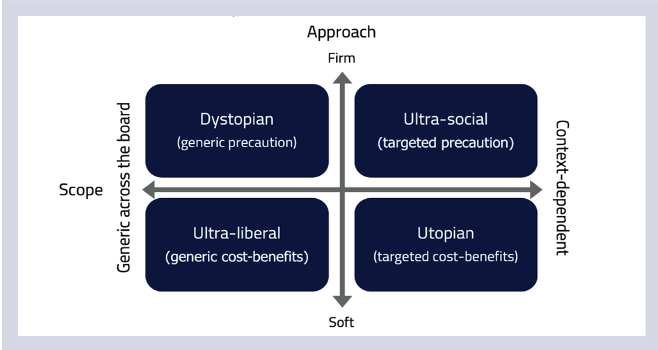

Figure 2: Proposed scenario,

cost-benefit analysis on a case by case or domain-based basis. The fear for across-the-board precaution regulation tends to be used by industry to push for self-regulation. Indeed, it could run the risk that its negative side effects (i.e., hampering innovation and growth) would outweigh the benefit of avoiding uncertain general risks (unfair discrimination, growing distrust). On the other hand, under conditions of radical uncertainty, a simple cost-benefits analysis based on assessment of risk may end up minimising the latter to push innovation and growth at the expenses of fairness and trust. Different from what is suggested in the AI White Paper - an approach across-the-board and irrespective of sector features is suggested - our analysis, while pointing out barriers and challenges, suggests that there are very concrete solutions that can be tailored to concrete situations and that there are no irremediable dangers and risks.

Following on the considerations above and the discussion of Section 2, in order to move toward a pragmatic policy-oriented instrument that can bring shapers and makers together, a scenario approach (see figure 2) was followed. The vertical axis addresses 'Algorithms and Data Governance', meaning the algorithms with their underlying data. The neutral concept of 'Governance' can be ensured by different mechanisms including law and regulation, co-regulation, self-regulation, as well as by market mechanisms. Governance of algorithms and of their underlying data concerns issues of personal data protection, privacy, machine data, transparency, liability, but also the drivers that can boost AI deployments. Along the vertical axis the approach is ranging from 'firm governance' mostly through law and regulation to 'soft governance' mostly by way of self-regulation and market functioning. So, in brief, the vertical axis is about the approach and instruments. On the horizontal axis, on the other hand, the dimension is 'Scope' ranging from 'context-dependent' to 'generic across-the-board'. Context-dependent means a governance regime with specific approaches and instruments tailored to contextualised sectorial instances of both opportunities and risks. Generic across the board is when the same approach and instruments are applied irrespective of contextual and sectorial specificities in terms of potential opportunities and risks.

As usual, scenarios aim at uncovering extremes and their potential impact, but they are not an end in themselves. Rather they are instrumental to pinpoint radical discontinuities, from which one can subsequently derive more nuanced and pragmatic regulatory and policy levers that position themselves somewhere in the middle of the space determined by the axes.

Dystopian. Under this scenario there is a hardening of regulation on the basis of the precautionary principle across-the-board, irrespective of applications and sector specific situations. As regulation is introduced on the basis of the worst scenario, negative side effects are such that no regulation could be preferred on the same precautionary grounds. With the intention of not downplaying dangers, horizontal and non-specific regulation would build a negative public discourse that would block innovators. Transparency requirements becomes mandatory in all sectors and create barriers especially for innovators and SMEs. Public procurement 'data sovereignty clauses' induce large players to withdraw from AI for urban ecosystems. Strict liability sanctions block AI in healthcare, while limiting space of self-driving experimentation. The support measures to boost European AI are not sufficient to offset the unintended effect of generic precaution and as a result Europe accumulate additional lags vis-à-vis its global competitors.

Ultra-social. Under this scenario regulation inspired by the precautionary principle is adopted but with a vertical approach that takes into consideration sector and application specific aspects. Though regulatory, this approach enables to consider nuances and different situations and is fairer and more inclusive in that it does not block and hamper AI innovation tout-court as in the previous scenario. Regulation and other instruments aimed at positive results concerning European values, transparency, responsibility, equal access and non-discrimination, etc. are introduced. Following the best spirit of the EU data strategy, incentives are successfully created for data sharing in both B2C and G2C, without too much unnecessary emphasis on clauses rhetorically addressing issue of digital sovereignty.

Utopian. This scenario is utopian for it would combine the vertical specific and thorough analysis of dangers and negative side effects with a self-regulatory approach on the side of both innovators and business incumbents. Whereas this would be an ideal solution, it seems unrealistic that such approach would be adopted on a large scale if let entirely to voluntary self-regulation.

Ultra-liberal. This is probably the ideal scenario for large platform incumbents, in that actual operationalisation is left to self-regulation based on cost-benefit analysis. What makes it ultraliberal is the horizontal and generic focus of cost-benefit that, on the one hand will not hamper experimentation and innovation, but on the other does not ensure that specific conditions of risk and uncertainty are fully factored in so that individuals and society as a whole are protected from negative side effects.

GENERIC AND CONTEXT DEPENDING OPPORTUNITIES AND POLICY LEVERS

GENERIC

There are five issues that, though from slightly different angles, are considered strategic and a potential source of barriers and bottlenecks: data, organisation, human capital, trust, markets. The availability and quality of data, as well as data governance are of strategic importance. Strictly technical issues (i.e., inter-operability, standardisation) are mostly being solved, whereas internal and external data governance still restrain the full potential of AI Innovation. Organisational resources and, also, cognitive and cultural routines are a challenge to cope with for full deployment. On the one hand, there is the issue of the needed investments when evidence on return is not yet consolidated. On the other hand, equally important, are cultural conservatism and misalignment between analytical and business objectives. Skills shortages are a main bottleneck in all the four sectors considered in this report where upskilling, reskilling, and new skills creation are considered crucial. For many organisations data scientists are either too expensive or difficult to recruit and retain. There is still a need to build trust on AI, amongst both the final users (consumers, patients, etc.) and intermediate / professional users (i.e., healthcare professionals). This is a matter of privacy and personal data protection, of building a positive institutional narrative backed by mitigation strategies, and of cumulating evidence showing that benefits outweigh costs and risks. As demand for AI innovations is still limited (in many sectors a 'wait and see' approach is prevalent) this does not favour the emergence of a competitive supply side. Few start-ups manage to scale up, and many are subsequently bought by a few large dominant players. As a result of the fact that these issues have not yet been solved on a large scale, using a 5 levels scale of deployment maturity (1= not started; 2= experimentation; 3= practitioner use; 4= professional use; and 5= AI driven companies), it seems that, in all four vertical domains considered, adoption remains at level 2 (experimentation) or 3 (practitioner use), with only few advanced exceptions mostly in Manufacturing and Healthcare. In Urban Mobility, as phrased by interviewed experts, only lightweight AI applications are widely adopted, whereas in the Climate domain we are just at the level of early predictive models. Considering the different areas of AI applications, regardless of the domains, the most adopted ones include predictive maintenance, chatbots, voice/text recognition, NPL, imagining, computer vision and predictive analytics.

MANUFACTURING

The manufacturing sector is one of the leaders in application of AI technologies; from significant cuts in unplanned downtime to better designed products, manufacturers are applying AI-powered analytics to data to improve efficiency, product quality and the safety of employees. The key application of AI is certainly in predictive maintenance. Yet, the more radical transformation of manufacturing will occur when manufacturers will move to 'service-based' managing of the full lifecycle from consumers preferences to production and delivery (i.e., the Industry 4.0 vision). Manufacturing companies are investing into this vision and are keen to protect their intellectual property generated from such investments. So, there is a concern that a potential new legislative action by the European Commission, which would follow the principles of the GDPR and the requirements of the White Paper, may

Table 1: Generic: concerns, opportunities and policy levers.

Concerns/ opportunities	Description	Policy/Regulatory lever
Data	Data governance (data preparation, data flows, data sharing) still a challenge for many organisations, also due to lack of inter-operability and standardisation. This limits the availability of data to train algorithms.	Regulate and/or stimulate European interoperability, and support cooperation of EU organisations in international standardisation and interoperability initiatives. In addition, mandate open access for non-personal data collected by public bodies.
Organisation	Organisations are conservative structures and, like any innovation, also AI faces internal resistance (between units or between companies). In some cases, this approach is driven by reliability or liability concerns.	Ensure certainty by regulating liability in AI applications. A clear regulation on liability for autonomous machines in factories would also ease potential tensions with trade unions.
Human capital	Skills shortage, the risk of replacement of human labour, and AI / human complementarity are a source of challenges and concerns.	Labour policy should protect workers rather than jobs by a 'flexicurity' system, where labour flexibility in the regulation of contracts and social security are combined. In parallel, strong support for specific (AI) training and re-training.
Trust	AI relies on vast amount of data and on the algorithms processing them to extract predictions on the basis of which decisions are taken. For full adoption of AI, users need to trust data processors and the results of the algorithms, especially when sensitive data are used to make decision affecting the data subjects.	Regulation should allow a sandbox approach when deploying AI, similar to phase I to III clinical studies of medical drugs. In a sandbox, potential issues could be identified, and trust built before widespread deployment. Another policy initiative would be setting up RTI projects to pilot an AI-driven privacy management system (see section 5).
Markets	Large companies often acquire small ones that have launched innovative AI solutions that risk of not being developed either because the acquisition is just to remove competition or because the AI solutions are not profitable in the short term.	New anti-trust approaches, such as considering companies' data assets when assessing merger requests and the price as signal of incumbent buying an emerging threat.

be inappropriate for the machine-data that are at the core of AI in manufacturing. On the other hand, there are issues that could be better clarified and streamlined by a common EU regulatory approach such as the issue of liability (extremely important also in factories where risks may come from autonomous machines). For instance, the introduction of a mandatory product liability insurance for manufacturers of autonomous machines could shift the discourse from the user/driver to the producer. This could be an achievable solution in the near term to speed up the development of autonomous vehicles and machines. The traditional approach to liability will need to give way to more product-related liability coverage or hybrid coverage.

More space is needed for experimentation of AI in the manufacturing sector, as any new regulation resembling the GDPR approach applied to machine data risk to leverage the conservative culture that characterise many small and medium manufacturing companies and hinder innovation. In addition, to foster the development of AI, more space like the co-innovation labs should be created around the EU[10], to bridge the gap between research, development, market and society. In this way, also small and medium companies have the chance to understand their needs and see directly the impacts that specific AI technologies can have on their businesses.

10 https://ict.fbk.eu/partnerships/co-innovation-labs/

Table 2: Manufacturing: concerns, opportunities and policy levers.

Concerns/ opportunities	Description	Policy/Regulatory lever
Transparent regulation	There is a risk that machine data be treated as personal data from a regulatory perspective. However, the same transparency requirements could be detrimental for business and trade secrets. Liability of producers vis-à-vis users is unclear and creates uncertainty hindering innovation and uptake.	Introduce a regulation specific to machine data, with clarity on the interdependence with GDPR, ensuring competition interests. Regulate product liability for autonomous systems and obligatory insurance.
Replacement human labour	Job certainty for workers throughout life is threatened by automation in factories. It will also bring the need to invest strongly in training and re-training.	Labour policies based on flexicurity (basic income, less working hours, more certainty for workers) could be considered, including easier change of jobs through live-time and pension arrangements independent from the employer. The use of AI in training and work guidance should be stimulated.
Sustainability	Great opportunities to predict the lifetime of a product or machine, which are substituted only when needed (and also reduces the risks for workers/consumers).	Support the incentives for those products that have an incorporated prediction of their lifetime to create a culture of predictive maintenance and reap the effects on sustainability.
Predictive manufacturing	There is potential to radically transform the manufacturing sector and its processes (see above) leading to higher productivity and innovation.	Stimulating piloting of AI in production processes and for example in predictive maintenance also for SMEs, which are more conservative and risk averse.

URBAN MOBILITY

The adoption of AI in the management of urban mobility systems brings different sets of benefits for private stakeholders (citizens, private companies) and public stakeholders (municipalities, transportation service providers). So far only light-weight task specific AI applications have been deployed (i.e., intelligent routing, sharing apps, predictive models based on citizens' location and personal data). On the other hand, the most advanced and transformative AI applications, such as autonomous vehicles are lagging behind, especially if compared to the US or China. The key challenge for AI deployment in Urban Mobility sector is the need to find a common win-win business model across a diversity of public and private sector players with different organisational objectives, cultures, and managerial capabilities and with different access to different data. While personal data are mostly in the hands of few large private players, machine data coming from sensors in the cities, or from public transport, are owned by the municipalities. The challenge is, thus, finding a common ground to exploit the potentials. Not only combining personal data with machine data requires a business model which needs to realign the incentives of public and private sector, but the risk is the exclusion of small innovative companies that would be left out by individual agreements between the municipality and the large companies.

Table 3: Urban Mobility: concerns, opportunities and policy levers.

Concerns/ opportunities	Description	Policy/Regulatory lever
Data collection	The increased adoption of sensors in the urban environment raises concerns about the data that are gathered in relation to privacy. The improvement of services may be counterbalanced by a lack of trust.	Data collected by cities (on e.g. infrastructure and environment) should be open for access to stimulate innovation. The same applies for certain data collected by private enterprises operating under a license by cities (e.g. mobility providers).
Business model	Municipalities seek public values (traffic reduction, wellbeing) while the private companies seek profit. Well-balanced Public-Private partnerships are needed in the urban mobility context.	PPPs should be encouraged, with the right incentives for private companies, and clear public objectives, in particular for projects using AI (e.g. improve start-to-end travelling in time, comfort and safety).
Sustainable cities	The reduction of traffic due to a more intelligent transportation system is possibly one of the greatest public values that can be created in large cities.	Stimulate the debate on sustainable cities and link the efficiency goals of AI-based applications with the environmental impact to create a 'green culture' in the cities.
Personalised services	An integration of all the transport means (private/public), together with personal specific characteristic can lead to one single platform able to plan personalised journeys.	Work towards a personal agent to support people in travelling, finding the shop or bar they search for. This should not be paid for by advertising, but either directly by the person, or as part of municipal services.

One of the most interesting development close to scale up is the creation of platforms, which are fed by all different data sources of transport services (both private and public) and provide the citizens a targeted recommendation on the best way to travel, also based on personal preferences and characteristics.

Urban Mobility should focus on what is already potentially available now but that faces several barriers. Making use of the vast amount of data in the hand of the public administration requires AI experts capable of preparing the data to be used and, most importantly, a collaboration framework with private companies that hold personal data on citizens' travels and needs. A new business model and ways to cooperate should be found, so that both the public, the private and the innovators achieve their objectives. First, better public services that reduce traffic, make cities more sustainable and that meet citizens' needs. Second, profitable business models so that a cooperation with the private actors is possible to openly share data. Lastly, the Urban Mobility KIC stressed that for SMEs and start-ups the administrative burden is often too high to develop an interesting idea.

HEALTH

Healthcare is arguably the sector where AI could make the greatest impact in addressing societal challenges. Given rising demands and costs, AI could help doing more and better with the same. The Covid-19 crisis has shown how strained our National Healthcare Systems are, and AI solutions could help meet the current and possibly future crisis by releasing human resources from some tasks so that they could be allocated to most mission critical activities. So far, however, AI applications in healthcare have been confined to administrative tasks (i.e., Natural Language Processing to extract information from clinical notes or predictive scheduling of the visits) and diagnostic (machine and deep learning applied to imaging in radiology, pathology and dermatology). Important gains are being achieved in these two domains (reducing time healthcare professionals devote to routine and repetitive tasks; more accurate and faster diagnostic), but the potential of AI to transform healthcare is much wider as deployment trail significantly behind what the industry is already offering.

An achievable solution to reach in the near term is the use of AI applications to overcome healthcare data security challenges and improve patient trust. Some of these have already been tested to produce a highly accurate privacy analytics model that reviewed every access point to patient data and detected when the electronic health record was potentially exposed to a privacy violation, attack, or breach.

Longer term efforts, requiring strong EU wide cooperation and support, are needed to create a secure, trusted and coherent data space that allows access to health data of patients whenever and wherever it is needed.

Many initiatives are launched that remain small, and the uptake by hospitals is still limited. One issue is that data governance and AI tools still need to prove that data privacy is treated with extra care in addition to GDPR requirements, and that the tools produce a clear return on investment. One obvious and very important issue concerns personal data and patient trust. The opportunities here are great, as more use of AI in research and development could lead to a more personalised healthcare based on patients' data. A second challenge is that of finding a common language and understanding between data experts and healthcare professionals.

AI in the healthcare faces organisational and skill challenges. One priority is to support upskilling or reskilling of healthcare practitioners through tailored educational programmes, to understand AI's potential. Specific healthcare training should be provided to data scientists working in hospitals so that they can better understand healthcare practitioners needs. In addition, at the regulatory level it is important that new AI regulation is harmonised with other pieces of regulation already in existence (i.e. medical device regulation). The risk of introducing new certification systems (new guidelines for clinical trials) should be avoided. The development of new legislation and regulation at national and regional level should be constantly monitored as to reduce fragmentation.

Table 4: Health: concerns, opportunities and policy levers.

Concerns/ opportunities	Description	Policy/Regulatory lever
Health data	Health data is a specific class of personal data and very sensitive. An extensive use to feed AI tools can raise many concerns. Data ownership is also an issue especially because it varies widely across EU Member States.	Develop a specific framework of transparency for the use of patient's data in the hospitals that deploy AI-powered applications. The patients should be aware of how their data are used.
Healthcare fragmentation	Fragmented European healthcare systems and standards are a serious concern for individuals who move within or between EU countries and from one to another hospital or care worker.	Need to stop fragmentation of data storage and access, within and between Member States, so that patients can move around EU and be sure that their digital clinical records are accessible in other hospitals and care situations. This needs big efforts in standardisation and interoperability.
R&D and innovation	More patients' data available coming from digital application represent a great opportunity for R&D and innovation.	Stimulate and encourage the release of personal data to advance scientific research. The objectives should be clear, so that patients can trust the organisations.
Remote healthcare	With AI finally remote healthcare is becoming possible, as AI solutions can increasingly divert patients to appropriate solutions for their specific symptoms and underlying conditions.	Support the transition to remote healthcare by training healthcare professionals, starting from the simple tasks and diagnostic appointments.

CLIMATE AND ENERGY

EIT's climate-focused KIC – EIT Climate-KIC – notes that the predictive capacity of AI has given it an important role in driving more effective climate actions. Predictive analytics is one of the most common applications of AI and has helped propel this technology to centre stage in the climate domain. However, the initiatives that really use AI in this field are still relatively few, and there is a need to move beyond AI applications that reduce climate risks and the negative externalities of economic activity to applications that can help transform whole systems – ideally nested within larger portfolios of social, economic and financial innovations and missions. But where are we now? There is much to build on. Through the analysis of countless data sources, AI has been able to generate insights into the future. With the climate crisis, AI tools have been used mainly in two areas: land-management and extreme weather forecast. For the first category of applications, EIT Climate-KIC has launched an initiative called OpenSurface, which is a platform that combines and integrates multiple data sources to better monitor and forecast land-use change. OpenSurface uses authoritative land records in conjunction with satellite imagery and ground-sourced data to help prioritise resource allocation. The platform combines fourth-generation technologies such as artificial intelligence, secure ledgers, remote sensing, and the internet of things to automatically compare planned, authorised activities with how forests are actually being managed or depleted. The second category of AI applications is related to weather conditions. With data provided from weather stations around the world, measuring wind speed, air pressure, temperature, and many other data points, AI can then identify patterns and give a holistic view of how our planet is changing.

The biggest challenge is that AI, as with most transformative technologies, can certainly produce benefits in terms of efficiency and effectiveness but can also further harm climate. The first problem is related to energy consumption. While currently only using 2% of the world's energy, data centres without substantial investment in developing new materials, designs, and chips, will increase exponentially their electricity consumption. Another potential detrimental impact raised by several stakeholders in the field is what has been known as 'rebound effect'. It has long been known that efficiency gains in a given process or service can lead to growth in our reliance on this same process or service — to such an extent that the growth cancels out the gains. Therefore, the peculiarity of the climate sector poses the challenge of balancing the potential gain and negative side effects that AI itself may produce as a result of its own demand for power and increase in energy consumption.

The development of AI should proceed hand-in-hand with attention to the environmental impact of new technologies. Excessive focus on efficiency gains and cost-effectiveness could hide the detrimental effects that some innovations produce on the climate. EIT Climate-KIC holds the view that Europe should aim to lead in leveraging and balancing AI for sustainability. This requires a rethinking of the kind of return we want from AI innovation, which could be used for improving the sustainability of products and to reduce the environmental impacts of production processes and energy consumption.

Table 5: Climate and Energy: concerns, opportunities and policy levers.

Concerns/ opportunities	Description	Policy/Regulatory lever
Rebound effect	AI applications may have a negative impact on the climate, as the increased effectiveness is accompanied by more energy consumption (i.e., rebound effects).	The environmental impact of AI applications should become transparent, and presenting the effects on the climate, together with efficiency and other targets, should become the standard.
Privacy concerns	AI tools that monitor, predict and support efficient energy consumption (e.g. smart meter) are not always transparent regarding the use of personal data.	All AI applications should be proven GDRP compliant with guarantees for privacy and the use of data under AI supported awareness and consenting.
Climate risk prediction	Accurate predictions on extreme weather conditions can help policy makers to better prepare and manage climate-related risks.	Stimulate the research and experimentation of AI solutions in those areas more at risk of extreme events. Create innovation hubs in symbolic places in Europe, where the population suffered from extreme events.
Energy efficiency	AI driven systems can better monitor and understand energy consumption, land use and climate change, to help individuals and companies to make an efficient use of resources and reduce waste.	Encourage and support the integration and interoperability of energy systems, so that predictions and monitoring can be done more effectively.

AI GOVERNANCE REGIMES: THEIR ASSESSMENT

Overall the contents of the five tables presented in the previous section and the proposed policy levers cover the entire spectrum of the EC AI White paper, whose logic is summarised in the picture below. As a matter of fact, the five tables contribute to make the general principles and envisaged initiative contained in the White Paper more operational and more sector specific. In doing so, they follow the spirit of the message implicit in the scenarios exercise where one does not want to achieve one of them, but rather combine the elements of each scenario into a set of action that maximise the positive components and minimise the negative ones of each scenario.

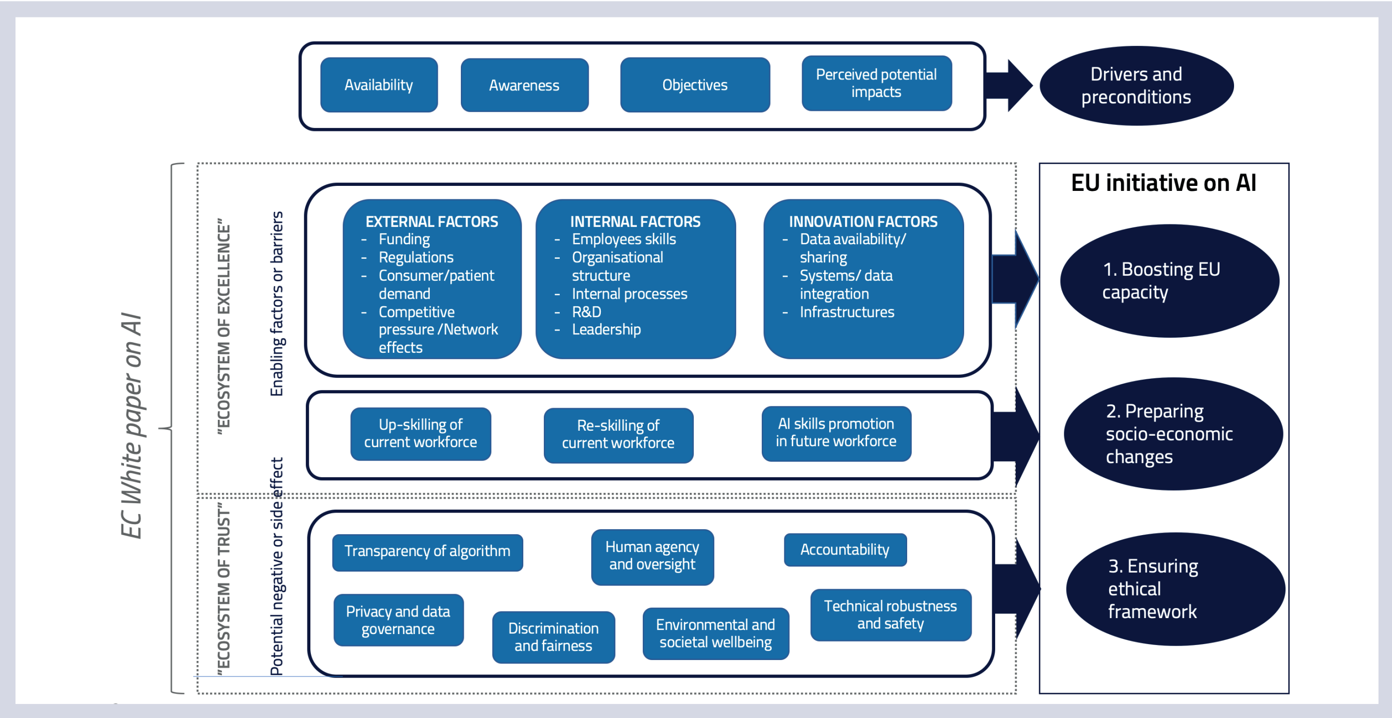

Figure 3: Schematic logic of the EC AI White Paper, Source: authors' elaboration

Based on the considerations in the previous sections, a qualitative assessment of the four scenarios with regard to four dimensions (Growth, Innovation, Fairness, and Trust), representing high-level policy objectives, was conducted. The scenarios have been scored in the strict order from least (1) to most impact (4) with regard to the four dimensions of assessment, thus providing a relative comparison between the scenarios. This has been depicted in the spider diagram overleaf. Note that using a strict order forces a strong discrimination between the scenarios and leaves less room for nuances.

The spider diagram, therefore, magnifies differences and has to be seen as a tool to give a quick insight into the relative strengths and weaknesses of the scenarios, rather than absolute differences. The utopian scenario would score the highest in terms of both innovation and growth, whereas the ultra-social is best in terms of fairness and trust. The ultra-liberal scores well on both innovation and growth, and less so on fairness and trust. Moving out from these extreme scenarios and looking pragmatically at a combination of elements that can be extracted from them, the tables in the previous sections – one common and one for each sector – have raised a number of issues and then mapped against possible policy and regulatory levers.

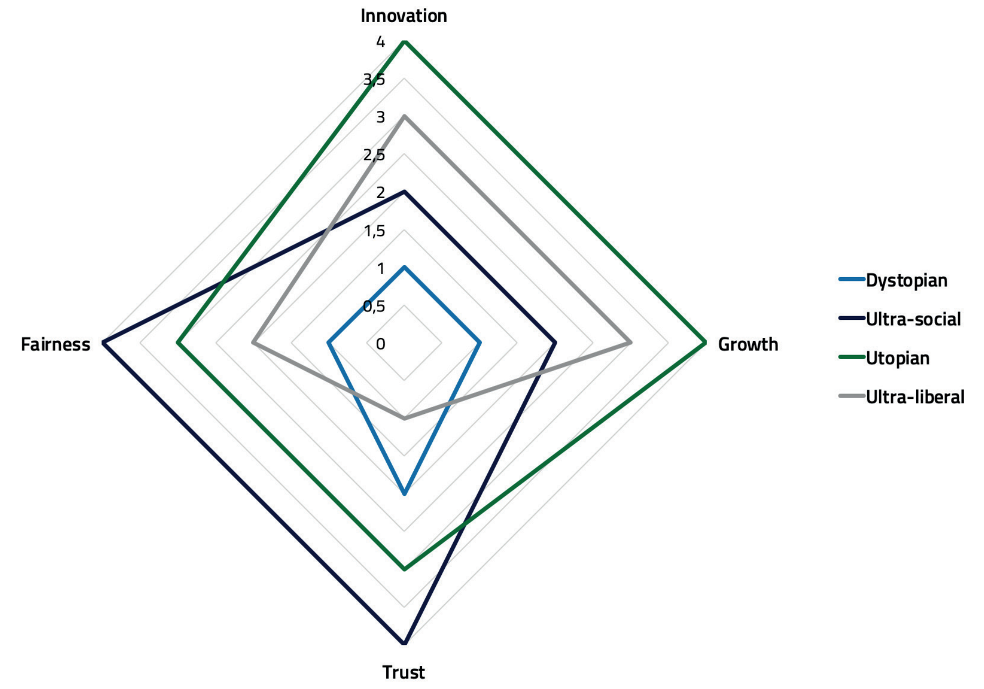

Figure 4: Scenarios assessment,

CONCLUSION

The chapters above give an impression of the state of affairs, generally and in the four sectors addressed specifically. They show difficulties as well as opportunities and suggest how these can be handled by specific policy levers. It is also obvious that the four sectors have very different requirements and that overall, across-the-board regulative solutions maybe good for one sector, but hinder innovation in another. The horizontal axis in the set-up of the scenarios plays an important role in coming to effective policy development.

Four main conclusions can be drawn from the above chapters:

- To ensure effective policy in the area of AI it is necessary to take context (sectors of application) into account.

- Policies regarding application of AI on personal data should be allowed to differ from policies regarding application of AI on machine data, especially in certain application sectors.

- General regulation or policy measures can be considered in relation to algorithm transparency and explainability

- Regulation should be adaptable and flexible, whilst minimising and mitigating risks and ensuring human rights and European values.

In addition, the following principles contribute to increase the positive impact of AI applications:

1. **Privacy management supported by AI.** Data sovereignty by subjects could be supported through developing an ecosystem where all transfer of data will be done with guaranteed GDPR compliance, under auditing of the regulator. Personal data will be under control of the subject and stored in a cloud. It can be transferred with consent/contract between the subject and other parties. The consent management will be done with AI support (a personal AI data management guard) which will automatically agree to data exchange if it is standard GDPR compliant and/or agreed once by the subject for such situation. When the AI agent concludes there is doubt, the subject will be warned and gives consent or not, thus teaching the AI agent for the future. R&I projects could be stimulated to develop these ideas.

2. **Counterfactual checks and Algorithm explainability.** This is an approach proposed by Watcher et al where they say there are three goals: inform and help data subjects understand why a particular decision was reached; provide them with grounds to contest adverse decisions; and help them understand how to achieve an outcome given current decision-making model. This would need regulation as well as policy and R&I actions.

3. **Sandbox-based regulation.** Regulation should stimulate to use a sandbox approach when deploying AI, similar to phase I to III clinical studies of medical drugs. In a sandbox, potential issues could be identified, and trust built before widespread 5.1deployment. The rules for a sandbox methodology could be different per sector, with in the Health sector a system close to the current stepwise approval processes for medicines and equipment.

REFERENCES

Al-shamasneh, A., & Obaidellah, U. (2017). Artificial Intelligence Techniques for Cancer Detection and Classification: Review Study. European Scientific Journal, 13(3), 342-370.

Ananny, M., & Crawford, K. (2016). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973-989. doi:10.1177/1461444816676645.

Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, P., De Prato, G., Feijoo, C., Fernandez Macias, E., Gomez, E., Iglesias, M., Junklewitz, H., López Cobo, M., Martens, B., Nascimento, S., Nativi, S., Polvora, A., Sanchez, I., Tolan, S., Tuomi, I., & Alujevic, L. V. (2018). Artificial Intelligence: A European Perspective. Joint Research Centre, European Commission. https://doi.org/10.2760/11251.

Barocas, S., & Selbst, A. (2016). Big Data's Disparate Impact. California Law Review, 104, 671-732.

Buiten, M. C. (2019). Towards Intelligent Regulation of Artificial Intelligence. European Journal of Risk Regulation, 10(1), 41-59. doi:10.1017/err.2019.8.

Datta, A., Sen, S., & Zick, Y. (2016, 22-26 May 2016). Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems. Paper presented at the 2016 IEEE Symposium on Security and Privacy (SP).

de Streel, A., Kramer, J., Bourreau, M. (2019). Ambitions for Europe 2024. Cerre White Paper. Retrieved from: https://www.cerre.eu/sites/cerre/files/cerre_whitepaper2024_digitaltelecommedia_0.pdf.

Domingos, P. (2012). A few useful things to know about machine learning. Commun. ACM, 55(10), 78-87. doi:10.1145/2347736.2347755.

Doshi-Velez, F., & Kortz, M. (2017). Accountability of AI Under the Law: The Role of Explanation. Berkman Klein Center Working Group on Explanation and the Law (retrieved from: https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf?sequence=3, 21/11/2019).

Dreyer, K., & Allen, B. (2018). Artificial intelligence in health care: brave new world or golden opportunity? Journal of the American College of Radiology, 15(4), 655-657.

Dwivedi, Y. K., Hughes, L., et al. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 101994.

EIT Health and McKinsey & Company, (2020), Transforming healthcare with AI. Impact on the workforce and organisations. Retrieved from: https://eithealth.eu/wp-content/uploads/2020/03/EIT-Health-and-McKinsey_Transforming-Healthcare-with-AI.pdf.

EPSC. (2019). Rethinking Strategic Autonomy in the Digital Age. Brussels: European Political Strategy Centre (retrieved from: https://ec.europa.eu/epsc/sites/epsc/files/epsc_strategic_note_issue30_strategic_autonomy.pdf , 14/10/2019).

European Commission. (2018a). Artificial Intelligence for Europe. COM(2018) 237 final, Brussels: European Commission.

European Commission. (2018b). Factsheet: Artificial intelligence for Europe. Retrieved from https://ec.europa.eu/digital-single-market/en/news/factsheet-artificial-intelligence-europe.

European Commission. (2018c). EU Member States sign up to cooperate on Artificial Intelligence. Retrieved from https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence.

European Commission. (2018d). Member States and Commission to work together to boost artificial intelligence 'made in Europe'. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/IP_18_6689.

European Commission. (2020a). White Paper on Artificial Intelligence. A European Approach to Excellence and Trust. COM(2020) 65 final, Brussels: European Commission.

European Commission. (2020b). A European Strategy to Data. COM(2020) 66 final, Brussels: European Commission.

European Parliament. (2020). Digital sovereignty for Europe. Brussels: European Parliament (retrieved from: https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/651992/EPRS_BRI(2020)651992_EN.pdf, 07/07/2020).

Fuller, T., & Metz, C. (2018, October 26). A.I. Is Helping Scientists Predict When and Where the Next Big Earthquake Will Be. The New York Times. Retrieved from https://www.nytimes.com/2018/10/26/technology/earthquake-predictions-artificial-intelligence.html (21/11/2019).

Goodman, B., & Flaxman, S. (2017). European Union regulations on algorithmic decision-making and a 'right to explanation. AI Magazine, 38(3), 50-57., 38(3), 50-57.

Hof, R. (2013, April 23). Deep Learning. The MIT Technology Review. Retrieved from https://www.technologyreview.com/s/513696/deep-learning/.

Jia, R., & Liang, P. (2016). Data Recombination for Neural Semantic Parsing. Paper presented at the Annual Meeting of the Association for Computational Linguistics, Berlin.

Klossa, G. (2019). Towards European Media Sovereignty. An Industrial Media Strategy to leverage Data, Algorithms and Artificial Intelligence. European Commission. Retrieved from: https://ec.europa.eu/commission/sites/beta-political/files/gk_special_report-european_media_sovereignty.pdf.

Nyholm, S., & Smids, J. (2016). The ethics of accident-algorithms for self-driving cars: An applied trolley problem? Ethical theory and moral practice, 19(5), 1275-1289.

Olhede, S., & Wolfe, P. (2017). When algorithms go wrong, who is liable? Significance, 14(6), 8-9. doi:10.1111/j.1740-9713.2017.01085.x

Olhede, S. C., & Wolfe, P. J. (2018). The growing ubiquity of algorithms in society: implications, impacts and innovations. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376(2128), 20170364. doi:10.1098/rsta.2017.0364.

Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information Boston: Harvard University Press.

Probst, L., Pedersen, B., Lefebvre, V., & Dakkak, L. (2018). USA-China-EU plans for AI: where do we stand. Report prepared for the European Commission, Directorate-General Internal Market, Industry, Entrepreneurship, and SMEs, Digital Transformation Monitor of the European Commission.

Scherer, M. (2016). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. Harvard Journal of Law & Technology, 29(2), 353-400.

Teich, D. (2019, April 25). Management AI: A Senate Bill Introduced To Begin Regulation Of Machine Learning Algorithms. Forbes. Retrieved from https://www.forbes.com/sites/davidteich/2019/04/25/management-ai-a-senate-bill-introduced-to-begin-regulation-of-machine-learning-algorithms/#5995188179e6 (21/11/2019).

Teich, D., & Tirias Research. (2018, January 24). Management AI: Bias, Criminal Recidivism, And The Promise Of Machine Learning. Forbes. Retrieved from https://www.forbes.com/sites/tiriasresearch/2018/01/24/management-ai-bias-criminal-recidivism-and-the-promise-of-machine-learning/#4ec6a9547c8a (21/11/2019).

von der Leyen, U. (2020). A Union That Strives for More. My Agenda for Europe. Political Guidelines for the Next European Commission 2019-2024. Retrieved from: https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en-.pdf.

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Transparent, explainable, and accountable AI for robotics. Science Robotics, 2(6), eaan6080. doi:10.1126/scirobotics.aan6080.

Wachter, S., Mittelstadt, B., & Russell, C. (2018). Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. Harvard Journal of Law & Technology, 31(2), 841-887.

ACKNOWLEDGEMENTS

In the context of their strategic innovation activities for Europe, five EIT Knowledge and Innovation Communities (EIT Manufacturing, EIT Urban Mobility, EIT Health, EIT Climate-KIC, and EIT Digital as coordinator) decided to launch a study that identifies general and sector specific concerns and opportunities for the deployment of AI in Europe.

The study followed a scenario-based approach to structure and assess the potential impact of different policy measures with regards to the regulation of Artificial Intelligence.

Digital Enlightenment Forum was contracted to write this report under the guidance of EIT Digital, supported by the contributing four EIT KICs. The strong involvement of the five EIT Knowledge and Innovation Communities in addition to external high-level experts from various disciplines formed the basis for a profound assessment of the different dimensions and their interrelations.

Digital Enlightenment Forum is particularly grateful to Cristiano Codagnone and Giovanni Liva (Univ. degli Studi di Milano), as well as selected advisors and experts for giving their time for interviews and discussions that helped to better understand the complexity of the issues involved and to formulate the results presented here.

This report is the third edition in EIT Digital's series of policy perspective reports. Edition one on the Digital Transformation of European Industry as well as edition two on European Digital Infrastructure and Data Sovereignty can be downloaded from EIT Digital's website at www.eitdigital.eu/newsroom/publications.

Publisher
EIT Digital
Rue Guimard 7
1040 Brussels
Belgium
www.eitdigital.eu

Contact
info@eitdigital.eu

EIT Climate-KIC, EIT Digital, EIT Health, EIT Manufacturing and EIT Urban Mobility are supported by EIT, a body of the European Union.

ISBN 978-91-87253-64-5