

Central European University Institut Barcelona d'Estudis Internacionals Academic Year 2019 – 2020

A Policy-First Approach to AI Development: A structured comparison of three key strategies

Dissertation submitted by Valentino Grassi in partial fulfilment of the requirements for the degree of ERASMUS MUNDUS MASTER IN PUBLIC POLICY

Supervisors: Prof. Cameran Ashraf and Prof. Aina Gallego

Electronic Signature

I hereby certify that this dissertation contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I hereby grant to IBEI and the Mundus MAPP Consortium the non-exclusive license to archive and make accessible my dissertation in whole or in part in all forms of media, now or hereafter known. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation.

Name: Valentino Grassi

Signature: Ven Ceri

Location and Date: Barcelona, Spain July 31, 2020 Word-count: 12,415 words

Abstract

This dissertation tries to question the relevance and effectiveness of the most influential national strategies for Artificial Intelligence (AI) development. The rising number of official plans and visions published by public institutions signals dynamic and flourishing policymaking activities surrounding the sector of digital technologies, for a long time devoid of consistent regulations. Increased policy activity however must correspond to exhaustive provisions and a comprehensive understanding of all the actors involved in the process of delivering AI products. This thesis analyses the underlying logics sustaining the three different strategies and the subsequent regulatory frameworks provided, matching these factors with the structural capabilities of each state. It then supports this study by running a policy evaluation scheme, based on criteria of *coherence, comprehensiveness, and human-centredness* to assess their effectiveness and feasibility. By applying a *policy-first* research approach, the dissertation concludes by offering policy recommendations to improve the scope of AI strategies, finding a balance between regulation and innovation.

Acknowledgments

I would like first of all to thank my supervisors, Professor Cameran Ashraf from CEU and Professor Aina Gallego from IBEI, for their help and support in selecting and further developing the content of this dissertation. The academic quality of the two institutions I attended was incredible and the international environment promoted by them has been a crucial and beneficial part of my master in Global Public Policy.

I feel I cannot be grateful enough for all the efforts and support my parents devoted to me during my whole academic career. Always pushing me to follow my ideas, my successes and mistakes, but most importantly my desire to try to understand everything surrounding me. From small to bigger things, I believe they positively fostered my quest for a better understanding of society and its dynamics, which led me to explore one of the most unpredictable phenomena of our times.

This thesis goes also to all my friends, the ones that really had to deal with me during the process of writing this dissertation, but that also inspired me constantly and supported me no matter what. I cannot stress enough how important it is for me to have been in contact with so many new faces and their incredible and interesting life stories. All of them will always have a place in my mind, since they contributed to enrich my vision of the world, cultivating everyday the importance and beauty of diversity and discovery.

Table of Contents

Electronic Signature	i
Abstract	ii
Acknowledgments	iii
List of Tables	v
List of Abbreviations	vi
1. Introduction on Artificial Intelligence	1
2. Research Question and Methodology	3
3. Why AI Strategies?	5
4. The AI Arms Race Narrative	8
5. The European Union Vision of AI	11
5.1 Policy Framework	11
5.2 Capabilities of the European AI Ecosystem	13
5.3 Policy Evaluation	15
6. China's Strategy to Lead in AI	17
6.1 Policy Framework	17
6.2 Capabilities of the Chinese AI Ecosystem	19
6.3 Policy Evaluation	21
7. United States of America and its self-regulating AI Strategy	23
7.1 Policy Framework	23
7.2 Capabilities of USA AI Ecosystem	25
7.3 Policy Evaluation	
7.4 California's Privacy Regulation as a Catalyst for an EU-US Partnership on AI	27
8. Conclusions	
9. Bibliography	

List of Tables

Table 1 - European Union AI Strategy Evaluation	16
Table 2 - China AI Strategy Evaluation	
Table 3 - USA AI Strategy Evaluation	
Table 4 - Overall AI Strategies Comparison	

List of Abbreviations

AI	Artificial Intelligence
AIDP	New Generation Artificial Intelligence
	Development Plan
BD	Big Data
ССАР	California Consumer Privacy Act
ССР	Chinese Communist Party
EU	European Union
ML	Machine Learning
US	United States of America

1. Introduction on Artificial Intelligence

Big Data, Artificial Intelligence, and Machine Learning are creating enormous opportunities and almost infinite scenarios in basically every aspect of our society. However, academics and decisionmakers are afraid of what negative consequences such technological developments could produce. To understand the big picture, constantly developing and harder to control for policymakers, it is important to capture what these phenomena represent from a historical perspective.

Artificial Intelligence falls within the definition of general-purpose technologies (Benanti 2018; Helpman 1998), determined by a high degree of pervasiveness. Expanding and propagated at an exponential speed, AI affects industrial productiveness and the economy, subsequently entering every aspect of our life, transforming and re-shaping fundamental concepts and practices conceived so far (Allen 2019; Girasa 2020). From this perspective, we can interpret the industrial revolutions as fundamental steps marked by the creation and arrival of specific technological instruments (Desjardins 2019; Lim 2019). The first two revolutions were then characterized by the introduction of new means of energy production: first the creation of steam engines, then by the advent of chemical and electricity. Moreover, the third step is represented by the shift from stationary machinery to flexible ones, where a software component can adapt and shape hardware according to need (Benanti 2018; B. D. Mittelstadt et al. 2016; Lim 2019). Artificial intelligence becomes part of this last segment of the industrial revolution, but with a unique exception: for the first time the machine itself is the agent capable of setting the agenda. While the development of AI and its related elements is going further and further, the software is created to be more and more autonomous.

The revolution happening nowadays is based upon two fundamental pillars. First of all, as stated by Moore's law, the computational power is exponentially growing since the 50s, however, in the next few years, it will grow at an unprecedented pace, never encountered in human history (Brynjolfsson and McAfee 2016). Secondly, the increased availability and supply of mobile devices such as smartphones and notebooks, doubled the distribution of multimedia contents every two years, because 90% of current data available nowadays has been created in the last two years (Reinsel et al. 2018).

Artificial Intelligence can be described as the ultimate result of the combination of this series of unique and new factors in technological evolution that created the conditions for an almost natural rise of AI (Lim 2019). Thus, AI public policies have become more relevant in this sector, after years in which the private sector was ahead of policymakers in terms of knowledge and expertise. Nowadays, public institutions are trying to reduce this gap by publishing their key strategies and vision for AI, which, as this thesis will demonstrate, are crucial in shaping the direction of specific regional versions of AI.

2. Research Question and Methodology

The AI scenario is one of the most complex and yet interesting to research about, considering the exponential scale at which the technology is progressing, both in its design and implementation phases. However, taking into account the enormous capabilities enabled by these fundamental innovations (High-Level Expert Group on Artificial Intelligence 2019; Einav and Levin 2014), the policy strategies that key players are committing themselves to follow are what is going to define the future of AI.

This dissertation aims to look at the three most influential AI strategies, linking them to their real-world capabilities in terms of the regulatory framework, infrastructure, data availability, privacy culture, and investment trends. The central research question is asking whether nation-states are adequately designing strategies and policies to regulate and address the risks of AI developments, or if these plans are merely propagandistic documents. Evaluation criteria are then applied to assess the consistency and robustness of these approaches.

The main research tool consists of official document analysis and comparison (Verba 1967), to discover diverging and shared values and policies between the cases taken into consideration. Each strategy is then evaluated following three general criteria:

Coherence with the AI Ecosystem measures the extent to which states designed a vision for AI which takes into account the structural capabilities of its AI ecosystem, in terms of size, the maturity of the tech sector, and its legislative framework. Moreover, it tries to capture whether the goals set out by policymakers are propagandistic or instead feasible.

Scope of the Strategy aims at evaluating if the AI national plans are covering all the crucial and necessary aspects of AI, from design to implementation processes, if they include provisions and regulations for data-collection procedures. Finally, it takes into account whether the strategy is mainly addressed to domestic or foreign businesses.

Human-Centredness assesses the overall ethical and ideological aspects of the AI strategies, describing if they place human beings and individual citizens being affected by these technologies at the center of their development. Factors such as fairness and transparency requirements are taken into account, to understand whether AI is going to be developed to ultimately benefit its consumers or not.

Instead of using the more popular *technology-first approach*, through which a sector is analyzed based on the current technologies implemented and how they are performing, this dissertation is centered around a *policy-first approach* (Roberts et al. 2019). Looking at current technologies would most likely produce a short-term perspective on the AI phenomenon, soon becoming outdated following the high-speed pace at which AI technologies are being developed and implemented. By looking at what policy and decision-makers are doing and what their vision for AI is, it is possible to draw a more reliable and predictive picture of the Artificial Intelligence landscape. Moreover, the focus on different AI strategies is functional to show how AI is not a single identifiable "entity", with the same characteristics and uses everywhere. On the contrary, the diverging AI national strategies and their ecosystems represent three fundamentally asymmetric roadmaps for the evolution of this disruptive technology, with differing consequences depending on the current and future regulatory frameworks in place (Roberts et al. 2019; van den Berg and Keymolen 2017).

3. Why AI Strategies?

The importance and magnitude of the so-called AI revolution on every sector of the economy and its consequences for the global geopolitical order are stimulating a multitude of actors in pursuing a so-called "Race for AI" (Castro, McLaughlin, and Chivot 2019), often compared to the one occurred with the invention of nuclear power (Whittaker et al. 2018). This fact is represented by the number of states that, in the last decade, urged their ministries to draft, design, and publish specific AI strategies. It is important to stress that the competitive environment arising around AI indicates that a wait-and-see strategy could be dangerous and costly, both for private businesses and public institutions (Bughin 2018). Moreover, McKinsey calculated a global profit pool of \$1 trillion that the technology will foster and produce by 2030, suggesting that companies should begin to adopt AI at scale within the next 3 or 4 years (Chui 2017), the same reasoning applies to nation-states of course. According to PWC, business revenues from AI could exceed 38 billion dollars by 2030 and contributing around 12.8 trillion dollars to the whole global economy in the same year (PWC 2017). Most of the predictions seem to go in the same direction, with small differences between the various estimates of the future size of AI, signaling that the technology is a real opportunity, and an issue, impossible to avoid (Lim 2019).

Currently, many debates have arisen regarding the "umbrella term of AI" (Calo 2017), with a particular focus on ethical debates and risky future implementation of this technology (Lim 2019; Tolan et al. 2019; Parikh, Teeple, and Navathe 2019; van der Voort et al. 2019). Most of these researches tend to apply a technology-first approach, looking at early adoptions of AI to then raise concerns and analysis regarding the unintended effects of these projects or policies (Bennett Moses and Chan 2018; Thewissen and Rueda 2019; Lim 2019), without comprehensively addressing the policy aspects of the issue. There is a tendency towards producing policy recommendations and guidelines to foster a better and fairer AI design, however, often underestimating the importance of current and future policy trends signalled by public institutions.

The lack of comprehensive and broad comparison of the most influential strategies for AI can also be linked to the fact that only in the second part of 2020 all the official plans were finally published, with the EU delivering its White Paper in February. However, once again, researchers seem to have preferred a single-case approach to the study of AI, casually comparing some features to the more mature Chinese AI ecosystem (Roberts et al. 2020; B. D. Mittelstadt et al. 2016; Witt 2019). This dissertation wants to go one step further, by linking the dots of the global AI scenario, relying on official strategies and policy and industry trends of each actor. Finally, the possibility to build a "western democratic" partnership of AI is discussed in regards to the EU and the US, two players willing to cooperate but characterized by very different conceptions of AI and its regulatory framework.

At least 36 states from every continent have currently worked and published plans and guidelines on how to develop, invest, and implement AI (Future of Life Institute 2020). International organizations as well are producing their vision for an AI future, recognizing the importance of the technology by advocating for coordinated and cooperative strategies in this field of work. However, as we will see in regards to the European Union, possible conflicts between supranational and national guidelines could arise, especially in terms of ethical issues and privacy concerns (Andersen 2018). This trend further highlights the public attention that the topic of AI raised in the last years, with increasing public concerns regarding the technology, citizens, and consumers' association started asking for their governments to publicly communicate what their intentions are (B. Mittelstadt 2019). Negative externalities of early AI adoptions already hit citizens or businesses directly, bringing to light several issues that need to be addressed from the very first phases of AI development (Parikh, Teeple, and Navathe 2019; Floridi et al. 2020).

However, one of the first issues the "AI society" is facing is the one of overabundance and magnitude of the AI strategies itself. Meaning that this rise of AI as a central topic in governments and the media is resulting in many guidelines which sometimes set unfeasible and overly optimistic goals and objectives as if states were more concerned to say that they plan to develop AI rather than put this process into practice (Cath et al. 2017). At the same time, as pointed out by Calo, "*That artificial intelligence lacks a stable, consensus definition or instantiation complicates efforts to develop an appropriate policy infrastructure*" (Calo 2017).

This is a very important problem since it could create the conditions for a slow and soft approach to AI, something already occurred in many cases in regards to climate change strategies and policy goals published in the last one or two decades (van der Voort et al. 2019). This is why the goal of this section is, among the others, to assess whether the AI strategies taken into consideration, reflect real capabilities of the states that designed them in terms of infrastructure, investment trends, policies implemented, talent availability, and public acceptance.

4. The AI Arms Race Narrative

Before getting into the deeper analysis of the three major players' strategies in AI, it is crucial to understand how the debate has been approached in the last years since it has been fuelled by a specific narrative over the development of AI. As mentioned before, the impact of AI is shaking the policy debate, getting to the point where the pursue of automation itself is questioned by many, pointing the fingers at its controversial and negative outcomes on society (The Economist Intelligence Unit 2017; Tolan et al. 2019; Oh et al. 2017; van den Berg and Keymolen 2017).

This arena full of decisionmakers and policy-influencers indicating their AI-orientation reminds a so-called "arms race scenario", because of the intensity and rapidity of investments, and also because of the importance and emphasis placed by policymakers, political leaders and experts on the necessity to establish AI frameworks (Larsson 2020). As expected, the United States and China are at the moment the two major players on the stage, although with different strategies in place, reflecting their geopolitical and economic positions, being the largest investors overall in the sector (Mou 2019). For many experts, this transition towards an AI-integrated economy and society will be the key contest determining the next global leader, considering also the cultural and sociological changes that this technological breakthrough is expected to influence (Roff 2019). By no coincidence, the recent tariffs war ongoing between President Donald Trump and Xi Jinping has serious implications for goods and services related to the creation of AI solutions. The Huawei case, which saw the Trump administration accusing and banning the Chinese government-sponsored firm, is a clear example of the increasing tensions surrounding an ongoing tech-war (Inkster 2019; The Economist Intelligence Unit 2017)

To better understand the geopolitical and technological implications of AI, this dissertation runs a comparison of the three key strategies published by three vital key players in the AI sector: The United States of America, the People's Republic of China, and the European Union. These three actors do not only represent the most influential AI markets and development hubs, but each of them is also bringing a unique and different approach to AI, which in some cases present even strongly contrasting arguments, especially in regards to ethics and privacy (Larsson 2020). They also represent three different state and society models, which consequently translated their socioeconomic values int their vision for AI, and consequently,

in the plans they published. Moreover, by focusing solely on the contraposition of the US and China, there would have been the risk of reproducing once again a "Cold War Scenario", distorting the much broader spectrum of the AI geopolitical phenomenon (Horowitz 2018).

The "AI Arms Race" narrative (Manuel and Singh 2019; Roberts et al. 2020; Castro, McLaughlin, and Chivot 2019; Inkster 2019; Barkin 2020) itself it is, in fact, problematic, since it tends to oversimplify the very complex dynamics behind the transition to such a disruptive technological innovation (Girasa 2020; Tomasik 2013). In the last 5-7 years, there has been a clear focus on the speed and intensity of investments and adoption of AI technologies in both the US and China. Policymakers, experts, and the media highlighted some very specific and high-sounding statistics, such as the fact that China could overtake the US in AI patents by 2020, as analyzed by the UN agency for Intellectual Property. (World Intellectual Property Organization 2019). Much attention has been given to the highly-promising goals set by the Chinese Communist Party, such as the intention of becoming the global leader in AI by 2030 (Roberts et al. 2019).

This narrative approach is then often combined with a sort of fear for the Chinese approach to AI, and its non-democratic governance (Cath et al. 2017; Oh et al. 2017) contributes to intensifying a national security and military approach at the debate on technological development. Even though it is understandable that, finally, technology became to be considered as a highly crucial infrastructure and essential part of present society (Floridi et al. 2020) thus originating concerns regarding national security, these issues should not take away the attention from the overall context. A "resort to arms" scenario is likely to produce unsupervised and inhuman application of AI in the economy and social contexts. As the New York based AI Now Institute found after reviewing the debate surrounding the technology in the last decade: "Empirical factors like where AI produced in either country will eventually be deployed, what purpose such systems will be put to, whether they work, and which communities bear the risk of bias and other harms are rarely discussed within the *arms race* discourse" (Crawford et al. 2019).

Signs of resilience to the uncontrolled and misuse of AI had been launched by Google's employees, which in 2018 protested against the involvement of the company in a Pentagon program to use image recognition for military purposes (Shane and Wakabayashi 2018). However, it is not realistic to imagine that the faith of a "responsible" and non-harmful AI will solely rely on employees' actions, since for example in this case, after Google dropped out of

the contract, Microsoft and Amazon regularly signed it. "This view of progress tends to see all calls for restraint, reflection, and regulation as a strategic disadvantage to US national interest. It turns accountability into a barrier to progress and suppresses calls for oversight" (Crawford et al. 2019). Finally, if the West is worried about the Chinese approach to AI because of its authoritarian and oppressive practices (Roberts et al. 2020), however, it should use it as a reminder of how cautious the approach to AI in some aspects has to be, but instead, it is resulting in calls for faster implementation of AI, undermining problems and negative outcomes arising from it (Roff 2019).

The intersection and competition of these three main different strategies (together with the ones designed and implemented by other major players such as Singapore, South Korea, Japan, and so on) will determine how AI and innovation will develop in the next years. However, considering the strategic importance of these technologies, the whole process will continue to be inevitably shaped and influenced by geopolitics. The United Nations Interregional Crime and Justice Research Institute (UNICRI) started in 2015 a program regarding Artificial Intelligence and Robotics, followed by the recent opening of the UNICRI Centre for Artificial Intelligence and Robotics (International Telecommunication Union 2018). These "small" steps suggest that there would be the chance to cooperate at the supranational level to create an international regulatory framework for AI, BD, and ML, however, as already happened during the "nuclear revolution", serious tensions will inevitably arise, with greater chances to culminate in conflicts.

5. The European Union Vision of AI

5.1 Policy Framework

On February 2020, the Commission finally released the White Paper on AI (European Commission 2020), outlining its overall vision of the technology being considered, and explaining the intended roadmap to be followed in the incoming years. The first and probably most important feature of the European strategy is the human-centric approach which is set as a fundamental pillar, and also as a goal, of the overall development of AI. The plan calls for "a coordinated European approach on the human and ethical implications of AI as well as a reflection on the better use of big data for innovation" (European Commission 2020). The opening sentence of the strategy it is incredibly important in the context of all the other strategies already published, by defining AI as a human-centered technology, the Commission is essentially calling for all its member states to pursue a very specific approach.

The issues regarding ethics and real-life implementations of AI discussed in the first part of the dissertation seem then to be, at least in theory, well-addressed by the EU Commission, which places as well a clear focus on the source of algorithms: big data. The EU is aware of its disadvantaged position in AI, this is why, according to the Commission (European Commission 2019), it decided to pursue quality over quantity. This means that a more supervised and intentional development of AI is preferred by European decisionmakers rather than an uncontrolled and exponential rise of the sector in the continent.

For Artificial Intelligence tools to function and to be developed, a huge amount of data is required to train algorithms and machine learning systems (Azzone 2018; van der Voort et al. 2019). Often the analysis of the state of AI in the world tends to undermine the importance of data collection practices in each country (Vydra and Klievink 2019): different conceptions of privacy and human rights will deliver very different AI versions. Data and specific data "cultures" are exactly some of the factors that are determining an asymmetric advancement of AI in China, EU, and the US for example (Cath et al. 2017).

The first element of the overall EU approach to AI, but more in general digital technology, is of course the General Data Protection Regulation (GDPR). Published in 2016 and then implemented in 2018, is considered to be the most advanced and comprehensive data protection

policy currently available and impacting almost 900 million users and all the companies operating in the EU single market (Anisetti et al. 2018). It essentially expanded the definition of personal information, including also online identifiers such as IP addresses, and it increased the level of fines for organizations failing to comply (Voigt and Von dem Bussche 2017). Moreover, it tightened the consent rules governing data collection and it includes for the first time the "right to be forgotten", enabling users to request the erasure of their data from company records.

Article 5 of the Regulation outlines six main principles to be followed when dealing with data, that shall be:

- a. Processed lawfully, fairly and in a transparent manner
- b. Adequate, relevant and limited to what is necessary
- c. Retained only for as long as necessary
- d. Collected for specified, explicit and legitimate purposes
- e. Accurate and, where necessary, kept up to date
- f. Processed in an appropriate manner to maintain security.

More specifically, principles a, b, e are the most impactful for the AI industry and for the approach that the EU has decided for them to follow. If point *a* is a structural requirement for data managers, that must follow EU laws, the other two points considered address very specific issues of data collection practices. A "limited to what is necessary" use and collection of data have the aim of addressing a crucial problem of big data and AI, which is the "over-collection of data" (Elahi, Wang, and Xie 2017). Mobile devices had been, in many cases, working as "trojan horses", requiring the users to permit GPS positions or microphone access to receive certain services (such as driving maps apps) but then collecting more data than necessary and using them to profile the user and his habits (Elahi, Wang, and Xie 2017). Thus, service providers are now required to delimit and define the scope of their systems, to use only the right amount of data needed to perform a specific action. However, this specific requirement has been criticized for being too vague and problematic, since it lacks a proper and crucially necessary clear definition of what an "adequate, relevant, and limited" use of data is (Kazim and Soares Koshiyama 2020; Veale 2020). These guidelines could thus have the opposite effect of fostering innovation, by pushing developers to design AI tools which could then be labelled as unfair following the interpretation of misleading requirements.

Finally, point *e* acts as a standard quality requirement for big datasets, requiring data collectors to use only accurate (labelled and transparent) data, but most importantly to update them. These two requirements were designed to give a response to ethical questions regarding the composition itself of datasets, which were often found to be filled with inaccurate data (Desouza and Jacob 2017; Natarajan, Frenzel, and Smaltz 2017) or biased data, for example, underrepresenting minorities and reinforcing prejudicial inequalities (Natarajan, Frenzel, and Smaltz 2017; B. D. Mittelstadt et al. 2016; van der Voort et al. 2019). Point *e* thus seems to correctly address the issue of dataset composition, however, further technical clarification of the "label and transparency" requirements is missing once again.

5.2 Capabilities of the European AI Ecosystem

The strategy released by the Commission in February addresses the necessity of both a policy and a regulatory framework, outlining step by step the vision of the Union on this broad issue. The main problem the Union needs to overcome is undoubtedly the fragmentation of its institutional structure: The Member States need in fact to be receptive and willing to adopt this strategy. However, often the most sensitive issues tend to be politicized at the level of the European Council; even in the case this matter would be treated as a "technical" one, slow and scattered implementation would hinder the benefits of the strategy (European Commission 2016; Kazim and Soares Koshiyama 2020). To overcome these obstacles, the White Paper outlines two crucial principles:

- An ecosystem of Excellence representing the Policy Framework
- An ecosystem of Trust representing the Regulatory Framework

At the center of these two pillars, there is overwhelming attention to the final users and the citizens overall, which at first sight seems promising and decisive in its approach, however further analysis from several experts tends to highlight the weaknesses of some policy strategies. First of all, the Commission, in building an Ecosystem of Excellence, targets the aim of attracting over 20 billion of investment in the sector per year (European Commission 2020), supporting it with several EU funding programs, such as the Digital Europe Programme, Horizon Europe and the European Structural and Investment Fund. However, the approach taken by the High-Level Expert Group called to draft the plan seems to be erroneously focused solely on the technical and industry side of the phenomenon. The overall policy framework

does not address the necessity for buyers (private and public) to acquire analytical skills that are fundamental for assessing the technological needs of their business and to understand the benefits of AI systems' integration (Veale 2020). Thus the link between offer and demand is absent from the White Paper, and combined with the lack of specific infrastructure plans, except for the creation of "excellence and testing centers for research", the provisions seem to be pretty vague (Kazim and Soares Koshiyama 2020). In this regard, the main flaw could be identified as the lack of plans or strategies to develop and eventually regulate European datacentres: facilities that are mostly entirely provided by private businesses (Castro, McLaughlin, and Chivot 2019). In this way, the EU is failing to capitalize on the opportunity to provide the real and safe infrastructure needed by AI and Machine Learning systems to be trained and for Big Data sets to be stored, losing the chance to provide clear and structured environmental requirements to be more sustainable (Masanet, Shehabi, and Koomey 2013).

It is, however, fair to note the fact that, the democratic value-standpoint adopted by the EU, especially in the GDPR, appear to have tangible effects in influencing the overall market and other policymakers in drafting their strategies and services, thus with the EU functioning as a sort of standard-setting institution in terms of ethics (Floridi et al. 2020; Castro, McLaughlin, and Chivot 2019).

On the other hand, the regulatory framework presents the pros and cons, even though it tries to include new elements to the legal debates surrounding AI (Nicholson Price II 2018). Full compliance with EU rules is it, of course, the backbone of this framework, deriving from them the "Seven Key Requirements of AI":

- a. Human agency and oversight
- b. Technical robustness and safety
- c. Privacy and data governance
- d. Transparency
- e. Diversity, non-discrimination, and fairness
- f. Societal and environmental wellbeing
- g. Accountability

While the requirements itself are necessary and human-centered, the Commission still must update the EU legal framework to include the whole AI development process under its safeguards, while limiting its scope to the release of the product or service on the market (Kazim and Soares Koshiyama 2020). Moreover, the strategy explicitly indicates high-risk sectors within AI, such as healthcare, transport, energy, and parts of the public sector, which essentially represents areas in which technological tools have a tangible and mostly direct human impact. Doing so, the Commission runs the risk of transforming AI as a "high-cost venture", encouraging risk aversion approaches and less willing to invest in those sectors (Kazim and Soares Koshiyama 2020; Veale 2020). Experts suggest on the contrary that such a compliance mechanism could be not beneficial for fostering a sustained AI development, while a real-time intervention and governance setting would be more feasible and effective (Cath et al. 2017).

5.3 Policy Evaluation

As repeatedly stated, the European Approach to AI is probably the most socially-beneficial one in terms of the initial and fundamental perspectives on the human role in the overall process of AI integration and development. The White Paper seems to successfully understand the crucial needs of the EU framework and responding to citizens' calls for a fair and reasonable evolution of algorithms (B. D. Mittelstadt et al. 2016; Cath et al. 2017). Good intentions nevertheless require strong and effective measures to be realized, combined with the intrinsic difficulty of dealing with decision-making and investment fragmentation, the provisions recommended by the Commission are very poor in terms of procedural arrangements. Moreover, it is important to recall that most of the Member States are also drafting their own AI strategies, with a variety of conflicting and asymmetric objectives which could cause further frictions and conflicts (Larsson 2020).

The risk-based approach chosen by the Commission is considered to be very controversial since it could limit the potentialities of AI in terms of investment attraction: on one hand, the EU wants to attract more funds from local and foreign actors, on the other one however it is not providing an effective and fast-adapting monitoring and control system over AI development (Veale 2020). The goal of EU policymakers should be more consistent in building a fair but also feasible environment for AI to grow in, instead of focusing its regulatory efforts mostly on the design and delivery stages. To be encouraged, the investment needs to be directed towards the AI sector trying to balance the scope of simplification measures, aimed at attracting capital and talent, and the need for more supervision over the technologies developed, something that currently the Commission is failing to do, placing most of the *costs of risk* on AI producers and distributors (Liu, Lin, and Chen 2019; Katyal 2019).

In terms of capabilities, according to the Center for Data Innovation the European Union, in comparison with the US and China, scores very high in the metrics regarding the availability of AI talent, which however is dispersed since these people are often attracted by higher salaries and opportunities in the US market (Castro, McLaughlin, and Chivot 2019). The Commission plan does not address sufficiently the need for better funding and attraction of talents, and a crucial multidisciplinary approach to the field. By requiring more "ethics figures" in the AI industry, the recommendations do not provide a clear understanding of how the sector and the academic framework should re-organize to include and foster the development of such figures (Veale 2020). Moreover, relying on funding institutions already in place, it is relatively harder for the EU to centrally direct funds towards specific AI projects, worsening the fragmentation of finances intended for Artificial Intelligence.

EU AI Strategy

Policy Evaluation

Coherence with the AI Ecosystem

Very coherent with the logic and requirements of GDPR

It does not address the importance of AI talents leadership of the EU

It does not solve problems related to policy fragmentation of member states

Scope of the Strategy

It extends to all domestic and foreign AI actors

It takes into account every aspect, from design to implementation of AI technologies

It succesfully address data-collecting and dataownership

Human-Centredness

It applies all EU legal and constitutional human rights and democratic values and regulations

It focuses most of its attention on final users/citizens

It defines high-risk AI applications to mitigate negative consequences



6. China's Strategy to Lead in AI

"AI is a vital driving force for a new round of technological revolution and industrial transformation, and accelerating AI development is a strategic issue to decide whether we can grasp opportunities"

With this few words, General Secretary of the Communist Party of China Xi Jinping told the Politburo and its key decision-makers in 2018 why its country was moving so fast and powerfully in developing a leading AI sector. In response to the trade war with the US and its President Donald Trump, the Chinese leadership is well aware that its fast rates of growth and the transition from a heavy industry and manufacturing economy to a more services-oriented one are not going to happen without serious negative externalities (Inkster 2019). As many analysts foresee, labor force disruptions and the role of healthcare are just some of the enormous challenges that the second-largest economy is going to face in the present and the nearest future (Daxue Consulting 2020).

6.1 Policy Framework

The most important strategic document issued by Chinese officials is undoubtedly the 2017 "New Generation Artificial Intelligence Development Plan" or AIDP, which, contrary to many expectations, provides a non-centrally enacted initiative (Sheehan 2018). It is based on a strong direction and vision provided by the Central Committee, which however has the aim of functioning as a "wish list" for private businesses and local governments. The actual innovation and industrial transformation are going to be driven by the two actors aforementioned, through a system of incentives, huge tax benefits, and party promotions based on economic performances (Roberts et al. 2020). However, China seems to be characterized by a more coherent, even if debatable, approach to AI as a whole, with the additional public-private partnerships initiated with the selection of "AI Champions": private firms in charge of developing specific and very influential plans, counting now 15 national champions since 2017 ("China Adds Huawei, Hikvision to Team Spearheading Country's AI Efforts" 2019).

The plan is very well structured, and it prescribes three crucial areas of focus according to the AIDP: international competition, economic development, and social governance (China

Institute for Science and Technology Policy at Tsinghua University 2018). Since the 1990s, Chinese military policy is mainly based on the "Shashoujian" philosophy, which pursues an asymmetric competition approach to the US: instead of competing on similar capabilities, the Chinese army aims at developing new and asymmetric ones to surpass its competitors (Roberts et al. 2020). This strategy seems to be applied to the AI sector as well, where the economic and military global supremacy needs to be achieved with the use of alternative or new technology, such as Artificial Intelligence. Thus, this approach is supported by the AIDP provisions for *international competition*, privileging technology developments rather than increased spending in the "traditional" sectors of the army and the economy.

In the *economic development* section, as repeatedly stated by Xi Jinping, there is a strong focus on the new round of industrial revolution happening with AI. Policymakers are well aware that the short-term consequences of (early) AI adoptions will partially disrupt the Chinese workforce, further exacerbating the growing domestic inequality rates (Stiglitz and Korinek 2017). However, with the AIDP, the government aims at grabbing that 26% boost in GDP by 2030 through AI implementation forecasted by many analysts, such as PWC (PWC 2017). To boost growth and mitigate negative externalities, and to reduce the gap in terms of the aforementioned lack of AI talents, in addition to the AIDP, the government is implementing structural reforms also in its educational system. With the "National Medium- and Long-term Education Reform and Development Plan" the country "seems to be preparing better than other middle-income countries to deal with the longer-term challenges of automation" (Roberts et al. 2020).

The Social Credit System is probably the most (in)famous Chinese implementation of AI: a credit-score system, based on behavioral data gathered and processed through AI, to determine whether citizens are acting in a "good or bad" way, thus sanctioning the low-score citizens and "promoting" the high-ranking ones (Grassi 2018). China plans to further increase the use of AI to empower "its potential for understanding group cognition and psychology", relying also on the technology integration in the rising number of smart cities and the internet of things infrastructures.

This is however only one aspect of the *social governance* implications of AI since greater importance is given to the healthcare sector, which is going to test the Chinese government's capabilities, facing an aging population and increasing access to health services. For this reasons, the "Healthy China 2030" strategy aims at achieving through AI a switch from

treatment to prevention, to sustain the growing necessities of its population (Natarajan, Frenzel, and Smaltz 2017; EIT Health and McKinsey & Company 2020).

6.2 Capabilities of the Chinese AI Ecosystem

Crucial to the study and research of Chinese policy and technological trend to understand what is the state of the underlying ecosystem and context surrounding a specific sector, avoiding biased or under/overestimating predictions (Mann 1999). In the case of Artificial Intelligence, the PR strategy is pretty clear: the CCP tended to state very optimistic and far-reaching goals in their declarations and official documents, focusing less on the problematics that this technological revolution should seriously address (Larsson 2020). Adopting a policy-first approach then results in a better prediction and understanding of what China intends to do with AI since a technology-first perspective could easily be outdated following the rapid and exponentially increasing adoption of AI systems (Ding 2018).

One of the main strengths of the Chinese AI ecosystem is the aforementioned top-down approach for the proliferation of the technology adopted by the ruling party, which results in more predictability for domestic and foreign investments and the rising amount of economic benefits for early adopters and R&D plans (Castro, McLaughlin, and Chivot 2019). In terms of importance, however, the real comparative advantage derives from the structural fact that Chinese users essentially "skipped" several technological development trends, shifting from a mainly rural society to a highly tech-based one. China is the country with the highest number of active internet users in the world, roughly 900 million, and a smartphone penetration rate of 60%, meaning that at least 851 million people possess a smart device (Daxue Consulting 2020). This record-numbers in terms of active users simply characterize the main strength of the overall AI exosystem, which is the availability and easy access to the huge amount of consumers data that is produced daily (Roberts et al. 2020). To develop and train AI and Machine Learning systems it is crucial to have abundant data, and especially so-called "pervasive data": datasets with very specific information regarding the user/device, such as location, clicks, and actions, and so on. Because of its smartphone usage rates, China is producing an enormous amount of big data, filled with deeper information (collected through mobile devices) which are easily accessible by the private and public sectors, as a result of lax and vague privacy and data protection regulations (Ding 2018).

In contrast to the EU GDPR, where companies need to put effort and money to lawfully collect data, Chinese companies have way less oversight from public authorities; moreover, China also "closed" its ecosystem, requiring companies to share data only domestically and forcing foreign companies (such as Apple) to hold Chinese users' data in Chinese datacentres (Vydra and Klievink 2019). All these factors create probably the best ecosystem for a rapid and sustained development of Artificial Intelligence overall, without taking into consideration ethical and privacy issues of course. In addition to all of these factors, the Chinese AI market is also supported by a large use of Government Guidance Funds (GGF), set up by state-owned companies and local governments to invest heavily in domestic AI start-ups and foreign acquisition for technology transfer strategies (Castro, McLaughlin, and Chivot 2019).

However, the Chinese ecosystem also presents serious weak points that policymakers and market actors will inevitably face and possibly try to fix soon. If the country has more than abundant data resources, it is lacking the human capital to succeed globally in the AI sector. China in fact can count on a talent pool of nearly 39,000 AI researchers, which is less than half of the US one, which counts roughly 78,000 researchers (Ding 2018). US AI workers tend in fact to graduate in leading universities and to gain at least 10 years of experience in their field, thus putting pressure on CCP leaders to restructure and direct more investments in the education sector, which however needs more years than other policies to show significant results (China Institute for Science and Technology Policy at Tsinghua University 2018; Kroeskop 2018).

Another common oversimplification of China usually comes up when analyzing its internal market structure, often understood as a highly-oligarchic one where the Communist Party simply decides who wins and who does not, especially when we think about the "AI National Champions" initiative (Meissner 2017). On the contrary, even though the central government has the authority to create monopolies and oligarchic structures, further studies showed how the Chinese domestic market is characterized by brutal and crowded competition, both from domestic and foreign companies (Lee 2018; Ding 2018; Sheehan 2018; Daxue Consulting 2020). This highly competitive framework it is viewed by many as a factor adding complexity to the sustained development of AI technologies, and the many peculiarities present in this specific context further highlights the fact that rather than competing in a "race", the US, the EU, and China are essentially developing different and asymmetric versions of Artificial Intelligence technologies (Roberts et al. 2020).

6.3 Policy Evaluation

A policy-first approach to the case of China reveals that its strategy and AI ecosystem, rather than more advanced or mature, are just very different from the United States and the European Union. The central government plays the role of a catalyst, driving private and local/public investments more in the direction of AI projects. Compared to the other two models taken into consideration, China is the one investing more resources and policymaking efforts in designing and creating the "perfect" context for AI proliferation based on Chinese capabilities, both in terms of infrastructures and users' behaviors. However, some of its strengths, such as data abundance and lax privacy laws, could easily become very dangerous weaknesses, running the risk of highlighting social tensions (such as inequality and age gaps) and trusting AI "too much", by not addressing quickly enough its potential negative externalities (B. D. Mittelstadt et al. 2016; Roberts et al. 2020). On the other hand, the willingness to adopt and test innovative policies and technological innovations in sensitive sectors, such as healthcare, if successful, could determine whether or not China will dominate the AI global competition.

In regards to innovation and (self)regulation, the Chinese case presents a strategy in which regulation is expected to be designed only after a rather high (or full) potential of AI is reached. Thus, in the first phase quantity of patents and inventions is preferred and stimulated by lax regulations, similar to what is happening in the US, however, China seems to underestimate the "export-factor" of AI (Puller 2006). To fully unleash the economic benefits of a wide AI implementation, the Chinese AI sector needs to take into account foreign legislative frameworks and technical standards, to then properly enter in those markets. The logic behind the EU White Paper on Human-centred approach is to work as a standard-setting document not only domestically, but challenging foreign visions for AI, by requiring external actors to adhere to EU moral and technical standards.

If China will pursue solely a self-regulating approach to AI it runs the risk of producing state of the art innovations with few chances to be implemented in the US or EU (Floridi et al. 2020). At the same time, the Chinese strategy is the one that best addresses and draws on its strength-points and it is coherent with the market and industry structure currently available (Roberts et al. 2020). With its "wish list" AI approach, the CCP is presenting AI to market actors mostly as an economic opportunity, considering the fiscal and direct benefits awarded to firms

developing this technology. However, this "philosophy" could also contribute to the creation of an "AI bubble" for investors: revenues could be lower than the ones currently forecasted, especially in the case where strong requirements and restrictions on AI tools will be in place in the next future.

China AI Strategy

Policy Evaluation

Coherence with the AI Ecosystem

Very coherent in regards to lax privacy regulations and datacollection methods

It correctly takes into account the capabilities of private sectors giants and start-ups

It is based on the actual infrastructure and investment capabilities of the sector



Table 2 - China AI Strategy Evaluation

Scope of the Strategy

It addresses all the relevant sectors that will benefit the most from AI

Its wish-list structure places private firms in an almost leading role

It does not address issues arising when exporting Chinese AI tools to foreign markets



Human-Centredness

Indivudal rights traded for greater overall societal benefits

It does not protect final users from AI outcomes on legal basis

It specifically aims at steering individual behavior through AI



7. United States of America and its self-regulating AI Strategy

7.1 Policy Framework

The United States, compared to other countries, did not have a fully coordinated AI strategy until 2019, when under President Trump the *American AI Initiative* was launched, to collect and coordinate all the state-led initiatives to boost AI development (Soni 2020). However, since the last year of Obama's second mandate, Artificial Intelligence became one of the main priorities of US governments.

The strategy published on February 11, 2019, is based on five central pillars:

- 1. Pursuing technological breakthroughs
- 2. Ensure the development of adequate technical standards
- 3. Training workers with skills useful to develop and implement AI technologies
- 4. Safeguarding American values, such as civil liberties and privacy, while fostering public trust in AI
- 5. Defending US technological advantage in AI, while promoting an innovationsupportive international environment

This plan differs from the Chinese one since instead of working as a "wish-list" for the private sector, it calls for each federal agency to draft guidelines, regulatory and non-regulatory approaches to guide the use and development of AI in every sector (Future of Life Institute 2020). The American Initiative was then reinforced by the further announcement of ten AI principles, issued by the With House Office of Science and Technology Policy (OSTP):

- 1. Public trust in AI
- 2. Public participation
- 3. Scientific Integrity and information quality
- 4. Risk assessment and management
- 5. Benefits and costs
- 6. Flexibility
- 7. Fairness and non-discrimination
- 8. Disclosure and transparency
- 9. Safety and security

10. Interagency coordination

These principles and recommendations seem to be in line with those published by the European Union Commission as well since both strategies tend to apply a risk-assessment approach (Cath et al. 2017; Floridi et al. 2020). They focus in fact on the role of local and federal agencies to assess whether an application of AI can be defined as high-risk or low-risk, consequently applying different degrees of governmental control over the development of the technology. This approach is however criticized by some experts (Veale 2020; Kazim and Soares Koshiyama 2020) because it could have serious negative effects on AI proliferation, making investments in the sector less reliable and riskier than others. For this reason, the US white paper insists various times on the concept of *flexibility*, intending to produce guidelines that can then be re-adjusted depending on the context (Castro, McLaughlin, and Chivot 2019). This is crucial considering the US AI ecosystem solely based on the aforementioned capabilities of the private tech sector and its close, but problematic, relationship with the US Army, which is constantly increasing its spending on military AI (Shane and Wakabayashi 2018; Oh et al. 2017). Besides, these pillars and principles present the same flaws of the EU counterparts: they are broad, vague, and regulatory bodies are being very slow in designing and implementing these requirements.

Of all the mild and vague recommendations given by the White House, a real key pillar of the whole strategy is the call for the construction of an International Collaboration on AI with likeminded partners. The US seems in fact to recognize the value of cooperation and shared strategies to lead in such a competitive sector, which however is in contradiction with an overall international disengagement strategy adopted under Trump's presidency (Robinson and Thierfelder 2019). In this case, the States are calling for partners such as the EU, Korea, and Japan to join them in creating a "western AI loose alliance" (Tomasik 2013) to counterbalance the rising role of Chinese AI developments. The synergy between the US and EU strategies could then function as a catalyst for future AI partnerships, in the landscape of an ongoing trade war against China, the US tech sector could come closer and closer to the Old Continent once again.

There are also serious divergencies between these two actors since the EU took a very bold and specific approach to regulation in regards to digital businesses. Starting from taxation (Guarascio 2018), the Union fiercely opposed the hegemony of the few tech giants (Google, Facebook, etc), all founded and based in the US, engaging in several important trials at

European courts. Since 2015, EU Commissioner responsible for media and information, Margrethe Verstager, has been very influential in determining the European approach to the IT sector, which finally led to the design and implementation of the famous GDPR (Witt 2019). The Privacy Regulation itself could represent a bigger obstacle for the creation of a strong partnership for AI since both European and foreign businesses are required to adhere to the strict requirements provided in the regulatory framework. Thus, on one hand, the vagueness of the regulations issued by both the EU and the US could work as a catalyst for building concrete and shared vision on AI, on the other hand however this factor could produce two incompatible versions of AI ecosystems, obstructing the creation of the so-called *International AI Environment*.

7.2 Capabilities of USA AI Ecosystem

After the internet, "born" in 1969 at UCLA, US researchers were also the first ones to develop AI systems, such as computer scientist John McCarthy who coined the term "Artificial Intelligence" back in 1956 at the legendary Dartmouth Conference. Nowadays, the US is still considered to be the AI global leader (Castro, McLaughlin, and Chivot 2019), since it excels in the total amount of investments, both public and private, in the sector and it maintains the highest rankings in terms of research and development publications and projects. Finally, the US holds a crucial advantage in the semiconductor and chipmaking industries, and its AI talent is more elite compared to the other competitors (Castro, McLaughlin, and Chivot 2019; Roff 2019).

However, compared to its direct competitor China, the US had a confusing approach to a centrally planned strategy for AI, since both presidents Obama and Trump initially refused to invest policymaking resources to produce a plan. This can be explained by the untold intention of the States to not impede the rapid development of AI technologies to catch up with the enormous number of AI patents issued every single month by Chinese researchers (Larsson 2020). The American approach to this technology is focused around the supremacy of private businesses and AI start-ups to drive the flourishment of artificial intelligence, counting on the hegemony and resources of tech giants such as Alphabet Inc. (Google), Amazon, IBM, Microsoft, Apple and so on (Floridi et al. 2020). Moreover, the underlining logic of US policymakers, coherent with the American culture on private entrepreneurship and capitalism,

is to make the private AI sector self-regulating (Foster 2008), since they know what standards are the best ones for the technology to remain profitable and sustainable. However, this approach is based on the "false" assumption that private businesses take into account factors such as transparency, fairness, and ethics on the same level of importance as profit (Puller 2006). It is very risky to believe that profit will be sacrificed in the name of justice, as more and more cases of misconduct by tech companies are investigated by public agencies (Foroohar 2019).

7.3 Policy Evaluation

Trump administration's plan thus reflects this stance, since it produced a strategy that "makes everybody happy": on one hand these ethical guidelines and "requirements" have the aim of responding to public calls for greater AI accountability and fairness, on the other hand, the private sector has been reassured that innovation would not be obstructed by excessive and rigid regulations (Girasa 2020). The overall strategy can be then summarised as a private-self-regulated framework, where who develops and implements AI is the one tasked with designing rules and limits to its scope and the number of risks taken for it to work. Once again, the "AI Race Narrative" discussed previously seems to be the logic behind this type of approach, since the direction taken by the government appears to be more concerned about the quantity, rather than the quality and fairness, of AI "production", to keep the pace with China (Castro, McLaughlin, and Chivot 2019). However, because of the structural differences between the two societies, this strategy could be more dangerous for Americans rather than for its competitors, with the risk of developing a version of AI incompatible with the democratic and ethical foundation of US society (Roff 2019; Tomasik 2013).

In the current scenario, there are few structural possibilities for the US to cooperate with the EU in terms of AI development since the concept of flexibility is very different among the two regulatory systems (Floridi et al. 2020). The privacy culture is also radically distinct in the two continents, starting with the fact that there is no constitutional provision in the States that acknowledges and protects individual privacy (Landwehr et al. 2012), while the 4th Amendment is only concerned with government's intrusions in certain situations (Barret 2019).

7.4 California's Privacy Regulation as a Catalyst for an EU-US Partnership on AI

The various structural factors seem to suggest a difficult conversation between the US and the EU to find a common framework to foster AI cooperation activities. In 2018 however, the 5th economy in the world, California, issued a new privacy regulatory framework described as very similar to the GDPR, at least in terms of its underlying logic. The California Consumer Privacy Act (CCPA) went into effect on January 1, 2020, but since its announcement has already produced the result of influencing the general public to ask for more accountability and control over individual data by US citizens and users (Soni 2020). California is considered to be a very influential public policy player in the States, usually leading the way with progressive and innovative reforms, such as the one regarding the highly-debated prison system (Petersilia 2008).

The direction the Californian parliament decided to take in regards to privacy and tech company has even more structural consequences on the United States because almost every single major tech business is based in the *Golden State*, with Silicon Valley being the densest conglomerate of tech companies in the World (over 2,000) (Castro, McLaughlin, and Chivot 2019). Thus, the CCAP is estimated to affect almost 40 million residents and over half a million US companies, which makes the magnitude of this bill global, in the same way, the European regulation did. Both acts target every business and data-collection activity conducted on respectively Californian and European soil, consequently broadening the scope of the two frameworks. For these reasons, considering the firms already complying with GDPR and now required to act according to the CCAP as well, these two regulations are considered to be "de facto global standards for data privacy and protection" (Barret 2019), especially in the Western area of the globe.

The big step forward taken by California, if proved successful in balancing the overwhelming presence of data-collecting companies in the State, should be taken as a baseline standard by the central government as well, being the most solid way to pursue to build the desired "international AI environment" prescribed by the American AI Initiative. To do so, policymakers would need to align the US and Federal States' Constitutional provisions to the one introduced in 1972 by Californians, which prescribe the right of privacy as one of the inalienable rights of all citizens (Floridi et al. 2020; Barret 2019). In addition to this, is worth noticing that most of the guidelines outlined by the American AI Initiative are already

contained in the CCAP, placing US decisionmakers in the position to either put into practice the governmental white paper or not.

USA AI Strategy

Policy Evaluation

Coherence with the AI Ecosystem

It coherenlty takes into account US tech sector global leadership

It does not address data infrastructure management

It stresses the importance of competing with foreign actors such as China



Table 3 - USA AI Strategy Evaluation

Scope of the Strategy

It applies to almost every federal agency

It calls for building an international AI environment based on shared values

It does not address effectively private sector self-regulation



Human-Centredness

It is theoretically in line with the EU general guidelines on human rights

It does not properly assess the risks related to AI design rather than implementation

Difficult to protect privacy without constitutional provision of privacy as a right



8. Conclusions



AI Strategies Comparison

Table 4 - Overall AI Strategies Comparison

One of the aims of this thesis is to challenge the popular idea that innovations and new technologies are nearly impossible to regulate, and the rise of AI reinforced this narrative. However, policymakers and governments regulating new tech are nothing new, and usually, early regulations have helped to produce successful results (van den Berg and Keymolen 2017). Extensive research on the negative consequences of improper uses of AI has been very popular in the last decade, however, the impacts of this technology ultimately rely on humans. These risky outputs affecting society do not depend solely on the "black box" code underlining them, but mainly on who is the final user and for what purposes AI is used. All these aspects can be safely regulated by governments, and this dissertation aimed at evaluating how these actors are doing so.

From the structured comparison carried out in the thesis, it is clear that the most influential players in AI still need to improve their policy design and decision-making efforts. None of the cases taken into consideration scores sufficiently in all of the three criteria being used for the evaluation. Each strategy tries to find a balance between regulation and innovation, following specific priorities depending on the structural factors characterizing the country.

The European Union scores overall better than its counterparts since its *White Paper on Artificial Intelligence*, combined with the provisions outlined in the GDPR, represent the most

comprehensive plan of all the ones analyzed. It places individuals and citizens at the core of its AI vision, while at the same time trying to address all the actors involved in every process of technology design, development, and implementation. It fails however to build a strategy consistently based on its real capabilities, setting goals that seem to be difficult to realize considering the problem of policy coordination among the EU Member States and the lack of sufficient domestic and foreign investments. The strategy planned by the Commission sets a solid and fair base to build a robust AI ecosystem, however, more efforts are desperately needed to build a tech market influential and profitable enough to empower the large amount of EU AI talent which is currently dispersed in foreign markets.

The case of China tells us that the authoritarian and centralized political structure of the Communist Party is surprisingly producing innovative and updated policymakers, capable of anticipating their counterparts in many aspects of digital technologies regulation. From a technical perspective, the strategy drafted by the second economy in the world is effective in building a comprehensive and functioning ecosystem surrounding the AI sector. From stimulating huge private/public investments in promising technologies to empowering small start-ups and supporting them with massive sources of data, the Chinese government is probably most aware of where it needs to direct its efforts to build a leading AI industry. However, concerns regarding the absence of democratic and human-centered values in their strategy still reinforce the idea of future use of the technology as a tool to maintain a sociopolitical status-quo. Once again, the destiny of Chinese AI technologies will be determined by the rise (or not) of an international AI environment and commonly-agreed standards for the industry.

Finally, the United States of America, currently considered to be the global leader in AI, basically still needs to take a bold position on what vision they have for the sector. The strategy presented by the Trump administration obtains the lower score in the comparison, mainly because US regulators are afraid of losing their leadership by overregulating the private sector. The American economy structure and research capabilities are both very reliant on private investments, and in the field of AI, considering that US tech giants are essentially the main developers of these technologies. However, by not providing a clear vision of AI, the States run the risk of dealing with consistent regulations when AI will be way more mature and embedded in the market structure. Moreover, a focus on the previously analyzed *Arms Race Narrative* is likely to produce an incompatible and divisive version of AI in different environments, reducing the beneficial potentials of this technology.

Overall, states need to considerably improve their policy strategies for AI, by taking into account the concrete risks of uncontrolled and unpredictable development of this disruptive technology. A coordinated effort to address infrastructural needs and capabilities is required to sustain a robust and fair AI ecosystem, starting from data collection. Once again, the GDPR should be further enhanced and set as a global standard (Barret 2019), since it is capable to address the first place in which bias and other negative consequences arise. Finally, an international consensus and cooperative AI environment are inevitably necessary, but the current state of national AI strategies signals a worrying lack of coherent domestic visions of Artificial Intelligence.

9. Bibliography

- Allen, Gregory C. 2019. Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security. Center for a New American Security Washington, DC.
- Andersen, L. 2018. "Human Rights in the Age of Artificial Intelligence." Access Now.
- Anisetti, Marco, Claudio Ardagna, Valerio Bellandi, Marco Cremonini, Fulvio Frati, and Ernesto Damiani. 2018. "Privacy-Aware Big Data Analytics as a Service for Public Health Policies in Smart Cities." *Sustainable Cities and Society* 39 (May): 68–77. https://doi.org/10.1016/j.scs.2017.12.019.
- Azzone, Giovanni. 2018. "Big Data and Public Policies: Opportunities and Challenges." *Statistics & Probability Letters* 136 (May): 116–20. https://doi.org/10.1016/j.spl.2018.02.022.
- Barkin, Noah. 2020. "Export Controls and the US-China Tech War." *Mercator Institute for China Studies*, March, 10.
- Barret, Catherine. 2019. "Are the EU GDPR and the California CCPA Becoming the de Facto Global Standards for Data Privacy and Protection?" *Scitech Lawyer* 15.3.
- Benanti, Paolo. 2018. *Le Macchine Sapienti: Intelligenze Artificiali e Decisioni Umane*. I edizione. Bologna: Marietti 1820.
- Bennett Moses, Lyria, and Janet Chan. 2018. "Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability." *Policing and Society* 28 (7): 806–22. https://doi.org/10.1080/10439463.2016.1253695.
- Berg, Bibi van den, and Esther Keymolen. 2017. "Regulating Security on the Internet: Control versus Trust." International Review of Law, Computers & Technology 31 (2): 188–205. https://doi.org/10.1080/13600869.2017.1298504.
- Bughin, Jacques. 2018. "Wait-and-See Could Be a Costly AI Strategy." *MIT Sloan Management Review* 59 (4): 1–4.
- Calo, Ryan. 2017. "Artificial Intelligence Policy: A Primer and Roadmap." UCDL Rev. 51 (399): 28.
- Castro, Daniel, Michael McLaughlin, and Eline Chivot. 2019. "Who Is Winning the AI Race: China, the EU or the United States?" *Center for Data Innovation*, August, 106.
- Cath, Corinne, Sandra Wachter, Brent Mittelstadt, Mariarosaria Taddeo, and Luciano Floridi. 2017. "Artificial Intelligence and the 'Good Society': The US, EU, and UK Approach." *Science and Engineering Ethics*, March. https://doi.org/10.1007/s11948-017-9901-7.
- "China Adds Huawei, Hikvision to Team Spearheading Country's AI Efforts." 2019. Tech in Asia. September 4, 2019. https://www.techinasia.com/china-adds-huawei-hikvision-ai-nationalteam.
- China Institute for Science and Technology Policy at Tsinghua University. 2018. "China AI Development Report 2018."
- Chui, Michael. 2017. "Artificial Intelligence the next Digital Frontier?" *McKinsey and Company Global Institute* 47: 3–6.
- Crawford, Kate, Roel Dobbe, Theodora Dryer, Genevieve Fried, Ben Green, Elizabeth Kaziunas, Amba Kak, et al. 2019. "Al Now Report 2019."
- Daxue Consulting. 2020. "The AI Ecosystem in China 2020." https://daxueconsulting.com/wpcontent/uploads/2020/03/AI-in-China-2020-White-Paper-by-daxue-consulting-2.pdf.
- Desjardins, Jeff. 2019. "Visualizing the AI Revolution in One Infographic." Visual Capitalist. April 9, 2019. https://www.visualcapitalist.com/ai-revolution-infographic/.
- Desouza, Kevin C., and Benoy Jacob. 2017. "Big Data in the Public Sector: Lessons for Practitioners and Scholars." Administration & Society 49 (7): 1043–64. https://doi.org/10.1177/0095399714555751.
- Ding, Jeffrey. 2018. "Deciphering China's AI Dream." Future of Humanity Institute. https://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf.

Einav, Liran, and Jonathan Levin. 2014. "The Data Revolution and Economic Analysis." *Innovation Policy and the Economy* 14 (1): 1–24.

EIT Health, and McKinsey & Company. 2020. "Transforming Healthcare with AI." https://eithealth.eu/wp-content/uploads/2020/03/EIT-Health-and-McKinsey_Transforming-Healthcare-with-AI.pdf.

Elahi, Haroon, Guojun Wang, and Dongqing Xie. 2017. "Assessing Privacy Behaviors of Smartphone Users in the Context of Data Over-Collection Problem: An Exploratory Study." In , 1–8. IEEE.

European Commission. 2016. "The European Single Market." Text. Growth - European Commission. July 5, 2016. https://ec.europa.eu/growth/single-market_en.

———. 2019. "Building Trust in Human-Centric Artificial Intelligence." Brussels: European Commission.

————. 2020. "White Paper On Artificial Intelligence - A European Approach to Excellence and Trust." White Paper. Brussels: European Commission.

- Floridi, Luciano, Josh Cowls, Thomas C. King, and Mariarosaria Taddeo. 2020. "How to Design AI for Social Good: Seven Essential Factors." *Science and Engineering Ethics*, April. https://doi.org/10.1007/s11948-020-00213-5.
- Foroohar, Rana. 2019. Don't Be Evil: The Case Against Big Tech. Penguin UK.
- Foster, Sophie. 2008. "Self-Regulation with More Accountability and Transparency." *Fijian Studies: A Journal of Contemporary Fiji* 6 (1/2): 236.
- Future of Life Institute. 2020. "National and International AI Strategies." Future of Life Institute. July 2020. https://futureoflife.org/national-international-ai-strategies/.

Girasa, Rosario. 2020. Artificial Intelligence as a Disruptive Technology: Economic Transformation and Government Regulation. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-35975-1.

Grassi, Valentino. 2018. "Individualism in China: The Social Credit Issue." Rome, Italy: LUISS Guido Carli.

Guarascio, Francesco. 2018. "EU Plans New Tax for Tech Giants up to 5 Percent of Gross Revenues."

High-Level Expert Group on Artificial Intelligence. 2019. "A Definition of AI: Main Capabilities and Disciplines." European Commission.

- Horowitz, Michael C. 2018. "Artificial Intelligence, International Competition, and the Balance of Power." *Texas National Security Review* 1 (3): 22.
- Inkster, Nigel. 2019. "The Huawei Affair and China's Technology Ambitions." Survival 61 (1): 105–11.
- Katyal, Sonia K. 2019. "Private Accountability in the Age of Artificial Intelligence." UCLA L. Rev. 66: 54.

Kazim, Emre, and Adriano Soares Koshiyama. 2020. "Lack of Vision: A Comment on the EU's White Paper on Artificial Intelligence." SSRN Electronic Journal.

https://doi.org/10.2139/ssrn.3558279.

Kroeskop, J. R. 2018. "Algorithms and Public Policy." Utrecht University.

- Landwehr, Carl, Dan Boneh, John C. Mitchell, Steven M. Bellovin, Susan Landau, and Michael E. Lesk. 2012. "Privacy and Cybersecurity: The Next 100 Years." *Proceedings of the IEEE* 100 (Special Centennial Issue): 1659–73. https://doi.org/10.1109/JPROC.2012.2189794.
- Larsson, Stefan. 2020. "On the Governance of Artificial Intelligence through Ethics Guidelines." Asian Journal of Law and Society, 1–23.
- Lee, Kai-Fu. 2018. *AI Superpowers: China, Silicon Valley, and the New World Order*. Houghton Mifflin Harcourt.
- Lim, Tai Wei. 2019. Industrial Revolution 4.0, Tech Giants, and Digitized Societies. Springer.
- Liu, Han-Wei, Ching-Fu Lin, and Yu-Jie Chen. 2019. "Beyond State v Loomis: Artificial Intelligence, Government Algorithmization and Accountability." *International Journal of Law and Information Technology* 27 (2): 122–41.
- Mann. 1999. "Framing China." Media Studies Journal, 102–7.

Manuel, Anja, and Pavneet Singh. 2019. "Compete, Contest and Collaborate: How to Win the Technology Race with China." *Stanford Cyber Policy Center*, October, 16.

- Masanet, Eric, Arman Shehabi, and Jonathan Koomey. 2013. "Characteristics of Low-Carbon Data Centres." *Nature Climate Change* 3 (7): 627–30. https://doi.org/10.1038/nclimate1786.
- Meissner, Mirjam. 2017. "China's Social Credit System. A Big-Data Enabled Approach to Market Regulation with Broad Implications for Doing Business in China." *Mercator Institute for China Studies*, 13.
- Mittelstadt, Brent. 2019. "AI Ethics Too Principled to Fail?" SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3391293.
- Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. 2016. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* 3 (2): 205395171667967. https://doi.org/10.1177/2053951716679679.
- Mou, Xiaomin. 2019. "Artificial Intelligence: Investment Trends and Selected Industry Uses." International Finance Corporation, 8.
- Natarajan, Prashant, John C. Frenzel, and Detlev H. Smaltz. 2017. *Demystifying Big Data and Machine Learning for Healthcare*. 1st ed. Boca Raton : Taylor & Francis, 2017.: CRC Press. https://doi.org/10.1201/9781315389325.
- Nicholson Price II, W. 2018. "Artificial Intelligence in Health Care: Applications and Legal Issues." SSRN Electronic Journal, April.
- Oh, Changhoon, Taeyoung Lee, Yoojung Kim, SoHyun Park, Saebom Kwon, and Bongwon Suh. 2017. "Us vs. Them: Understanding Artificial Intelligence Technophobia over the Google DeepMind Challenge Match." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2523–34. Denver Colorado USA: ACM. https://doi.org/10.1145/3025453.3025539.
- Parikh, Ravi B., Stephanie Teeple, and Amol S. Navathe. 2019. "Addressing Bias in Artificial Intelligence in Health Care." *JAMA* 322 (24): 2377. https://doi.org/10.1001/jama.2019.18058.
- Petersilia, Joan. 2008. "Influencing Public Policy: An Embedded Criminologist Reflects on California Prison Reform." *Journal of Experimental Criminology* 4 (4): 335–56.
- Puller, Steven L. 2006. "The Strategic Use of Innovation to Influence Regulatory Standards." *Journal* of Environmental Economics and Management 52 (3): 690–706. https://doi.org/10.1016/j.jeem.2006.07.002.
- PWC. 2017. "Sizing the Prize: What's the Real Value of AI for Your Business and How Can You Capitalise?" PWC. https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysissizing-the-prize-report.pdf.
- Roberts, Huw, Josh Cowls, Jessica Morley, Mariarosaria Taddeo, Vincent Wang, and Luciano Floridi. 2019. "The Chinese Approach to Artificial Intelligence: An Analysis of Policy and Regulation." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3469784.
- ———. 2020. "The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation." *AI & SOCIETY*, June. https://doi.org/10.1007/s00146-020-00992-2.
- Robinson, Sherman, and Karen Thierfelder. 2019. "Global Adjustment to US Disengagement from the World Trading System." *Journal of Policy Modeling* 41 (3): 522–36.
- Roff, Heather M. 2019. "The Frame Problem: The AI 'Arms Race' Isn't One." Bulletin of the Atomic Scientists 75 (3): 95–98. https://doi.org/10.1080/00963402.2019.1604836.
- Shane, Scott, and Daisuke Wakabayashi. 2018. "'The Business of War': Google Employees Protest Work for the Pentagon." New York Times, April 4, 2018. https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagonproject.html.
- Sheehan, Matt. 2018. "How China's Massive AI Plan Actually Works." *MacroPolo*, February 12, 2018. https://macropolo.org/analysis/how-chinas-massive-ai-plan-actually-works/.

- Soni, Vishal Dineshkumar. 2020. "Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA." Available at SSRN 3624487.
- Stiglitz, Joseph E., and Anton Korinek. 2017. "Artificial Intelligence and Its Implication for Income Distribution and Unemployment." NBE Working Paper Series. Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/papers/w24174.pdf.
- The Economist Intelligence Unit. 2017. "Risks and Rewards: Scenarios around the Economic Impact of Machine Learning." The Economist.
- Thewissen, Stefan, and David Rueda. 2019. "Automation and the Welfare State: Technological Change as a Determinant of Redistribution Preferences." *Comparative Political Studies* 52 (2): 171–208. https://doi.org/10.1177/0010414017740600.
- Tolan, Songül, Marius Miron, Emilia Gómez, and Carlos Castillo. 2019. "Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia." In Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law -ICAIL '19, 83–92. Montreal, QC, Canada: ACM Press. https://doi.org/10.1145/3322640.3326705.
- Tomasik, Brian. 2013. "International Cooperation vs. AI Arms Race." *Foundational Research Institute* 5.
- Veale, Michael. 2020. "A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence." *European Journal of Risk Regulation*, 8.
- Verba, Sidney. 1967. "Some Dilemmas in Comparative Research." World Politics 20 (01): 111–27. https://doi.org/10.2307/2009730.
- Voigt, Paul, and Axel Von dem Bussche. 2017. "The Eu General Data Protection Regulation (Gdpr)." A Practical Guide, 1st Ed., Cham: Springer International Publishing.
- Voort, H.G. van der, A.J. Klievink, M. Arnaboldi, and A.J. Meijer. 2019. "Rationality and Politics of Algorithms. Will the Promise of Big Data Survive the Dynamics of Public Decision Making?" *Government Information Quarterly* 36 (1): 27–38. https://doi.org/10.1016/j.giq.2018.10.011.
- Vydra, Simon, and Bram Klievink. 2019. "Techno-Optimism and Policy-Pessimism in the Public Sector Big Data Debate." *Government Information Quarterly*, June, S0740624X18302326. https://doi.org/10.1016/j.giq.2019.05.010.
- Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Mysers West, Rashida Richardson, Jason Schultz, and Oscar Schwartz. 2018.
 "Al Now Report 2018." Al Now Institute at New York University.
- Witt, Mariah. 2019. "Don't Google It: The European Union's Antitrust Parade (Enforcement) against America's Tech Giants." San Diego Int'l LJ 21: 365.
- World Intellectual Property Organization. 2019. "WIPO Technology Trends 2019: Artificial Intelligence."