INSTITUTE OF
DIGITAL
SOVEREIGNTY

iDF RIGHTS
INSTITUTE

Artificial Intelligence

Challenges and Prospects for Human Rights in Europe

Coordinated by Bernard Benhamou

General Secretary of the ISN

CONTENTS

ARTIFICIAL INTELLIGENCE

Challenges and Prospects for Human Rights in Europe

I. INTRODUCTION

Artificial intelligence (AI) technologies have today countless applications, ranging from energy consumption management to early diagnosis of illnesses, industrial processes optimization, autonomous vehicles control and design of 'smart' weapons. The development of AI technologies has greatly increased data processing capabilities, thereby facilitating decision-making and the automation of complex tasks previously considered as lastingly reserved for humans. For the industry players, AI represents both a major technological challenge and an opportunity to establish a new landscape of industrial, financial and political power. Already eight years ago, *Google* CEO Sundar Pichai described the future of technologies in these terms: *"Looking to the future, the next big step will be for the very concept of the "device" to fade away. Over time, the computer itself—whatever its form factor—will be an intelligent assistant helping you through your day. We will move from mobile first to an AI first world."*[1]

More recently, in 2023, Mustafa Suleyman, co-founder of *DeepMind* (now a subsidiary of *Google*), described the central role that AI would play in the coming age in these terms: *"AI is far deeper and more powerful than just another technology.*

[1] Sundar Pichai, CEO of Google, Founders' Letter, April 28, 2016
https://googleblog.blogspot.fr/2016/04/this-years-founders-letter.html

The risk isn't in overhyping it; it's rather in missing the magnitude of the coming wave. It's not just a tool or platform but a transformative meta-technology, the technology behind technology and everything else, itself a maker of tools and platforms, not just a system but a generator of systems of any and all kinds."[2]

With the rise of generative AI,[3] these technologies could also transform our societies' relationship to cultural creation and knowledge transmission, thereby profoundly modifying the very functioning of our societies. Beyond their industrial and social effects, AI technologies could have lasting consequences on the organization of our societies and on our freedoms. Whether in terms of the organization of democratic processes, the conditions for conducting public debate or the very understanding of the notion of truth, these technologies could alter all political activities and functions. For Kate Crawford, a leading scholar of the societal implications of AI, the current and future scope of application of AI gives a major political role to tech players. To support this claim, Kate Crawford even paraphrases the famous words of Clausewitz on war, but this time describing artificial intelligence as the continuation of politics by other means:

"*Artificial intelligence is politics by other means...*

Kate Crawford

> *"Simply put, artificial intelligence is now a player in the shaping of knowledge, communication, and power. These reconfigurations are occurring at the level of epistemology, principles of justice, social organization, political expression, culture, understandings of human bodies, subjectivities, and identities: what we are and what we can be. But we can go further. Artificial intelligence, in the*

[2] Suleyman, Mustafa; Bhaskar, Michael. The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma (p. 78). Random House 2023

[3] In particular Large Language Models (LLM) such as ChatGPT by OpenAI or Google's Gemini, but also AI designed to generate images or videos from text, such as Midjourney, Stable Diffusion, DALL·E 3, Sora, etc.

> *process of remapping and intervening in the world, is politics by other means—although rarely acknowledged as such. These politics are driven by the Great Houses of AI, which consist of the half-dozen or so companies that dominate large-scale planetary computation.”*[4]

For historian Yuval Harari, the transformations made possible by AI deserve to be analyzed and controlled in order to prevent democratic societies from falling into autocratic tendencies due to uncontrolled use of these technologies: *“The technology that favored democracy is changing, and as artificial intelligence develops, it might change further. Information technology is continuing to leap forward; biotechnology is beginning to provide a window into our inner lives—our emotions, thoughts, and choices. Together, infotech and biotech will create unprecedented upheavals in human society, eroding human agency and, possibly, subverting human desires. Under such conditions, liberal democracy and free-market economics might become obsolete.”*[5]

The challenges associated with the integration of human concerns into AI development are now major social and political issues. Even more so given that some tech players now use their AI to further a ‘political agenda’. The development of these technologies in Europe should be accompanied by a reflection on the very nature of our social and political interactions, and on the way in which we could allow, or not, some of these interactions to be altered by tech players. This alignment between ethical and political concerns and the realization of AI is what Brian Christian, author of a seminal book on AI ethics, calls the *"AI alignment problem"*: *“Machine learning is an ostensibly technical field crashing increasingly on human questions. Our human, social, and civic dilemmas are becoming technical. And our technical dilemmas are becoming human, social, and civic. Our successes and failures alike in getting these systems to do “what we*

[4] Crawford, Kate. Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence (pp. 19-20). Yale University Press 2021

[5] Why Technology Favors Tyranny (The Atlantic, Oct 2018) https://www.theatlantic.com/magazine/archive/2018/10/yuval-noah-harari-technology-tyranny/568330/

***want," it turns out, offer us an unflinching, revelatory mirror."*[6] **Alongside the measures allowing for the development of next-generation AI in Europe, one of the main political and industrial challenges our societies will face is ensuring that these technologies remain a reflection of the principles of freedom and European democratic values.**

II. AI: A LEVERAGE TOOL TO RESHAPE SOCIETIES

A. Personal Data, User Profiles and Microtargeting

Since the earliest stages of the development of major Internet platforms, a process of users' data extraction has been established. This was particularly the case for *Google* with what Shoshana Zuboff refers to as the "*behavioral surplus*" in her book *The Age of Surveillance Capitalism.*[7] This process was invented by *Google* based on an analysis of user queries that were not useful to improving the search engine. Initially, *Google* only archived queries relevant to the optimization of its search engine results, but later began to archive all queries to create profiles of all its users. With this 'surplus' of information, *Google* has been able to extensively analyze the behaviors and interests of billions of users, allowing it to offer advertisers precise targeting based on each user's profile which was built on several thousand parameters on each individual. This is known as *microtargeting*. These developments have enabled *Google* to become the world's largest advertising company. Some economists even refer to *Google* as *"an advertising company that owns a search engine…"*.

[6] Brian Christian. The Alignment Problem (W. W. Norton & Company 2020)

[7] Shoshana Zuboff. The Age of Surveillance Capitalism (Profile Books, Jan 2019)

However beyond their use in advertising, these microtargeting data can have both social and political uses (as in the case of the *Cambridge Analytica/Facebook* affair). As Zuboff and Harari note, once user profiles were created, *Google* was able to analyze changes in individual behaviors, but also their consequences on specific groups or an entire population. That led Shoshana Zuboff to remark: *"We thought we were searching Google, but Google was searching us…"*.

"We thought we were searching Google, but Google was searching us...

Shoshana Zuboff

According to AI Editor of the *Financial Times* Madhumita Murgia, beyond search engines, all user interactions now form part of these incredibly detailed profiles subsequently used by AI to model user behavior to offer value-added services.

> *I started to unpick the structure of this flourishing data economy. Every time I interacted with an online product – say Google Maps, Uber, Instagram, or contactless credit cards – with a single click, my behaviour was logged by these little cookies. Combined with public information such as my council tax or voter records, along with my online shopping habits and real-time location information, these benign datasets could reveal a lot about me, from my gender and age, down to nuances about my personality and my future decision-making.*[8]

Which means that it is in *Google's* interest, and more generally that of all players whose business model relies on advertising, to maintain this means of retrieving user data. The 'philosophy' of hyper-transparency has thus become the cornerstone of the economic model of big Internet advertising platforms, bringing with it the eventual demise of the concept of privacy. Then *Google* CEO Eric Schmidt put it: *"If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place."*[9] By ensuring that these user data flows

[8] Madhumita Murgia: Code Dependent: Living in the Shadow of AI (Henry Holt and Co. 2024)

[9] Google CEO Eric Schmidt Dismisses the Importance of Privacy (Electronic Frontier Foundation, Dec 10, 2009)
https://www.eff.org/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy

never dry up, these platforms can continue to monetize their services. Tech companies have thus shifted from an economic model of 'data-productivism' to one of 'data-extractivism' in which the goal is to retrieve as much data as possible on their users. Using AI to analyze user data has consequently become a major activity for Internet advertisers like *Google* and *Facebook*. In a telling article titled *"What 7 Creepy Patents Reveal about Facebook",*[10] the *New York Times* described the methods used by Facebook back in 2018 to closely analyze the psychological profile and behavior of its users. These patents concerned the analysis of their consumption patterns, their geolocation analysis, changes in relationships with other users, views of text or audiovisual content, and identification of the authors of photos and videos posted online from tiny defects in their smartphone sensors.

These in-depth profiles collected by *data brokers* allow not only advertisers, but also States and political and religious groups, to produce targeted content designed to constantly influence users, shifting not just their consumption patterns but their ideological or religious beliefs. As emphasized by U.S. Representative Kathy Castor during Mark Zuckerberg's *congressional hearing,*[11] *Facebook* compiles highly detailed profiles not only on *Facebook* users but also on people who do not have an account with the platform and who visit other websites. These are referred to as '*shadow profiles*' and enable companies to consolidate and expand their profile databases to nearly all Internet users. So in addition to the data indexed by the search engine, data on users are key to training AI models which could help replace traditional search engines with "answer engines".[12]

Due to the ease with which they enable access to data that previously required an individual but also collective feat of memory, smartphones have often been described as our "exo-brain". With the rise of intelligent assistants, functions

[10] What 7 Creepy Patents Reveal about Facebook (New York Times, Jun 21, 2018) https://www.nytimes.com/interactive/2018/06/21/opinion/sunday/facebook-patents-privacy.html

[11] U.S. Rep. Castor Questions Mark Zuckerberg, CEO of Facebook (C-Span, Apr 11, 2018) https://www.youtube.com/watch?v=OIvq763F57k

[12] AI search could break the web (MIT Tech Review, 31 Oct 2024) https://www.technologyreview.com/2024/10/31/1106504/ai-search-could-break-the-web/

related to data analysis and processing could also be transferred to AIs. By proposing their replacements to traditional information search engines, companies like *OpenAI* and *Perplexity AI* aim to demonstrate that the functions of search engines like *Google* could be replaced by generative AI models.[13] These next generation services provide answers to user questions rather than simply a list of results in the form of links. As a result, they will act as new 'filters' between users and the information available on the Internet. This once again opens up the risk of abuse and imperceptible censorship or, conversely, algorithmic amplification of some content. In the years to come, questions related to knowledge transmission in the era of generative AI could thus become central from an economic, cultural, and political perspective.

B. A Shift in Knowledge Transmission

> *Education is the point at which we decide whether we love the world enough to assume responsibility for it and by the same token save it from that ruin which, except for renewal, except for the coming of the new and young, would be inevitable. And education, too, is where we decide whether we love our children enough not to expel them from our world and leave them to their own devices, nor to strike from their hands their chance of undertaking something new, something unforeseen by us, but to prepare them in advance for the task of renewing a common world.*
>
> *Hannah Arendt*[14]

The transmission of knowledge and information within human societies has significantly evolved since the emergence of computers and the rapid decline in the costs traditionally associated with data archiving. In the past, these processes required considerable efforts to archive and transmit information to future

[13] OpenAI Is Launching Search Engine, Taking Direct Aim at Google (Wall Street Journal, Jul 25, 2024) https://www.wsj.com/tech/ai/openai-search-engine-searchgpt-97771f86

[14] Hannah Arendt, Between Past and Future (pp. 252-253) (Viking Press 1961)

generations. Only information deemed worthy of interest was passed on. As Viktor Mayer-Schönberger reminds us: *"Since the beginning of time, for us humans, forgetting has been the norm and remembering the exception. Because of digital technology and global networks, however, this balance has shifted. Today, with the help of widespread technology, forgetting has become the exception, and remembering the default."*[15] Today, beyond the archiving of information, the collapse of data storage costs and mass data processing has given AI an unparalleled role in the collection, reshaping and transmission of human knowledge. Beyond memory, the development of AI technologies could also permanently alter the transmission mechanisms and the way we view history and culture in our societies. Not only that, but our democratic values and principles could also be called into question by the relinquishment of sovereignty over our minds in favor of AI which would 'relieve' us of an ever-growing share of our social, cultural and political lives. A mix of resignation and fascination for AI technologies could indeed prompt our societies to rely entirely on these new oracles and entrust them with potentially dangerous roles in organizing democratic debate and even conducting electoral processes. The impact of these technologies on public debate and more broadly on our democracies remains difficult to accurately ascertain. However, in the absence of tipping a significant proportion of the electorate, they could have crucial consequences on elections by helping to manipulate the undecided voter segments that, in democratic countries, often determine the

“Victoria Woodcock, [microtargeting] Operations Director was the most indispensable person in the campaign. If she'd gone under a bus, Remain would have won.

Dominic Cummings
Director of the pro-Brexit campaign

[15] Viktor Mayer-Schönberger: Delete: The Virtue of Forgetting in the Digital Age (Princeton University Press 2009)

outcome of elections. It was particularly by using these microtargeting technologies on voters via mass data analysis systems that the *Brexit* campaign organizers were able to win over the undecided voters.

In his reference book *"The Chaos Engineers",* Giuliano da Empoli discusses the methods used by Dominic Cummings, Director of the pro-*Brexit* campaign, to win over voters by using microtargeting to polarize the different segments of the British population:

> *For the new Dr. Strangeloves of politics, the game is no longer about uniting people around the smallest common denominator, but rather about inflaming the passions of as many groups as possible, and then adding them together, often without their knowledge. To win a majority, they don't converge toward the center, but instead, they unite the extremes. Thanks to the work of a team of scientists, Cummings was able to target millions of undecided voters whom his opponents didn't even know existed, sending them exactly the right messages, at the right time, to sway them into the Brexit camp. Measuring the precise impact of this complex activity on the vote is impossible, but all indications suggest it was significant. Cummings himself wrote that Woodcock [the head of the software used in the campaign] was the "most indispensable person in the campaign. If she'd gone under a bus, Remain would have won...".*[16]

C. European AI Act: Industrial and Political Objectives

The introduction by the European Union of an *Artificial Intelligence Act* or *AI Act* (adopted by the European Council in May 2024) marks a turning point in the conception of critical technologies in our societies. The *AI Act* introduces a framework for regulating high-risk practices that affect citizens of the European Union. The level of risk ranges from minimal for video games to unacceptable for social scoring or AI used for political manipulation. The *AI Act* also introduces greater transparency to improve user information on risky practices, while encouraging AI developers to consider these constraints in the design of their

[16] da Empoli, Giuliano. Les ingénieurs du chaos (pp. 162-163). JC Lattès 2019

services. This will particularly apply to services that impact public safety or fundamental freedoms. Lastly, the *AI Act* will lead to the creation of economic activities and jobs in the field of AI ethics, both within businesses and administrations.

Furthermore, this European regulation aims to prevent an evolution towards authoritarian control tools similar to China's *Social Credit*, a political and social system that uses data collected on the behavior of Chinese citizens, including eventually their genetic data. The EU *AI Act* therefore aligns with *UNESCO*'s recommendations on AI ethics which emphasize its commitment to democratic principles and freedoms: "*Privacy is a right. The Recommendation calls to put in place appropriate safeguards to protect this right, including addressing concerns such as surveillance. It explicitly states that AI systems should not be used for social scoring or mass surveillance purposes.*"[17]

With the evolution of AI another freedom long considered an inalienable right could also be challenged: the freedom to keep thoughts private and secret. Experiments have already been conducted to capture people's emotions and physiological reactions directly from the brain. In China, for example, authorities require some workers to wear headgear with electroencephalogram (EEG) sensors to monitor their emotional state.[18]

In addition, other more discreet yet more widely used and more accurate mechanisms have already made it possible to identify a person's beliefs through mass analysis of behavioral data. These techniques were initially designed to influence users for advertising purposes but have since been used for ideological, political or religious manipulation. As historian Yuval Harari points out, in what

[17] UNESCO Recommendation on the Ethics of Artificial Intelligence (Nov 23, 2021) https://www.ohchr.org/sites/default/files/2022-03/UNESCO.pdf

[18] With brain-scanning hats, China signals it has no interest in workers' privacy (MIT Tech Review, Apr 30, 2018) https://www.technologyreview.com/2018/04/30/143155/with-brain-scanning-hats-china-signals-it-has-no-interest-in-workers-privacy/

is now referred to as 'cognitive warfare', certain types of personal data, such as medical data, could play a decisive role.

> *"It is crucial to remember that anger, joy, boredom and love are biological phenomena just like fever and a cough. The same technology that identifies coughs could also identify laughs. If corporations and governments start harvesting our biometric data en masse, they can get to know us far better than we know ourselves, and they can then not just predict our feelings but also manipulate our feelings and sell us anything they want — be it a product or a politician. Biometric monitoring would make Cambridge Analytica's data hacking tactics look like something from the Stone Age. Imagine North Korea in 2030, when every citizen has to wear a biometric bracelet 24 hours a day. If you listen to a speech by the Great Leader and the bracelet picks up the tell-tale signs of anger, you are done for."*[19]

Based on the information gathered about people, it therefore becomes possible to act on a multitude of apparently insignificant signals that can have a significant impact on individuals' emotions and decision-making. These invisible signals could even prove more effective than direct messages inviting people to adopt a specific behavior.

The late integration of generative AI in the *AI Act* highlights the difficulty in crafting regulations on such rapidly evolving technologies. These generative AI technologies were only in their early stages of commercial development at the time the *AI Act* was drafted. It will be the procedures and methods for applying these texts that will determine their effectiveness in the years to come. However, a form of consensus is emerging on the need to control the risks stemming from these technologies through regulation. Even the tech industry players themselves acknowledge that a lack of regulation on artificial intelligence will lead to social and political abuse, which could ultimately harm the entire technology industry. Even Mustafa Suleyman, the co-founder of *DeepMind*, recognizes the importance and pioneering nature of the European regulation *(AI Act)* in helping control (or contain) potential risks of these technologies.

[19] Yuval Noah Harari: the world after coronavirus (Financial Times, Mar 20, 2020) www.ft.com/content/19d90308-6858-11ea-a3c9-1fe6fedcca75

> *"The reality is that containment is not something that a government, or even a group of governments, can do alone. It requires innovation and boldness in partnering between the public and the private sectors and a completely new set of incentives for all parties. Regulations like the EU AI Act do at least hint at a world where containment is on the map, one where leading governments take the risks of proliferation seriously, demonstrating new levels of commitment and willingness to make serious sacrifices. Regulation is not enough, but at least it's a start. Bold steps. A real understanding of the stakes involved in the coming wave. In a world where containment seems like it's not possible, all of this gestures toward a future where it might be."*[20]

The regulations considered to limit the risks of abuse of AI (particularly the European *AI Act)* mainly address the applications of technologies rather than their technological foundations. However, in some cases, this distinction becomes difficult to maintain. For example, some applications that produce hyper-realistic *deepfakes present such challenges*. These technologies can sometimes have serious consequences for people's reputation and even mental or physical health. Examples include AI applications designed to create pornographic images or videos from content found on social media. The harmful nature of these applications is such that even the highly influential *MIT Technology Review* recommended their prohibition.[21] However, once again, total prohibition measures may face both technological and legal challenges.[22]

[20] Suleyman, Mustafa; Bhaskar, Michael. The Coming Wave: Technology, Power, and the Twenty-first Century's Greatest Dilemma (p. 232). Random House 2023

[21] A horrifying new AI app swaps women into porn videos with a click (MIT Technology Review Sep 13, 2021)
https://www.technologyreview.com/2021/09/13/1035449/ai-deepfake-app-face-swaps-women-into-porn/

[22] Bans on deepfakes take us only so far—here's what we really need (MIT Technology Review Feb 27, 2024)
https://www.technologyreview.com/2024/02/27/1089010/bans-on-deepfakes-take-us-only-so-far-heres-what-we-really-need/

III. FUNDAMENTAL FREEDOMS IN DANGER?

In her work on the protection of freedom of thought, human rights expert Susie Alegre stresses the need to analyze the text of the *European Convention on Human Rights* in light of the development of new generations of AI.

> *The rights to freedom of thought, conscience, religion and belief and freedom of opinion are absolute rights protected in international law. Without freedom of thought or opinion, we have no humanity, and we have no democracy. Making these rights real requires three things:*
>
> *1. the ability to keep your thoughts private;*
>
> *2. freedom from manipulation of your thoughts;*
>
> *3. that no one can be penalised for their thoughts alone.*[23]

The advances in digital technologies and AI can introduce further limitations to the exercise of these fundamental rights and freedoms. It is now possible to 'infer' the thoughts and beliefs of someone based on seemingly trivial actions (for example, on social media). Researchers at the University of *Cambridge* in 2013, demonstrated that analyzing *Facebook* '*likes'* makes it possible to accurately

[23] See Vermeulen, B., 'Article 9', in P. van Dijk et al. (eds), Theory and Practice of the European Convention on Human Rights, 4th edn, Cambridge, Intersentia Press, 2006, p.752.
Susie Alegre. Freedom to Think: Protecting a Fundamental Human Right in the Digital Age (Atlantic Books 2023)

predict a range of highly sensitive personal attributes.[24] The algorithms used in this study were 88% accurate in determining the sexual orientation of individuals, 95% accurate in distinguishing between African Americans and Caucasian Americans, and 85% accurate in differentiating political affiliation between Republicans and Democrats. Religious beliefs (Christian/Muslim) were correctly identified in 82% of cases and the relationship status of users and substance use were accurately predicted in between 65% and 73% of cases. A study led in 2014 by Michal Kosinski of *Stanford* University even demonstrated that the personality judgement established by AI models based on social media data were more accurate than those established by humans (friends, family, spouses or colleagues).[25]

The freedom not to be manipulated is also increasingly being challenged with the 'democratization' of certain AI models. These models analyze personality characteristics from the abundant data on social media and then use targeted messages to influence their reactions and decisions. The first famous case was the fraudulent use of *Facebook* data for political manipulation by the company *Cambridge Analytica.*[26] The scandal, in which AI played a central role,[27] may well have been instrumental in the 2016 U.S. Presidential election, decided by a margin of 107,000 votes in three states (Pennsylvania, Wisconsin and Michigan), which represented only 0.09% of the total votes cast.[28]

[24] Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. The Proceedings of the National Academy of Sciences (PNAS) Apr 9, 2013 https://www.pnas.org/doi/full/10.1073/pnas.1218772110

[25] Youyou W, Kosinski M, Stillwell D. Computer-based personality judgments are more accurate than those made by humans. The Proceedings of the National Academy of Sciences (PNAS) Jan 27, 2015 https://www.pnas.org/doi/10.1073/pnas.1418680112

[26] Cambridge Analytica: how did it turn clicks into votes? (The Guardian, May 6, 2018) https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie

[27] Decoded: How Cambridge Analytica used AI (Politico Jan 28, 2020) https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-how-cambridge-analytica-used-ai-no-google-didnt-call-for-a-ban-on-face-recognition-restricting-ai-exports/

[28] How Trump won the presidency with razor-thin margins in swing states (Wahington Post Nov 11, 2016) https://www.washingtonpost.com/graphics/politics/2016-election/swing-state-margins/

A. AI & Disinformation: the New Threats to Democracy

> *Freedom of information is the one that makes it possible to verify the existence of all others.*
>
> Christophe Deloire
> Director General of *Reporters Without Borders*[29]

The evolution of AI technologies associated with social media already has considerable effects on the essence of our democracies. This ranges from undermining the information necessary for the exercise of citizens' free will through the creation of '*deepfakes'* to the coordination of artificial opinion movements on social media and the sending of targeted messages based on specific profiles established by data brokers. Each component of our democracies can now be attacked by political actors, States or hostile groups.

The novelty of AI lies in its ability to intrude into something that previously seemed inaccessible via traditional means of propaganda: the very functioning of the human consciousness. For Simon McCarthy-Jones, associate professor of psychiatry at *Trinity College* Dublin, altering this part of the human mind that contributes to forming our beliefs equates to depriving us of what makes us capable of participating in democratic processes.

> *To lose sovereignty over our minds is to lose our dignity, our democracy, and even our very selves. Such sovereignty is termed mental autonomy. This is "the specific ability to control one's own mental functions," which include attention, memory, planning, rational thought and decision making (Metzinger, 2013).*

[29] Christophe Deloire "Freedom of information is the one that makes it possible to verify the existence of all others." (INA 2012)
https://mediaclip.ina.fr/en/i22075282-christophe-deloire-freedom-of-information-is-the-one-that-makes-it-possible-to-verify-the-existence-of-all-others.html

> *Dignity, "the presumption that one is a person whose actions, thoughts and concerns are worthy of intrinsic respect, because they have been chosen, organized and guided" (Nuffield Council on Bioethics, 2002, p. 121) requires mental autonomy. Democracies, in which citizens choose the laws that bind them (Johnson and Cureton, 2019), are only possible if citizens are mentally autonomous. The ability to think freely is so essential to our identity that to violate it is to deprive us "of personhood altogether" (Halliburton, 2009, p. 868)."*[30]

On social media, users are targeted by algorithms that attempt to offer content matching their tastes, interests, partisan preferences, etc. This results in confinement within information bubbles and a now well-documented risk of political manipulation, with messages primarily appealing to emotion and demagoguery. This was notably the case in the U.S. in what is known as the *Cambridge Analytica* scandal, or during the *Capitol* attack, during which extremist groups like *QAnon* managed to coordinate thanks to *Facebook*'s microtargeting algorithms. International law expert Anu Bradford commented on the consequences of this assault on American politics and on digital platforms regulation:

> *Some commentators suggest that the January 6, 2021, attack on the US Capitol may have been a turning point in the US's approach toward regulating online content. Emily Bazelon described the event as showing how the "American marketplace of ideas clearly failed," implying that American techno-libertarian beliefs about free speech were among the causes of the event. The Capitol riots led to broad condemnation of the role that online platforms played in allowing harmful and dangerous speech to gather so much momentum that it ultimately led to violence.*[31]

More recently, the riots following the murders of children in Southport (UK) prompted reflection on the role of social media in the *post-Brexit* polarization of

[30] McCarthy-Jones S. The Autonomous Mind: The Right to Freedom of Thought in the Twenty-First Century (Frontiers in Artificial Intelligence. Sep 26, 2019) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7861318/

[31] Anu Bradford. Digital Empires_The Global Battle to Regulate Technology (p. 363) (Oxford University Press 2023)

public opinion in Britain.[32] Beyond misinformation and targeted message campaigns, other AI practices could undermine the very functioning of our democratic institutions.

The combination of data from social media and the processing power of AI systems can subtly and undetectably alter voting conditions for citizens. For Elena Kagan, a U.S. *Supreme Court* Justice, AI assisted *gerrymandering* could undermine the very foundations of democracy. This manipulation could even render some electoral districts “unlosable” in an imperceptible way for citizens. In her opinion: *“Gerrymanders will only get worse (or depending on your perspective, better) as time goes on […]. What was possible with paper and pen — or even with Windows 95 — doesn't hold a candle (or an LED bulb?) to what will become possible with developments like machine learning. And someplace along this road, 'we the people' become sovereign no longer”*[33].

> ***“At very little cost, it is now possible to create autonomous AI disinformation systems to respond to users while adapting their responses to millions of different profiles...***

Recently, generative AI has made it possible to create and customize, at low cost, messages sent to large groups of Internet service users. These new capabilities have enabled the ‘democratization’ of automated interference and manipulation of public opinion. In addition to State entities, these disinformation technologies have now become accessible to private actors, political groups and even

[32] ‘A polarisation engine’: how social media has created a ‘perfect storm’ for UK’s far-right riots (Carole Cadwalladr - The Guardian Aug 3, 2024) https://www.theguardian.com/media/article/2024/aug/03/a-polarisation-engine-how-social-media-has-created-a-perfect-storm-for-uks-far-right-riots

[33] Supreme Court Justice Elena Kagan warns AI-powered gerrymandering could undermine US democracy (Business Insider, Jun 28 2019) www.businessinsider.com/justice-elena-kagan-warns-ai-powered-gerrymandering-may-hurt-democracy-2019-6

individuals. For a negligible cost, it is now possible to create a "fully autonomous AI disinformation system".[34] By using AI systems, these robots can provide a specific response to millions of users by adapting their responses to their profiles. People targeted by these campaigns think they are interacting with real people who share their views. This concern for the future of democracy is also one shared by Ben Buchanan, *White House* Special Advisor on AI:

> *Even more insidiously, AI offers the prospect of automating propaganda and disinformation campaigns, adding more fuel to the new fire. AI can write text that seems genuine, generate videos that seem real, and do it all faster and cheaper than any human could match or detect. AI also shapes the terrain on which disinformation efforts unfold. It influences which stories appear on Facebook news feeds, which tweets show up on Twitter timelines, and which video pops up next on YouTube's autoplay. Major internet platforms have become arenas where worldviews clash every day, with malicious actors trying to coax corporate algorithms to make their messages go viral, and companies trying—and often failing—to use AI to keep disinformation and other forms of hate at bay.*[35]

B. New Forms of Political Manipulation

These political manipulations could be further amplified in the new shared virtual worlds envisioned by promoters of the *Metaverse*. The experiences users will take part in will rely heavily on AI, which will allow for the large-scale creation of hyper-personalized content and services. This new generation of services combining virtual reality and generative AI content could help disseminate new forms of mass disinformation. An article in the *Washington Post* under the compelling title *"Facebook misinformation is bad enough. The metaverse will be worse",* by AI

[34] Inside Countercloud: a Fully Autonomous AI Disinformation System (The Debrief, Aug 16, 2023) https://thedebrief.org/countercloud-ai-disinformation/

[35] Buchanan, Ben; Imbrie, Andrew. The New Fire: War, Peace, and Democracy in the Age of AI (p. 8) (MIT Press 2022)

researcher Rand Waltzman, considers what ‘AI enhanced’ political campaigns could look like in these new universes:

> *A political candidate is giving a speech to millions of people. While each viewer thinks they are seeing the same version of the candidate, in virtual reality they are actually each seeing a slightly different version. For each and every viewer, the candidate's face has been subtly modified to resemble the viewer. This is done by blending features of each viewer's face into the candidate's face. The viewers are unaware of any manipulation of the image. Yet they are strongly influenced by it: Each member of the audience is more favorably disposed to the candidate than they would have been without any digital manipulation. This is not speculation. It has long been known that mimicry can be exploited as a powerful tool for influence. A series of experiments by Stanford researchers has shown that slightly changing the features of an unfamiliar political figure to resemble each voter made people rate politicians more favorably. The experiments took pictures of study participants and real candidates in a mock-up of an election campaign. The pictures of each candidate were modified to resemble each participant. The studies found that even if 40 percent of the participant's features were blended into the candidate's face, the participants were entirely unaware the image had been manipulated. In the metaverse, it's easy to imagine this type of mimicry at a massive scale.*[36]

These ‘AI enhanced’ candidates could be just the first step towards the use of entirely ‘synthetic’ characters created by generative AI. It would then be possible to convince citizens without them even knowing that these characters are entirely artificial. This science fiction scenario resembles the 2002 Andrew Niccol film *S1m0ne*, in which a cash-strapped director, played by Al Pacino, creates an entirely artificial movie star. Everyday interactions with artificial characters could thus become a reality. With the latest generation of generative AIs, these artificial companions are already becoming commercial realities. Some services now allow users to create virtual partners in the form of photorealistic avatars that use generative AIs for their appearance and to generate interactions with their users.[37]

[36] Rand Waltzman. Facebook misinformation is bad enough. The metaverse will be worse (Washington Post, Aug 22, 2022) https://www.washingtonpost.com/opinions/2022/08/22/metaverse-political-misinformation-virtual-reality/

[37] The Perfect Girlfriend (Esquire Oct 3, 2024) https://www.esquire.com/news-politics/a62452522/ai-girlfriend/

The boundary between human and machine, which some tech players aim to abolish, could become increasingly difficult to discern. In the words of Laurence Devillers, professor of computer science applied to social sciences at Paris Sorbonne University: *"We are entering an age of inextricable relationships between humans and machines, a relationship of trust and affection within which the separation between living and artifact, today so clear, will become increasingly blurred."*[38]

C. Towards Emotional Relationships with AIs?

A recent example of a 'seductive' generative AI occurred with the launch of the voice version of *OpenAI*'s *Chat GPT-4o* assistant. Named *"Sky",* the assistant's voice seemed directly inspired by the voice in the Spike Jonze film *"HER"*. Released in 2013, the film's character played by Joaquin Phoenix falls in love with the female voice of his computer's operating system, portrayed by Scarlett Johansson. The *GPT-4o "Sky"* voice sounded so similar to the actress voice that she sought legal action against *OpenAI* for violation of her right of publicity.[39] This led *OpenAI* to remove the *Sky* voice from its virtual assistant. In her statement, Scarlett Johansson criticized the cavalier attitude of Sam Altman, *OpenAI* CEO, who, after several refusals from the actress, still decided to use the *Sky* voice. Johansson also advocated for legal solutions to be put in place to protect artists: *"In a time when we are all grappling with deepfakes and the protection of our own likeness, our own work, our own identities, I believe these are questions that deserve absolute clarity. I look forward to resolution in the form of transparency and the*

[38] Devillers, Laurence. Les robots émotionnels: Santé, surveillance, sexualité… : et l'éthique dans tout ça ? (p.40). Humensis 2019 see also on emotions detection by IA:
Laurence Devillers, Laurence Vidrascu, Lori Lamel. Challenges in real-life emotion annotation and machine learning based detection, Neural Networks, Volume 18, Issue 4, 2005, Pages 407-422, https://doi.org/10.1016/j.neunet.2005.03.007.

[39] Scarlett Johansson said she was forced to hire legal counsel to deal with Sam Altman and OpenAI (Fortune May 21, 2024)
https://fortune.com/2024/05/20/scarlett-johansson-chatgpt-sky-voice-lawyer-sam-altman-openai-her/

passage of appropriate legislation to help ensure that individual rights are protected."[40]

In September 2024, solely because of this conflict with *OpenAI*, Scarlett Johansson was included in the list of the *100 Most Influential People in AI 2024*[41]. From a technological and industrial perspective, this *OpenAI* controversy also highlights how much technological companies want users to develop emotional relationships with AI. For AI technologies to develop as part of the daily lives of users, they need to be reassuring, and gradually become part of their environment. As the pioneer of ubiquitous computing, Mark Weiser, stated: *"The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it."*[42] For companies developing these AIs, it is important that users feel confident so that they are willing to share more information that will subsequently allow developers to refine their AIs. An entire field of computer science research is devoted to humanizing human/machine relationships, but also to the risk of dehumanization that this emotional connection with machines could have on real human relationships.[43] One of the consequences of this dehumanization is the erosion of trust between individuals, and, with it, the questioning of *'conversation'* between citizens and the confrontation of opinions. This conversation, which not only forms the fabric of our daily lives but is also the very backbone of democracy.

[40] Scarlett Johansson's Statement About Her Interactions With Sam Altman (New York Times, May 20, 2024) https://www.nytimes.com/2024/05/20/technology/scarlett-johansson-openai-statement.html

[41] The 100 Most Influential People in AI 2024 - Time (Sep 5, 2024) https://time.com/collection/time100-ai-2024/

[42] Mark Weiser. The Computer for the 21st Century (Scientific American 1991) https://ics.uci.edu/~corps/phaseii/Weiser-Computer21stCentury-SciAm.pdf

[43] See also on this subject the work of Sherry Turkle at MIT on the risks that this 'humanization' of human/machine relationships could pose to interpersonal relationships.
Alone Together: Why We Expect More from Technology and Less from Each Other (Basic Books 2011)

IV. CONSEQUENCES OF AI ON DEMOCRACIES

A. Protective AIs vs. Hostile AIs?

In democratic societies, the detection of interference and the fight against attacks on democracy will increasingly rely on artificial intelligence technologies. Experts on these new AI technologies may thus be the only ones capable of verifying if public opinions have been manipulated or if the expression of votes has remained sincere.[44] Over time, considerable pressure could be placed on these experts and their role in the democratic process would inevitably generate suspicions. Citizens' control over the uses of these technologies will thus become even more crucial than before. At the same time, the scope of sensitive data needs to be redefined, because the developments of AI technologies have made it possible to analyze seemingly innocuous information to infer sensitive data about individuals.

Finally, citizens must also be better informed about the risks of new types of surveillance and manipulation which, as seen with the war in Ukraine, are now an integral part of the belligerence referred to as 'hybrid warfare'. Henry Kissinger described it in the following terms: *"Today, after the Cold War, the major powers and other states have augmented their arsenal with cyber capabilities whose utility derives largely from their opacity and deniability and, in some cases, their operation, at the ambiguous border of disinformation, intelligence collection, sabotage, and traditional conflict — creating strategies without acknowledged doctrines."*[45]

[44] Richard Ghevontian. La notion de sincérité du scrutin - Cahiers du Conseil Constitutionnel n° 13 - January 2003 https://www.conseil-constitutionnel.fr/nouveaux-cahiers-du-conseil-constitutionnel/la-notion-de-sincerite-du-scrutin

[45] Henry A. Kissinger, Eric Schmidt and Daniel Huttenlocher. "The Age of AI: And Our Human Future." (John Murray Press, 2021)

Today, AI already plays a key role in many disinformation campaigns. Whether it is by analyzing the profiles of campaigns targets or by crafting messages, images or videos using generative AI. Thousands and even millions of fake accounts can also be coordinated to create a false impression of grassroots support or opposition (known as "*astroturfing*"), causing voters to be influenced by what they think to be spontaneous trends in public opinion. Some experiments therefore aim to use conversational AIs to challenge the beliefs instilled by disinformation campaigns among some voters.[46]

Major Internet platforms also use AI for content moderation and to detect 'inauthentic behavior'. The main feature of social networks is making registration as simple as possible. This allows automated systems to register large numbers of accounts to set up large-scale political disinformation campaigns. With AI it becomes possible to create millions of fake accounts in just a few days. Fighting these practices involves analyzing the dynamic of exchanges between certain accounts, the nature of messages exchanged and the type of coordination between these accounts to amplify certain messages. *Facebook*, for example, uses AI systems to detect and delete these fake accounts on a massive scale. In 2023 alone, Facebook removed over 3 billion fake accounts.[47]

Rather than attempting to limit disinformation campaigns 'at the source', some specialists recommend using AI to combat their manifestations.[48] If these processes designed to protect democratic debates using AI were to become widespread, they themselves could come to be subject to criticism. Due to the considerable power over electoral processes that would be granted to these AI technologies, suspicions would inevitably fall on the developers of

[46] This Chatbot Pulls People Away From Conspiracy Theories (New York Times, Sep 12, 2024) https://www.nytimes.com/2024/09/12/health/chatbot-debunk-conspiracy-theories.html

[47] Inauthentic Behavior Meta Transparency Center (2024) https://transparency.meta.com/en-us/policies/community-standards/inauthentic-behavior/

[48] Stopping AI disinformation: Protecting truth in the digital world (Word Economic Forum, Jun 14, 2024) https://www.weforum.org/agenda/2024/06/ai-combat-online-misinformation-disinformation/

these technologies, further fueling the distrust already prevalent in electoral processes and, more broadly, in democratic institutions. In a way, this opacity of democratic institutions protected by AI would constitute the ultimate manifestation of the *"black box society"* effect that Frank Pasquale condemned in his book by the same name[49]. He argued in favor of greater transparency of the algorithms used by big platforms by opening the code to experts. He refers to the notion of qualified transparency which would allow independent experts groups to audit the algorithms code most critical to the democratic functioning of our societies.

B. New Ways to Fight Disinformation

Among the measures that help limit the impact of disinformation campaigns, some experts are now considering focusing on the receptivity of citizens rather than combating the effects of these campaigns once they have already occurred. Thus, the recent report of the *États généraux de l'information*, an independent, collective and collaborative consultation process on the right to information, proposed the development of *"prebunking"*[50] which, contrary to *"debunking"*, aims to raise the awareness of possible targets of disinformation in advance rather than taking action after the fact, when these campaigns have already progressed in the minds of citizens:

> *The working group considers the implementation of pre-bunking techniques (anticipatory debunking) essential for immunizing citizens against*

[49] "Black Box Society - The Secret Algorithms That Control Money and Information" by Frank Pasquale (Harvard University Press 2015)

[50] See on this subject the work of David Colon: La guerre de l'information: Les États à la conquête de nos esprits (Taillandier 2023) and the OECD report "Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity" (OECD, Mar 4, 2024) https://www.oecd.org/en/publications/facts-not-fakes-tackling-disinformation-strengthening-information-integrity_d909ff7a-en/full-report.html

> *disinformation campaigns. Education on the phenomenon of "weaponization," i.e., the use of the informational environment for strategic and influence purposes, is key, as individuals may become potential targets for the dissemination of false or biased information. The group recommends "test and learn" phases to determine the most effective ways to immunize citizens against disinformation. Support from academic research is crucial to identify the most effective measures. In addition to its functions of detecting and characterizing threats, Viginum should be tasked with a resilience mission, to structure anticipatory debunking efforts in France across all sectors, drawing inspiration from the Swedish Psychological Defence Agency (mpf.se).*[51]

As individuals, but also as society, we must ask ourselves about the political and social power handed to these companies to change the very shape of our societies. AI could well further amplify this power in the coming years, given that AI models can progressively consider users' actions to refine their responses. It has become possible to use the vast amount of data collected on individuals to enhance not information on their behavior but rather how their behavior can be shaped over time. Whether consumption patterns or ideological, religious or political convictions, several thousand parameters can be integrated into the messages sent to individuals to influence them.

The analysis of individuals' behavior via personal data extraction used to be a 'craft' approach; but now with AI, the industrialization (or democratization in the most negative sense of the term) of emotional manipulation of entire populations is now accessible at minimal cost. British international human rights lawyer Susie Alegre described the risks associated with the accumulation of personal data in these words: *"However, most of the discussion so far about the solutions has focused on privacy and data protection. But the fundamental problem with*

[51] Extract from the report of the États généraux de l'information - Protecting and Developing the Right to Information: an Urgent Democratic Imperative (Sep 12, 2024)
https://www.vie-publique.fr/files/rapport/pdf/295405.pdf
https://etats-generaux-information.fr/content/download/158281/file/EGI_RAPPORT_DE_PILOTAGE_ANGLAIS-2.pdf

techniques like behavioural microtargeting and the 'surveillance capitalism' model is not the data; it's how it is used as a key to our minds."[52]

C. A "Race to the Bottom" in Democracy?

> ***"AI will also improve the tools in the autocrat's arsenal, such as lethal autonomous weapons, cyber operations, and disinformation. In autocracies, ethical concerns will be nothing but minor speed bumps...***
>
> **Ben Buchanan**

Today the goal of many authoritarian regimes is for democracy to be perceived as merely a 'narrative', and moreover one that is less effective than non-democratic regimes. Among the arguments they advance against democratic regimes are their instability and their inability to design long-term public policies. Social media, and later AI, have thus become tools to convince their populations, but also those of democratic countries, of the benefits of authoritarian regimes. Today, AI represents numerous advantages for these regimes in conducting hybrid warfare against democracies. AI enables the propagation at little cost of their propaganda and a way to amplify narratives on the supposed effectiveness of authoritarian forms of governance, while simultaneously discrediting democratic regimes and increasing the polarization of their public opinions. However, one of the weaknesses of authoritarian regimes, as China proved at the beginning of the *COVID-19* pandemic, is the fear among local officials' of transmitting information that could be seen as a challenge to central

[52] Susie Alegre. Freedom to Think: Protecting a Fundamental Human Right in the Digital Age (Atlantic Books 2023)

power. It was the fear of being punished that led these officials to conceal the reality of the pandemic in the crucial first weeks.

For authoritarian regimes, AI also appears as a potential solution to the challenge of obtaining reliable information to improve both political control and local administration. Indeed, AI allows for the processing of a multitude of signals from the various sensors installed throughout cities and across the country enabling the control of citizens and reducing dependence on 'unreliable' human sources.

This greater degree of control over AI by authoritarian regimes is raising concerns about the ability of democracies to leverage these technologies to protect freedoms. Ben Buchanan, for example, underlines the need for the leaders of democratic countries to remain vigilant in the face of these new threats.

> *When it comes to geopolitical competition, a worrying possibility runs throughout this book: that AI will do more for autocracy than it will for democracy. In this view, autocrats will simply be better able or more willing to use AI to control people, information, and weapons. Unencumbered by privacy laws, they will have access to more data. They will use their central planning systems to drive algorithmic research and build ever-faster computers, turning autocracy's biggest weakness—the need for centralization—into its greatest strength. AI will help entrench power within the state, aiding in surveillance and repression and compelling private companies to assist as needed. For national security, the argument goes, AI will also improve the tools in the autocrat's arsenal, such as lethal autonomous weapons, cyber operations, and disinformation. In autocracies, ethical concerns will be nothing but minor speed bumps.*[53]

There is an increasingly significant tension between the commercial objectives of platforms whose business models are based on acquiring new users (and thus their personal data), and the legitimate concerns of public authorities in democratic societies to limit the spread of hate speech and disinformation

[53] Ben Buchanan, Andrew Imbrie. The New Fire: War, Peace, and Democracy in the Age of AI (p. 8) (MIT Press 2022)

campaigns. In this respect, protecting democracies must also involve a democratic debate on these platforms' business models, whose political role is likely to grow further with the rise of AI.

D. The Vulnerabilities of Democracies Confronted with AI

Democratic societies' openness to external influences has often been described by hostile regimes as a political and ideological vulnerability. Now, the use of digital technologies creates new attack surfaces to target our infrastructure and the social and political functioning of our societies. With the rise of AI technologies, the dynamics of conflicts could be profoundly altered in favor of cyber attackers. This is summarized by the authors of *The Age of AI*:

> *As advanced economies integrate digital command-and-control systems into power plants and electricity grids, shift their governmental programs onto large servers and cloud systems, and transfer data into electronic ledgers, their vulnerability to cyberattack multiplies; they present a richer set of targets so that a successful attack could be substantially devastating. Conversely, in the event of a digital disruption, the low-tech state, the terrorist organization, and even individual attackers may assess that they have relatively much less to lose.*[54]

The limits of the 'emancipatory' role of digital technologies have been studied since the Internet first launched into the public domain. As early as 2001, researchers at the *Carnegie Endowment for International Peace* expressed doubts regarding the ability of democracies to control these technologies better than authoritarian regimes:

[54] Henry A. Kissinger, Eric Schmidt and Daniel Huttenlocher. "The Age of AI: And Our Human Future." (John Murray Press, 2021)

> *Authoritarian regimes will have to continually adapt their measures of control if they want to counter effectively the challenges of future variations in information and communication technologies. It is quite possible that this task will prove too difficult and that use of ICTs will eventually play a role in the democratic revolution that has been so widely predicted. Over time, however, authoritarian regimes have weathered innumerable challenges posed by changing technologies, and they may prove up to the current challenge as well.*[55]

The *X* network *(formerly Twitter*) is often described as the social media platform of anger and immediacy leading to the polarization of opinions in information bubbles. This tendency has been even more pronounced since Elon Musk began deploying his vision of unrestricted freedom of expression. This has made *X* even more accessible for destabilization operations and interference by hostile states and radical political and religious groups. With the development of generative AIs, we could enter an even more favorable phase for the manipulation of opinions. Indeed generative AIs allows for the shaping of individual opinions to meet the demands of the ultimate information bubble: the one that will be built between each user and AI with no human intermediaries.

E. China: AI for Political Control of Populations

Many authoritarian regimes have demonstrated that they can limit the risks of internal destabilization and at the same time benefit from these technologies as a form of ideological control over populations. China initially developed its *Great Firewall* to control and censor all information flows coming and leaving its territory to avoid ideological contagion from democratic countries. Then, the Chinese authorities developed an unparalleled mass surveillance system: the *Social*

[55] The Internet and State Control in Authoritarian Regimes: China, Cuba, and the Counterrevolution Carnegie Endowment for International Peace report by Shanthi Kalathil and Taylor C. Boas (CEIP 2001) https://carnegie-production-assets.s3.amazonaws.com/static/files/21KalathilBoas.pdf

Credit System. This Orwellian system, developed by some of China's most powerful companies such as *Alibaba,* assigns a score to all Chinese citizens. The *Social Credit System* relies on both facial recognition technologies and AI algorithms designed to analyze individuals' behavior. The score assigned to each Chinese citizen aims to assess their social, financial and ideological 'behavior'.[56] A low *Social Credit* score prohibit individual from traveling by train or plane, accessing certain public services or loans. The goal behind the *Social Credit System* is to eliminate potentially disruptive elements in society. The aim, according to Hou Yunchun, former deputy director of the development research center of the *State Council*, is to make "*discredited people become bankrupt...*".[57] The Chinese regime thus demonstrated with the *Social Credit* that it could create a surveillance system and eradicate dissidence across its entire population. Just as rivalry existed between democracies about the use of technologies for economic purposes, another form of competition has emerged between authoritarian regimes to improve surveillance technologies targeting their citizens *via* AI: from real-time geolocation tracking, facial recognition to the analysis of citizens' activities and behavior. At the same time, these technologies have enabled foreign powers to influence the opinions of democracies. Autocracies have thus built on the social psychology experiments carried out by

> ***“In China, the Social Credit System enhances knowledge on individuals and companies and improves AI algorithms which in turn make it possible to analyze data in greater detail in order to better control populations…***

[56] China's Social Credit System (Infographics, Bertelsmann Foundation 2019) www.bertelsmann-stiftung.de/fileadmin/files/aam/Asia-Book_A_03_China_Social_Credit_System.pdf

[57] China blacklists millions of people from booking flights as 'social credit' system introduced (The Independent, Nov 22, 2018) www.independent.co.uk/news/world/asia/china-social-credit-system-flight-booking-blacklisted-beijing-points-a8646316.html

social networks to manipulate users' emotions.[58] Politically, it became possible to measure the impact of political campaigns on populations and to develop new know-how in the field of disinformation. It was precisely by relying on the knowledge of Russian experts in disinformation that the teams at *Cambridge Analytica* in 2016 launched, their campaign to favor Donald Trump's election.[59]

In her book *Weapons of Math Destruction,* mathematician Cathy O'Neil, two years before the revelations of the *Cambridge Analytica* scandal, described the "microtargeting" mechanisms used to manipulate undecided voters: *"Modern consumer marketing, however, provides politicians with new pathways to specific voters so that they can tell them what they know they want to hear. Once they do, those voters are likely to accept the information at face value because it confirms their previous beliefs, a phenomenon psychologists call confirmation bias."*[60]

As *MIT Tech Review* points out, successive examples of negligence on the part of the major Internet platforms have forced new AI players to better take better account of the political dimension of their activities, particularly when it comes to disinformation and interference.

> *Researchers have long expected adversarial actors to adopt generative AI technology, particularly large language models, to cheaply increase the scale and caliber of their efforts. The transparent disclosure that this has begun to happen—and that OpenAI has prioritized detecting it and shutting down accounts to mitigate its impact—shows that at least one large AI company has learned*

[58] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock: Experimental evidence of massive-scale emotional contagion through social networks (Proceedings of the National Academy of Sciences PNAS Jun 2, 2014)
https://www.pnas.org/doi/10.1073/pnas.1320040111

[59] Cambridge Analytica: links to Moscow oil firm and St Petersburg university (The Guardian, Mar 17, 2018)
https://www.theguardian.com/news/2018/mar/17/cambridge-academic-trawling-facebook-had-links-to-russian-university

[60] Cathy O'Neil. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy by Cathy O'Neil (p. 187). (Crown 2016)

> *something from the struggles of social media platforms in the years following Russia's interference in the 2016 US election.*[61]

China long used its technological power to promote its regime and the benefits of 'State-controlled capitalism' by the *Chinese Communist Party*, then later used its technological instruments to influence political debates in Europe and the U.S.. In this area, China closely analyzed the methods used by Russia in its disinformation campaigns. As shown in the report *Chinese Influence Operations. A Machiavellian Moment* by Paul Charon & Jean-Baptiste Jeangène Vilmer: published by the *Institut de Recherche Stratégique de l'École Militaire (IRSEM)*.

For the PLA: Russia is a Model to Imitate[62]

For its experience manipulating social media, at least since the annexation of Crimea in 2014 (closely followed in China, by the PLA notably), Russia "provided China with a model to imitate." In 2014, a member of the PLA General Staff Department wrote an article outlining three lessons from the "war on [the Russian] public opinion" in Ukraine: "take the offensive by pushing your narrative first, present legal arguments, and support it all with hard power" Several similar articles appeared in the following years, demonstrating a clear willingness, on the part of the Chinese military, to learn from the Russian example. In 2018, an article investigating RT's coverage of the American bombings in Syria recommended an "investigation of RT's communication methods [:] without

[61] Propagandists are using AI too—and companies need to be open about it (MIT Technology Review Jun 8, 2024) https://www.technologyreview.com/2024/06/08/1093356/propagandists-are-using-ai-too-and-companies-need-to-be-open-about-it/

[62] https://drive.google.com/file/d/1AhHevTlIOddtKcRaOl6pkUbZ1oXCOima/view

losing its 'objectivity,' we can silently influence the emotions and inclinations of the public and make it more dependent on information from our media outlets." RT has regularly been quoted as a model to follow in publications of the Chinese military, especially for its activities on social media. Analysts from the PLA National Defense University compared the Russian channel to "a propaganda aircraft carrier," highlighting its performance on *YouTube*.

Now, with *TikTok,* its video-sharing platform with over 1.5 billion users, China has one of the largest influence channels ever created for Western populations. As democracies have learned, the Chinese regime uses its tools to collect sensitive information - whether ideological, religious or medical - on its platform's users to influence them. In their Senate report: *TikTok Tactic: Opacity, Addiction, and Chinese Shadows*, senators Mickaël Vallet and Claude Malhuret noted the following with regard to the risks this platform represents:

> *"With its undeniable international success, TikTok could be effectively used to serve the interests and longevity of the Chinese regime, to the detriment of the security of global TikTok users data, the plurality of expression and the quality of information disseminated on its platform. As Bernard Benhamou, General Secretary of the Institute of Digital Sovereignty (ISN) emphasizes: "By definition, not using these data would be a form of professional misconduct on the part of the Chinese services".*[63]

Now, beyond propaganda designed to promote the Chinese regime, its focus is on interference in the main electoral processes of democracies. Recently, Chinese services have borrowed the methodologies of the Russian services in terms of disinformation to develop their own interference actions in U.S. electoral

[63] Extract from the report presented to the French Senate "La tactique TikTok: opacité, addiction et ombres chinoises" - n° 831 (2022-2023) submitted on July 4, 2023 https://www.senat.fr/rap/r22-831-1/r22-831-11.pdf

process.[64] It has become possible for them to artificially amplify messages to thousands and even millions of profiles by specifically targeting people who have shown sensitivity to certain subjects. This algorithmic magnification can take place without the sender of the original message even being aware of it.

For countries like Russia and China, hybrid warfare is not only a means to limit the risks of democratic contagion but also to restrict the innovation capabilities of opposing countries. Their technological capabilities threaten their regimes both economically and militarily as well as politically due to the attractiveness they can represent for their scientific elites.

V. TRANSHUMANISM, LONGTERMISM, EUGENICS: RADICAL IDEOLOGIES FOR A CERTAIN VISION OF HUMANITY

Technologies can be inspired by the political, philosophical or religious beliefs of their creators and can in turn transmit these beliefs to their users. In 1994, philosopher Umberto Eco humorously described what he considered to be the religious roots of the clash between *Microsoft* and *Apple,* by studying the differences between the first *MS-DOS* and *Macintosh* operating systems:

> *"The fact is that the world is divided between users of the Macintosh computer and users of MS-DOS compatible computers. I am firmly of the opinion that the Macintosh is Catholic and that DOS is Protestant. Indeed, the Macintosh is counter-reformist and has been influenced by the ratio studiorum of the Jesuits. It is cheerful, friendly, conciliatory; it tells the faithful how they must proceed step by step to reach – if not the kingdom of Heaven – the moment in which their document is printed. It is catechistic: The essence of revelation is dealt with via simple formulae and sumptuous icons. Everyone has a right to salvation. MS-*

[64] China's Advancing Efforts to Influence the U.S. Election Raise Alarms (New York Times, Apr 2, 2024) https://www.nytimes.com/2024/04/01/business/media/china-online-disinformation-us-election.html

> *DOS is Protestant, or even Calvinistic. It allows free interpretation of scripture, demands difficult personal decisions, imposes a subtle hermeneutics upon the user, and takes for granted the idea that not all can achieve salvation. To make the system work you need to interpret the program yourself: Far away from the baroque community of revelers, the user is closed within the loneliness of his own inner torment."*[65]

In contrast, today these technologies are at the origin of new ideologies aimed at accounting their effects on societies and on humanity. In the extreme, AI could lead in the future to the creation of new cults based on the opinions of the AI. These opinions could be perceived as emanating from a transcendence to which only AI has access. For some philosophers, human beings tempted to surrender to AI express a sort of modern 'weariness of the self'. This is the case for transhumanists who cultivate a belief in an idealized version of humanity that would follow the path of machines to escape the human condition (fallibility, mortality, aging, etc.). Susan Schneider remarks: "*Indeed, some suspect that synthetic intelligence will be the next phase in the evolution of intelligence on Earth. You and I, how we live and experience the world right now, are just an intermediate step to AI, a rung on the evolutionary ladder.*"[66]

A.Transhumanism: Towards Inhuman Humans?

For tech industry leaders, the transhumanist credo conveniently positioned itself to redefine all the ethical regulations surrounding technologies in the name of a distant higher interest for the human species. This comes at the expense of debate on the immediate risks that these innovations may pose to societies. As an example, for advocates of the 'longtermism' school of thought, the climate crisis is merely an epiphenomenon since it alone cannot threaten the very existence of

[65] The Holy War: Mac vs. DOS English translation of Umberto Eco's back-page column, "La bustina di Minerva," in the Italian newsweekly Espresso (Sep 30, 1994). https://www.agonia.net/index.php/prose/124266/The_Holy_War:_Mac_vs._DOS

[66] Susan Schneider: Artificial You: AI and the Future of Your Mind (Princeton University Press 2019)

humanity. Only humanity's ‘multi-planetary’ destiny is considered worth preserving, even at the cost of limiting the freedoms of the citizens of our current societies. According to transhumanist philosopher Nick Bostrom, it could be necessary to establish a global surveillance system in the best interests of humanity. This surveillance could aim to prevent the actions of humans that might hinder humanity's destiny, which according to Bostrom involves allowing humans to become multi-planetary.

> *“The connection with longtermism is that, according to Bostrom and Ord, failing to become posthuman would seemingly prevent us from realising our vast and glorious potential, which would be existentially catastrophic. As Bostrom put it in 2012, ‘the permanent foreclosure of any possibility of this kind of transformative change of human biological nature may itself constitute an existential catastrophe.’ Similarly, Ord asserts that ‘forever preserving humanity as it is now may also squander our legacy, relinquishing the greater part of our potential.’ Bostrom himself argued that we should seriously consider establishing a global, invasive surveillance system that monitors every person on the planet in realtime, to amplify the ‘capacities for preventive policing’*[67] *(eg, to prevent omnicidal terrorist attacks that could devastate civilisation).”*[68]

Bostrom’s apocalyptic view of the future might seem anecdotal if it had not been met with approval from some of Silicon Valley's the most powerful libertarian CEOs, including Elon Musk, Peter Thiel, Jeff Bezos and Samuel Bankman-Fried, the fallen cryptocurrency magnate.

[67] Bostrom, Nick (2013). Existential Risk Prevention as Global Priority. Global Policy, 4 (1), 15–31. https://existential-risk.com/concept.pdf

[68] Why longtermism is the world’s most dangerous secular credo (Émile P Torres in Aeon Essays Oct 19, 2021)
https://aeon.co/essays/why-longtermism-is-the-worlds-most-dangerous-secular-credo

Extract from the Transhumanist Declaration[69]

- Humanity will be radically changed by technology in the future. We foresee the feasibility of redesigning the human condition, including such parameters as the inevitability of aging, limitations on human and artificial intellects, unchosen psychology, suffering, and our confinement to the planet earth.

- Systematic research should be put into understanding these coming developments and their long-term consequences.

- Transhumanists think that by being generally open and embracing of new technology we have a better chance of turning it to our advantage than if we try to ban or prohibit it.

These opinions may also be reflected in the design of technology, particularly in AI models. Elon Musk, for example, has attempted with limited success so far,[70] to create a generative AI that supports his libertarian vision and serve as a conservative counterpart to AI models like *ChatGPT* or *Gemini* developed by *OpenAI* and *Google*.

Beyond the sources used to train these generative AIs, which can be a critical factor in the responses they provide, their engineers can also introduce, with obvious risks of bias and discrimination, their own moral priorities. As Brian

[69] The Transhumanist Declaration;
Susan Schneider Artificial You: AI and the Future of Your Mind (Princeton University Press 2019)

[70] I tried X's 'anti-woke' Grok AI chatbot. The results were the opposite of what I expected (ZDNET Dec. 21, 2023) https://www.zdnet.com/article/i-tried-xs-anti-woke-grok-ai-chatbot-the-results-were-the-opposite-of-what-i-expected/

Christian pointed, some AI model designers might even challenge the validity of our societies' moral values in guiding their work.

> *"Some, for instance, worry that humans aren't a particularly good source of moral authority. "We've talked a lot about the problem of infusing human values into machines," says Google's Blaise Agüera y Arcas. "I actually don't think that that's the main problem. I think that the problem is that human values as they stand don't cut it. They're not good enough."*[71]

In contrast, Jaron Lanier, one of the pioneers of virtual reality, emphasizes the political and social dangers posed by movements like transhumanism and longtermism. He claims that these ideas represent a new form of technological determinism that carries major social and cultural risks: *"I do not think the technology is creating itself. It's not an autonomous process. The reason to believe in human agency over technological determinism is that you can then have an economy where people earn their own way and invent their own lives. If you structure a society on not emphasizing individual human agency, it's the same thing operationally as denying people clout, dignity, and self-determination…"*[72]

B. AI and Cultural Transformations of Societies

Generative AIs are already beginning to transform certain creative professions and could pose a threat to others, as illustrated by the recent strikes by writers, voice artists and actors in the U.S. over concerns of being replaced by generative AIs in the near future. For Yuval Harari, the next step could be entirely 'autonomous' cultural creations designed by AI which would colonize minds and alter human thoughts and dreams. While artistic creations draw on the author's experiences, culture and values, these new creations could potentially create new dogmas, and

[71] Brian Christian. The Alignment Problem, Machine Learning and Human Values (p. 247) (W. W. Norton & Company 2020)

[72] Lanier, Jaron: "Who Owns the Future?" (Simon & Schuster 2013)

even new cults, which could unpredictably transformation of the human imagination and the way in which societal behavior.

> *What will happen to the course of history when AI takes over culture, and begins producing stories, melodies, laws and religions? Previous tools like the printing press and radio helped spread the cultural ideas of humans, but they never created new cultural ideas of their own. AI is fundamentally different. AI can create completely new ideas, completely new culture. At first, AI will probably imitate the human prototypes that it was trained on in its infancy. But with each passing year, AI culture will boldly go where no human has gone before. For millennia human beings have lived inside the dreams of other humans. In the coming decades we might find ourselves living inside the dreams of an alien intelligence.*[73]

Many social norms and customs have been modified by the rise of digital technologies. Social media, for instance, have transformed our relationship with personal data disclosure. Devices like augmented reality glasses could represent further disruptions to interpersonal norms. The potential for continuous video and audio recording might become the new norm in conversations and exchanges. For tech industry players whose activity is based on the massive use of personal data for advertising purposes, any constraint on data collection becomes not only an economic obstacle but also a political nuisance. This led Eric Schmidt, former CEO of *Google*, to state that we need to change our conception of privacy rather than attempt to preserve it.[74] Indeed, some industry actors believe that it is preferable to modify our societies' social or cultural norms to enable the development of their technologies. In the extreme, it would not only be necessary to modify these social or cultural norms, but also the genetic heritage that

[73] Yuval Noah Harari argues that AI has hacked the operating system of human civilisation (The Economist Apr 28, 2023) https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation

[74] Google CEO Eric Schmidt Dismisses the Importance of Privacy (Electronic Frontier Foundation Dec 10, 2009) https://www.eff.org/deeplinks/2009/12/google-ceo-eric-schmidt-dismisses-privacy

constitutes humans themselves to make them 'compatible' with certain technological evolutions.

C. From Transhumanism to Eugenics

> *"Captain, although your abilities intrigue me, you are quite honestly inferior. Mentally, physically. In fact, I am surprised how little improvement there has been in human evolution. Nothing ever changes, except man. Your technical accomplishments? Improve a mechanical device and you may double productivity. But improve man and you gain a thousandfold. I am such a man. Join me. I'll treat you well..."*
>
> Khan Noonien Singh, genetically augmented tyrant addressing Captain Kirk in the "*Star Trek"* episode *Space Seed* (1967)[75]

New AI services could be made accessible by modifications that are not connected to changes to social and cultural codes, but rather to alterations of the genetic code itself. Beyond their medical applications, advances in genomics could therefore be used to create new generations of services designed to 'augment' human capabilities. Next-generation services could be rooted not in traditional human–machine interfaces or virtual reality but directly in connected to users' brains. This is the goal of companies specializing in neurotechnologies, such as *Neuralink* founded by Elon Musk in 2016: to develop an implanted brain–computer interface. The initial goal is to restore new forms of mobility or interaction for people with disabilities, but the next stage is wellness services, and even 'memory extensions' and altered consciousness, could be created for 'augmented' humans. Chinese neuroscientist Mu-ming Poo, of the *Shanghai-based Institute of Neuroscience*, warns of the risks and the ethical issues related to neurotechnologies abuses: *"a more difficult ethical issue is its use for augmenting normal brain functions rather than for therapeutic purposes, applications such as non-invasive*

[75] http://www.chakoteya.net/StarTrek/24.htm

neuromodulation of sleep/wakefulness, attention, emotion, as well as learning and memory functions".[76]

To achieve this fusion between humans and machines, some people already envision genetic modifications that would make humans better suited to connecting with technological systems.[77] There would be some sort of continuity between the concepts of transhumanism and eugenics that would, according to their advocates, genetically enhance human beings to increase their 'performance' and allow them to better interact with the technological universe. This would raise the exact opposite question to that addressed by specialists in *Artificial General Intelligence (AGI):* it is not about when AI will become equal or superior to humans, but rather when humans will cease to be human once 'augmented'. And at what point these genetic modifications and/or brain implants will transform these individuals into 'non-humans'. What distinctively human capacities will they lose by having a connected brain? Which brain characteristics will be altered by genetic modifications? Which current functions of our brains will be deemed unnecessary by the creators of these genetic or technological modifications?

> ***"It is not a question of when AI will become equal or superior to humans, but rather when 'augmented' individuals will stop being human... At what point will genetic modifications [and/or brain implants] transform these individuals into 'non-humans'...***

[76] Mu-ming Poo, China's new ethical guidelines for the use of brain–computer interfaces, National Science Review, Volume 11, Issue 4, April 2024, nwae154, https://doi.org/10.1093/nsr/nwae154 https://academic.oup.com/nsr/article/11/4/nwae154/7668215

[77] Neurotechnology: Scientific and Ethical Challenges (Parliamentary Office For Scientific and Technological Assessment (OPECST Jan 2022)
https://www2.assemblee-nationale.fr/content/download/481455/4686960/version/1/file/OPECST_2022_0032_neurotechnologies_eng.pdf

Paradoxically, it is the simultaneous rise of artificial intelligence and biotechnologies that could force our societies to redefine the boundaries between what we consider as socially or biologically acceptable and what is not for the future of the human species. Democratic debates like those that took place on bioethics must also be had to determine what modifications could be permitted and which, as for AI regulation, would entail unacceptable risks for each individual and for our societies as a whole.

VI. TECHNOLOGICAL PROSPECTS OF AI AND POLITICAL DEVELOPMENTS

A.Still Far from Artificial General Intelligence

When generative AI presents statements that seem true but are in fact false, their developers often use the term 'hallucination'. This is explained by the fact that these systems lack common sense and an understanding of the context in which events take place. AI cannot differentiate facts from incorrect statements. Despite efforts to perfect their models and fine-tune the data used to train AI, they can still provide responses that contradict obvious facts. In a way, all responses generated by AI are merely hallucinations, deemed useful or acceptable by those on the receiving end.

There is an ongoing debate among AI researchers regarding the potential emergence of Artificial General Intelligence (AGI) or even *Artificial Superintelligence (ASI)*. Some believe the fundamental limits of AI technologies could render this impossible. Similar debates exist on the possibility of designing fully fledged quantum computers that could transform both the global computing landscape and the field of AI. World specialist in quantum entanglement, Alain Aspect, who was awarded the 2022 Nobel Prize for Physics, does not rule out the possibility that fundamental barriers could prohibit the realization of large quantum computers:

> *The "ideal" quantum computer, which is the one we hear about, does not yet exist. Personally, I think it's so hard to pull off that I doubt I'll ever see it in my lifetime. I am not pessimistic; however, I firmly believe that when something is very difficult technically, but not impossible because of a fundamental law of physics, it always ends up being achieved. Because if we ever stumbled on a fundamental limit, we would have answered this essential question of the 'threshold' between the classical world and the quantum world. At the moment, we don't have the faintest idea. So one of two things: either we manage to build the quantum computer and it will be wonderful, or we come up against a limit, and all the people who have put a lot of money into this research will be very disappointed…*[78]

If the possibility of creating *AGI* as imagined by proponents of transhumanism, was brought into question by scientific obstacles, this would bring a halt to the 'narratives' that have convinced AI investors that an *El Dorado* was within reach. The design of *AGI* by tech creators is still far off. Although it is possible for AI to carry out complex procedures based on vast datasets during their training, the simplest tasks done by humans seem to still be beyond their capacity. In 1988, robotics expert Hans Moravec commented on the limits of AI in these words: *"It's comparatively easy to make computers exhibit adult-level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception or mobility."*[79]

[78] Alain Aspect, Nobel Prize in Physics winner, in Le Nouvel Esprit Public June 25, 2023 https://www.lenouvelespritpublic.fr/podcasts/446

[79] Moravec, Hans. Mind children: The future of robot and human intelligence. Harvard University Press (1988)
https://digitaleconomy.stanford.edu/news/the-turing-trap-the-promise-peril-of-human-like-artificial-intelligence/

Three decades later, even proponents of transhumanist theories like Ray Kurzweil and Nick Bostrom acknowledge that today's most powerful machines are still a long way off from the capacities of the human brain. For philosopher Susan Schneider, the enthusiasm surrounding the capabilities of new AI should not obscure the fact that these remain limited.

> *Of course, speed is not everything. If the metric is not arithmetic calculations, your brain is far more computationally powerful than Summit. It is the product of 3.8 billion years of evolution (the estimated age of life on the planet) and has devoted its power to pattern recognition, rapid learning, and other practical challenges of survival. Individual neurons may be slow, but they are organized in a massively parallel fashion that still leaves modern AI systems in the dust.*[80]

But beyond computing power or the ability to adapt to new situations, it is the very concept that a machine could one day acquire a new form of consciousness that continues to be at the heart of discussions within the AI expert community.

Despite the certainty of proponents of transhumanism and technological singularity, fundamental limitations to the achievement of *AGI* may well be discovered in the years to come despite the considerable financial, technological and human resources that have been devoted to the development of AI. But even before these AIs can one day be seen —if they ever are— as comparable to human intelligence, their social, political and cultural effects will be significant for our societies.

B. From Science Fiction to Industrial Reality

Although current artificial intelligences are still far from the sentient beings depicted in science fiction—whether in *2001: A Space Odyssey, Ex Machina,* or *Her*—many of the technologies we use today were originally envisioned by science fiction writers. For instance, the creators of *Star Trek* conceived several

[80] Susan Schneider: Artificial You: AI and the Future of Your Mind (Princeton University Press 2019)

innovations that have since become commonplace, including mobile phones, tablets, portable diagnostic devices, intelligent voice interfaces, and shared virtual worlds. Amazon's head of Devices even remarked that the voice-activated computer on *Star Trek* served as the "north star" for the design of the AI voice assistant *Alexa*.[81]

The connection between high-tech industries and the world of science fiction has never been stronger, yet science fiction also often explores a broad range of dystopian futures, including themes of pervasive surveillance, dehumanization, psychological manipulation, and extensive genetic modification. While the emergence of malevolent AI is a common trope, the real challenge for future generations may lie in preventing scenarios that blend aspects of the surveillance state depicted in *Minority Report* and the eugenics-driven society of *Gattaca* from becoming reality. In *Minority Report*, the oracles could be seen as a narrative device representing an embodied version of crime predictions generated by AI models. As Eric Schmidt, Henry Kissinger, and Daniel Huttenlocher have noted, as AI becomes increasingly integrated into every aspect of our lives, voluntary submission to AI's opinions or predictions will no longer be a distant fantasy, but a concrete reality.

> *As AI is woven into the fabric of daily existence, expands that existence, and transforms it, humanity will have conflicting impulses. Confronted with technologies beyond the comprehension of the nonexpert, some may be tempted to treat AI's pronouncements as quasi-divine judgments. Such impulses, though misguided, do not lack sense. In a world where an intelligence beyond one's comprehension or control draws conclusions that are useful but alien, is it foolish to defer to its judgments? Spurred by this logic, a re-enchantment of the world may ensue, in which AIs are relied upon for oracular pronouncements to which some humans defer without question.*[82]

[81] How Star Trek inspired Amazon's Alexa (VentureBeat, Jun 7, 2017) https://venturebeat.com/ai/how-star-trek-inspired-amazons-alexa/

[82] Henry A. Kissinger, Eric Schmidt and Daniel Huttenlocher. "The Age of AI: And Our Human Future." (John Murray Press, 2021)

However, there are policy and high-security domains where voluntary submission to the opinions of AI poses significant risks. In this context, Henry Kissinger, who played a key role in international disarmament negotiations, reminds us that negotiations between blocs—particularly between the United States and the Soviet Union—were initiated as soon as nuclear-armed powers came into direct confrontation. We are still far from witnessing similar negotiations aimed at limiting the offensive use of AI. This is especially true given that AI can be employed by organizations and groups acting on behalf of states, thus providing a degree of plausible deniability for the instructing powers.

Nevertheless, the specific challenges associated with integrating AI into the nuclear chain of command make negotiations in this highly contentious area increasingly likely. Ben Buchanan, White House Special Advisor on AI, affirms that, at present, AI technologies are not suitable for conflicts involving nuclear weapons systems.

> *Since the military applications of AI favored by the warriors are distinct from the scientific applications preferred by the evangelists, most machine learning systems are not ready for adversarial environments like war against opponents with good hacking capabilities; much more iterative testing and evaluation must be done before these systems can be considered ethically and strategically appropriate for combat. Due to these major concerns about security and robustness, some high-stakes uses of machine learning—such as in nuclear command-and-control systems—are manifestly unwise for the foreseeable future.*[83]

Nuclear security experts at the *European Leadership Network* also emphasize that the risks are too great to consider deploying AI systems for the authorization of nuclear strikes. They even advocate for the implementation of an international moratorium in this domain.

[83] Buchanan, Ben; Imbrie, Andrew. The New Fire: War, Peace, and Democracy in the Age of AI (p. 244). MIT Press 2022

> *From a global security perspective, AI introduces risks of miscalculations, misperceptions and misinterpretations, whether from system failures, vulnerabilities, or misuse, potentially leading to inadvertent or accidental escalation [...] If incorporated into strategic decision- making systems, these models pose some of the gravest risks due to their opaque nature, unpredictability, and susceptibility to cyber- attacks[...] The 'black box' nature of some AI models makes its decision- making process very challenging to decipher with our current understanding of these models [...] If integrated in strategic decision-making systems, this might leave no accountability systems and no method of verification for AI predictions and decisions. Considering the significant risks posed by cutting-edge deep learning models, P5 states should impose a moratorium on the integration of these models into specific NC3 areas with the highest integration risks.*[84]

C. Toxic Economic Models and AI Data

Economic models based on attention and the hyper-individualization of messages entail an ever-increasing extraction of personal data. There is now a tension between the objectives of the manufacturers of AIs who seek to use ever-growing datasets to improve their AI models and citizens who could be stripped of their personal data, and unable to control what subsequently becomes of it (and of themselves). As organizational psychologist Tomas Chamorro-Premuzic explains, a vicious circle derived from the economic model of advertising platforms has been established. A development that has had individual, social and political consequences.

> *The quantification of our attention—and AI's ability to weaponize that information—creates a vicious cycle: since our attention is scarce and information is plentiful, the battle for our attention exacerbates. Netflix is competing with Twitter, Twitter is competing with the New York Times, and the*

[84] Alice Saltini: AI and nuclear command, control and communications: P5 perspectives (The European Leadership Network Nov 2023) https://www.europeanleadershipnetwork.org/wp-content/uploads/2023/11/AVC-Final-Report_online-version.pdf

> *New York Times is competing with Instagram; they're all competing for our precious time and our even more precious focus. Their algorithms crave our attention, and their business models depend on it, which makes our attention highly valuable, not least because there is so little of it left to capture after the algorithms consume it. It also leads to yet more information overload, which threatens to distract us even more. The result is a degradation of focus that causes attention deficit hyperactivity disorder (ADHD)–like behaviors, such as restless hyperactivity, rapid boredom, and impulsivity.*[85]

There is a noticeable gap in the reasoning of many AI specialists, such as Neil Lawrence,[86] who cannot (or do not wish to) imagine that restrictions on access to personal data could be implemented. Only minimal forms of compensation are considered by companies that use this data for their profit, which prolongs a tempting but flawed analogy: data is seen as the "new oil" of the 21st century economy. In contrast, Shoshana Zuboff advocates for imposing restrictions on the massive extraction of data and, consequently, on the economic models of these platforms. The ever-increasing collection of personal data now poses not only a political risk of manipulation but also an economic risk, to the point that even experts from the *International Monetary Fund* have expressed concerns: *Most transactions involving personal data are unbeknownst to users, who likely aren't even aware that they have taken place, let alone that they have given permission. This gives rise to what is known in economics as an externality: the cost of privacy loss is not fully considered when an exchange of data is undertaken. The consequence is that the market's opacity probably leads to too much data being collected, with too little of the value being shared with individuals."*[87]

As the scandals surrounding the uncontrolled use of data by major Internet platforms have demonstrated, these companies cannot be relied upon to uphold

[85] Chamorro-Premuzic, Tomas. I, Human: AI, Automation, and the Quest to Reclaim What Makes Us Unique (p. 40) (Harvard Business Review Press 2023)

[86] Neil D. Lawrence. The Atomic Human: What Makes Us Unique in the Age of AI (Public Affairs 2024)

[87] Let's Build A Better Data Economy by Yan Carrière-Swallow and Vikram Haksar (International Monetary Fund March 2021) https://www.imf.org/external/pubs/ft/fandd/2021/03/pdf/how-to-build-a-better-data-economy-carriere.pdf

democratic principles when doing so directly conflicts with their industrial and financial objectives. An example of such corporate behavior is discussed by Max Fisher in his book *The Chaos Machine*, where he examines *Facebook's* role in fostering the ideological radicalization of its users.

> *I later learned that, a short time before my visit, some Facebook researchers, appointed internally to study their technology's effects, in response to growing suspicion that the site might be worsening America's political divisions, had warned internally that the platform was doing exactly what the company's executives had, in our conversations, shrugged off. "Our algorithms exploit the human brain's attraction to divisiveness," the researchers warned in a 2018 presentation later leaked to the Wall Street Journal. In fact, the presentation continued, Facebook's systems were designed in a way that delivered users "more and more divisive content in an effort to gain user attention & increase time on the platform." Executives shelved the research and largely rejected its recommendations, which called for tweaking the promotional systems that choose what users see in ways that might have reduced their time online.*[88]

When not considered by AI developers upstream, and above all when deliberately omitted, these moral concerns can have lasting consequences on users of these technologies. Steven Kerr of Ohio State University sums it up as follows: *"If the reward system is so designed that it is irrational to be moral, this does not necessarily mean that immorality will result. But is this not asking for trouble?"*[89]

VII. AI AND HEALTH: INDUSTRIAL STRATEGIES AND POLITICAL CONSEQUENCES

[88] Max Fisher. The Chaos Machine The Inside Story of How Social Media Rewired Our Minds and Our World (pp. 14-15). (Quercus Editions Ltd 2022)

[89] Steven Kerr. The Academy of Management Executive (1993-2005) Vol. 9, No. 1 pp. 7-14 (Feb 1995) https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Motivation/Kerr_Folly_of_rewarding_A_while_hoping_for_B.pdf

A. Health Data: Political and Industrial Prospects

The evolution of AI technologies, in addition to assisting in the diagnosis of diseases, also makes it possible to analyze data from sensors embedded in our environment for the early detection of pathologies. This approach can be applied not only at the individual level but also at the level of entire populations. Even before the *COVID-19* pandemic, many scientists had suggested integrating sensor networks and AI systems to detect pandemics at an early stage and develop protective measures. In an extreme scenario, according to the research firm, *Frost & Sullivan* the most 'efficient' response in terms of combating pandemics would be to establish a global network of sensors. This network would enable the early detection of biological threats worldwide. However, as the authors note, such a proposal, which would be the most important technological market opportunity ever conceived for the Internet of Things *(IoT),* would face opposition from public opinion due to its intrusive nature:

> *The simple answer might for enterprises, cities, and national governments to collectively create a massive global network of sensors to detect viruses. However, this would require planning and implementation on a global scale that would tax the very foundations of democracy and obligate governments to place the needs of the planet ahead of the needs of their citizens. The most logical solution is often the most difficult to implement. The amount of planning required to make this solution a reality would, arguably, make it one of the most significant achievements in the history of mankind. […] I find this to be the "holy grail" of IoT opportunity in the long term.*[90]

The fascination with 'techno-efficiency' seen among some IoT tech players echoes the '*technological solutionism*' described by Evgeny Morozov in his book *To Save Everything, Click Here*.[91] This fascination is even more pronounced when

[90] The Next Generation of IoT – Addressing the Coronavirus and Preventing Future Outbreaks (Frost & Sullivan, Jan 31, 2020) ww2.frost.com/frost-perspectives/the-next-generation-of-iot-addressing-the-coronavirus-and-preventing-future-outbreaks/

[91] To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist (Evgeny Morozov, Ed. Penguin 2013)

it comes to AI technologies. The trend is not limited to authoritarian regimes but is also seen in economic actors or governments who assess the benefit/risk ratio of these technologies and consider that democracy and the protection of freedoms can become adjustment variables.

> ***“The only privacy that's left is the inside of your head. Maybe that's enough...***
>
> **Excerpt from the movie**
> ***"Enemy of the State"***
> ***(Tony Scott 1998)***

IoT technologies have diversified as new methods of analysis based on AI have been developed, whether these be *wearables* (connected headsets, rings, glasses or wristbands fitted with sensors), drones for thermal imaging or urban sensor networks.[92]

AI can thus reveal or infer information on a person’s health from seemingly trivial data. As cardiac or circulatory disorders can be predicted by analyzing a person’s movements over time, *Facebook* has even developed patents related to the continuous analysis of its users’ movements and behavior that assess their health parameters.[93] In 1998, in the Tony Scott movie *Enemy of the State*, the *NSA* official played by Jon Voight, remarks on surveillance: *“The only privacy that's left is the inside of your head. Maybe that's enough.”*[94] This ultimate obstacle to knowing everyone’s thoughts could be overcome in the coming years if AI is capable of directly decoding the unfolding of visual or

[92] Internet of Things for Current COVID-19 and Future Pandemics: An Exploratory Study by M. Nasajpour, S. Pouriyeh, M. Parizi, M. Dorodchi, M. Valero, H. Arabnia Dept of Information Technology and Dept of Software Engineering and Game Development, Kennesaw State University, Department of Computer Science, University of North Carolina and University of Georgia (submitted Jul 22, 2020) arxiv.org/pdf/2007.11147.pdf

[93] What 7 Creepy Patents Reveal about Facebook (New York Times, Jun 21, 2018) nytimes.com/interactive/2018/06/21/opinion/sunday/facebook-patents-privacy.html

[94] Enemy of the State (Movie Transcript) http://movietranscript.blogspot.com/2016/01/1998-enemy-of-state-english-transcripts.html

thought representations in the brain. Over time, these innovations - which are already sparking interest among researchers when it comes to creating new communication methods for people with disabilities - could be used for 'non-medical' applications to analyze brain processes. This could enable subtle manipulation of people's behavior and thoughts. Several companies in China, for example, are experimenting with headsets that monitor brainwaves detect stress, anger, or drowsiness in employees.[95] The *Meta* group labs are also beginning to use generative AI models to reconstruct mental imagery by analyzing brain activity monitored by magnetic resonance imaging *(MRI)* systems.[96] Likewise, *EEG* sensors connected to AI systems are now being used to translate brain scans into words and sentences.[97]

The measures taken to ensure the security of the data collected by these sensors will determine whether these technologies are used for prevention, control or manipulation of individuals. As several authors have noted, there is a fine line between the prevention of risky behavior and the conditioning of users via connected devices. This trend towards controlling populations could in effect mark the transition from social engineering to social control assisted by AI following the example of the Chinese *Social Credit System*.[98]

[95] En Chine, des capteurs cérébraux pour surveiller les émotions des employés (Slate May 1, 2018) www.slate.fr/story/161173/en-chine-des-capteurs-cerebraux-pour-surveiller-les-emotions-des-employes

[96] Meta recreates mental imagery from brain scans using AI (Venture Beat, Oct 18, 2023) https://venturebeat.com/ai/meta-recreates-mental-imagery-from-brain-scans-using-ai/

[97] Scientists use AI to decipher words and sentences from brain scans (Science.org, May 1, 2023) https://www.science.org/content/article/scientists-use-ai-decipher-words-and-sentences-brain-scans

[98] The Social Credit System: Not Just Another Chinese Idiosyncrasy (Eunsun Cho - Journal of Public and International Affairs, Princeton University, May 1, 2020) https://jpia.princeton.edu/news/social-credit-system-not-just-another-chinese-idiosyncrasy

B. Moving the Center of Gravity of the Health Economy

As AI technologies gain momentum, the collection and processing of health data have become strategic challenges for technology companies. Advances in connected medical devices now enable the establishment of more cost-effective, personalized prevention and monitoring systems, capable of continuously tracking users' activities and physiological parameters. These devices, which interact with users daily, could facilitate the early diagnosis of chronic diseases such as cancer, diabetes, asthma, and cardiovascular disease.

Recent versions of the Apple Watch, for example, which include electrocardiogram (ECG) functions, are expected to detect sleep apnea by analyzing users' sounds and movements through AI systems. These connected health devices may also help modify user behavior to manage risk factors like a sedentary lifestyle and obesity.

The health insurance sector has become a key target for major technology companies, enabling them to gain a foothold in the insurance market while shifting the focus or the center of gravity of the healthcare economy toward prevention. Health expenditures are currently concentrated on treatment,[99] but with next-generation AI services in connected health, these companies are poised to create a global market for health prevention. A report by *Goldman Sachs* on connected health[100] estimated that the introduction of connected devices in the healthcare sector could result in savings of $305 billion annually in the United States alone. Of these savings, $200 billion would stem from improvements in the prevention and management of chronic diseases, particularly cardiovascular

[99] Santé : « En France, la culture de la prévention n'est clairement pas acquise » (David Simard, Le Monde Aug 27, 2023) www.lemonde.fr/idees/article/2023/01/05/sante-en-france-la-culture-de-la-prevention-n-est-clairement-pas-acquise_6156771_3232.html

[100] The Digital Revolution comes to US Healthcare (Goldman Sachs Equity Research 2015) www.anderson.ucla.edu/Documents/areas/adm/acis/library/DigitalRevolutionGS.pdf

disease, asthma, and diabetes. These savings would represent nearly 10% of total health expenditure in the United States ($4.3 trillion in 2023).[101]

However, the potential for prevention measures to infringe upon individual freedoms is a growing concern, particularly given the substantial impact such measures could have on the fiscal balance of developed nations. This dynamic of risk prevention was evident in the *HR1313* bill, introduced in the U.S. Congress in 2017 by the Trump Administration. The bill aimed to implement large-scale genetic testing in workplaces to prevent and detect diseases early. It proposed penalties of $5,000 annually for employees who refused to undergo genetic screening.[102] The bill was ultimately not adopted, thanks to opposition from Democratic members of Congress.

C. Health Data: What Strategy in Europe?

In the field of AI applied to healthcare, the creation of a *Health Data Hub* by the French Ministry of Health represents a key component of France's open health data strategy. This platform is designed as a one-stop shop for accessing all health data on French citizens, with the aim of promoting the development of new AI-driven services in healthcare, including diagnostic tools, treatment monitoring, and preventive measures. However, as noted by healthcare professionals and medical IT experts, hosting this hub on *Microsoft Azure*[103] raises both risks to data

[101] National Health Expenditures 2023 (U.S. Centers for Medicare & Medicaid Services 2019) https://www.cms.gov/data-research/statistics-trends-and-reports/national-health-expenditure-data/nhe-fact-sheet

[102] Employees who decline genetic testing could face penalties under proposed bill (Washington Post, March 11, 2017) https://www.washingtonpost.com/news/to-your-health/wp/2017/03/11/employees-who-decline-genetic-testing-could-face-penalities-under-proposed-bill/

[103] « L'exploitation de données de santé sur une plate-forme de Microsoft expose à des risques multiples » (article published in Le Monde on Dec 10, 2019) www.lemonde.fr/idees/article/2019/12/10/l-exploitation-de-donnees-de-sante-sur-une-plate-forme-de-microsoft-expose-a-des-risques-multiples_6022274_3232.html

sovereignty and concerns over missed opportunities to cultivate critical expertise within the French digital health ecosystem.[104]

Regarding the hosting of the *Health Data Hub*, a report by the French Data Protection Authority (*CNIL*) in 2020 emphasized that "the shift in hosting solutions for the Health Data Hub and other health data repositories hosted by companies subject to U.S. law should occur as soon as possible."[105] Access to health data is a crucial strategic issue for technology companies, as French health data represents one of the most comprehensive and structured datasets among developed nations. Due to its quality and breadth, these datasets in France and Europe provide a powerful leverage point for training AI models and enhancing their performance.

In response to the controversy surrounding the choice of *Microsoft*, the French government initially announced plans to relocate hosting to French or European companies.[106] However, despite parliamentary efforts to obtain information on the progress of this transition, it has yet to materialize.[107] The protection of health data will likely become even more significant with the upcoming implementation of the *European Health Data Space*[108], as part of the broader European data strategy.[109] In this context, the recent decision by the *CNIL* to authorize *Microsoft* to host the

[104] Health Data Hub: « Le choix de Microsoft, un contresens industriel ! » (interview with Bernard Benhamou in Le Point, June 18, 2020) www.lepoint.fr/technologie/health-data-hub-le-choix-de-microsoft-et-un-contresens-industriel-10-06-2020-2379394_58.php

[105] In wake of the Schrems II, CNIL challenges use of Microsoft cloud storage to host public health data lakes (the Health Data Hub case – Part 1) Hogan Lovells Oct 15, 2020) https://www.hoganlovells.com/en/publications/in-wake-of-the-schrems-ii-cnil-challenges-the-use-of-microsoft-cloud-storage-to-host-public-health-data-lakes-the-health-data-hub-case-part-1_1

[106] Microsoft doit se retirer du Health Data Hub, d'après la Cnil (L'Usine Digitale Oct 9, 2020) www.usine-digitale.fr/article/microsoft-doit-se-retirer-du-health-data-hub-d-apres-la-cnil.N1014634

[107] Le député Philippe Latombe saisit la Cada pour obtenir le "benchmark" de l'hébergement du Health Data Hub (TIC Pharma May 11, 2023) https://www.ticpharma.com/story?ID=2268

[108] European Health Data Space https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en

[109] Communication from the Commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the Regions. A European strategy for data, February 19, 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066

European *EMC2* project raises concerns, particularly due to the risks posed by U.S. extraterritorial laws—especially the Foreign Intelligence Surveillance Act (FISA).[110] This Act allows U.S. intelligence agencies to request data processed by U.S. companies, regardless of its physical location.[111]

A key challenge for public authorities in France and Europe will be ensuring that AI developments in healthcare do not undermine the social model in favor of systematized control over individuals. This risk becomes even more pronounced if genomic data are integrated into these platforms. A shift towards hyper-individualization in health coverage could erode the social model, which is based on solidarity and the pooling of risks across the population. Soon, safeguarding the social welfare model will increasingly intersect with concerns around digital sovereignty.

At present, social networks combined with connected health devices already enable major platforms to enhance their profiles of users. This integration provides opportunities for the development of AI models tailored to precise health risk assessments for individuals. As Pascal Demurger, Managing Director of *MAIF*, has stated: *"This has upended the world of insurance. Insurers traditionally had very little data on their clients but a large number of clients. Thanks to big data, we can now collect a large amount of behavioral data on each person."*[112]

Rather than directly entering the healthcare sector, which requires significant and high-risk long-term investments, major Internet platforms can leverage user data

[110] HDH: des fournisseurs français attaquent la décision de la Cnil favorisant Microsoft (Le Monde Informatique March 18, 2024)
https://www.lemondeinformatique.fr/actualites/lire-hdh-des-fournisseurs-francais-attaquent-la-decision-de-la-cnil-favorisant-microsoft-93256.html

[111] L'extension des prestataires américains devant collaborer avec la NSA fait polémique (Jean-Marc Manach - Next April 22, 2024)
https://next.ink/135019/lextension-des-prestataires-americains-devant-collaborer-avec-la-nsa-fait-polemique/

[112] Santé: faut-il faire payer les assurés en fonction de leur mode de vie ? (Le Monde, Sept 6, 2016)
www.lemonde.fr/economie/article/2016/09/06/assurance-votre-vie-privee-vaut-bien-une-ristourne_4993378_3234.html

to offer insurance and disease prevention services. With access to detailed user data, these platforms can accurately model health risks for individuals and optimize their insurance services for profit. *Google* (through its parent company Alphabet) has already ventured into the health insurance sector with its *Coefficient* division[113], while health and fitness services have become a strategic priority for Apple, particularly through its connected watches and smartphones. Additionally, Facebook is researching medical applications in virtual environments as part of its metaverse initiative.[114]

AI technologies have the potential to reshape all aspects of healthcare systems. If non-European platforms dominate the development of AI technologies related to care, prevention, and health insurance within Europe, they could undermine our social models by prioritizing economically driven approaches over social or democratic considerations. This may lead to the implementation of preventive measures linked to variations in health insurance premiums based on individual behavior. In the absence of a coherent industrial policy—critical for fostering European alternatives—these changes could have significant collateral effects, including violations of individual rights and increased risks related to mass surveillance.

[113] Verily (Alphabet) se lance dans l'assurance avec sa nouvelle division, Coefficient (L'Usine Digitale, Aug 25, 2020) www.usine-digitale.fr/article/verily-alphabet-se-lance-dans-l-assurance-avec-sa-nouvelle-division-coefficient.N996864

[114] Meta is forcing Apple into virtual reality (Quartz Mar 11, 2022) https://qz.com/2141093/meta-is-using-quest-2-to-force-apple-into-the-metaverse-early

VIII. AI ARCHITECTURE AND GOVERNANCE

Commercial large language models (*LLMs*) were initially developed in a vertical and centralized manner, with each new generation of models incorporating an increasing number of parameters (ranging from several billion to trillions in the most recent models). For example:

- In May 2020, GPT-3 had 175 billion parameters and was trained on 45 terabytes of text data.
- In March 2023, GPT-4 surpassed 1 trillion parameters and was trained on over 1,000 terabytes of data.[115]

This race for computational power has been characterized by intense competition among major cloud providers, who have leveraged their computing and storage capabilities to develop successive generations of AI, raising concerns about the potential emergence of an AI duopoly between *Microsoft* and *Google*[116]. As Tim Wu, an expert in the regulation of U.S. technology companies, argues in his book *The Curse of Bigness,*[117] the risks associated with monopolies are not only economic but also political. The major tech companies already exert significant influence over virtually all aspects of human activity, including the functioning of democratic systems. With the continued growth of AI, and the extreme concentration of data within these large platforms, their power could extend even further.

However, the construction of these new "cathedrals" of AI code and data may face new economic and political challenges. The financial, human, and energy resources required to build and maintain these AI systems could complicate the

[115] LLMs hitting 2 trillions parameters (IEEE Future Directions Nov 14, 2023) https://cmte.ieee.org/futuredirections/2023/11/14/llms-hitting-2-trillions-parameters/

[116] The AI Boom That Could Make Google and Microsoft Even More Powerful (Wall Street Journal, Feb 11, 2023) https://www.wsj.com/articles/the-ai-boom-that-could-make-google-and-microsoft-even-more-powerful-9c5dd2a6

[117] Tim Wu. The Curse of Bigness: Antitrust in the New Gilded Age (Columbia Global Reports 2018)

achievement of economic equilibrium, whether in terms of the cost of acquiring dedicated processors (such as *GPUs*), the expenses associated with training and operating these models in data centers, or the energy demands needed to run them.

While investors were initially enthusiastic about the prospect of creating *"Artificial General Intelligence,"* doubts are now emerging about whether the technological and economic objectives of large language models (*LLMs*) will be met. The companies behind the largest *LLMs*—led by *OpenAI*, *Microsoft*, *Meta*, and *Google*—are now confronted with questions regarding the technological, financial, and environmental sustainability of this AI-driven race for resource consumption. Investors such as Eric Schmidt predict that the increase in energy consumption driven by AI is inevitable, warning that the growing demand for energy and raw materials may make it impossible to meet global warming reduction targets, pinning hopes on the idea that AI itself may one day help limit CO2 emissions.[118]

In response to these risks, new LLM architectures may be developed in the future that are both decentralized and more modest in scale. As noted by the *Wall Street Journal*, the race for computational power may not be suitable for all corporate needs.

> *"A giant LLM [large language model] that's been trained on the entire World Wide Web can be massive overkill," said Robert Blumofe, chief technology officer at cybersecurity, content delivery and cloud computing company Akamai. For enterprise use cases, he said, "You don't need an AI model that knows the entire cast of 'The Godfather,' knows every movie that's ever been made, knows every TV show that's ever been made." Oliver Parker, vice president of global generative AI go-to-market at Google Cloud, said he has seen enterprises shifting to midsize*

[118] Former Google CEO Eric Schmidt Says 'We Are Never Going To Meet Our Climate Goals' – Pushes For AI And Data Centers To Solve The Crisis (Yahoo Finance, Oct 10, 2024) https://finance.yahoo.com/news/former-google-ceo-eric-schmidt-161538746.html

> *models in the last three months, in part because the models meet criteria capturing a lot more enterprise use cases.*[119]

Companies such as OpenAI are beginning to release smaller versions of their large models and are now exploring the development of AI solutions that operate locally.[120] This shift also offers the advantage of giving the companies using these AIs greater control over their data. Data protection, both for the data used to train these models and for the data input by users, remains a significant shortcoming of AI services hosted by major companies. New cryptographic methods are currently being investigated for processing encrypted data within AI systems, based on the principle of "homomorphic encryption." The goal of this research is to safeguard both personal data and industrial data. It is worth noting that Europe is home to some of the world's leading experts in cryptography and could potentially develop new generations of technologies combining AI and cryptographic security in the future.[121]

Another promising area of research aimed at protecting personal data used to train AIs is "federated learning," which the *French Data Protection Authority (CNIL)* defines as follows: *"Federated learning is a learning paradigm in which multiple entities collaboratively train an AI model without pooling their respective data. In practice, the entities involved in the learning process send the models trained on their local data to an orchestrator center to consolidate the global model. This*

[119] These AI Models Are Pretty Mid. That's Why Companies Love Them. Companies are looking for simpler and cheaper ways to deploy artificial intelligence (Isabelle Bousquette - Wall Street Journal, Jul 17, 2024) https://www.wsj.com/articles/these-ai-models-are-pretty-mid-thats-why-companies-love-them-710a0f72?mod=tech_lead_story

[120] OpenAI Slashes the Cost of Using Its AI With a "Mini" Model (Wired Jul 18, 2024) https://www.wired.com/story/openai-gpt-4o-mini/

[121] Paris-based Zama raises €67 million Series A to build privacy-preserving blockchain and AI (EU-Startups.com Mar 7, 2024) https://www.eu-startups.com/2024/03/paris-based-zama-raises-e67-million-series-a-to-build-privacy-preserving-blockchain-and-ai/

paradigm contrasts with centralized learning, in which all data is sent to a central server responsible for training the model."[122]

In the field of sensitive data, enhanced protection of health data could also involve the development of AI technologies on federated networks between local healthcare stakeholders across Europe. These initiatives should enable the scientific and economic valorization of health data by academic institutions or businesses within the French and European healthcare sectors.[123]

A foundational principle of the Internet could also be applied to the development of AI in the networked era: the principle of neutrality. According to this principle, the network behaves neutrally with respect to the data it transmits, and the network's intelligence is located at its edges (a concept also known as "end-to-end" architecture). The principle of net neutrality, coined by Tim Wu in 2003[124] and advocated by the EU during the United Nations' summit on Internet governance in 2005,[125] could also be applied to AI in the context of computing power. As *The Economist* succinctly puts it: *"The internet got better and faster by moving data closer to users. Now the same must happen with computing power."*[126] Distributed architectures may thus become the norm for AI design. In addition to the large language models promoted by cloud providers as part of their data

[122] Federated learning (CNIL 2024)
https://www.cnil.fr/fr/definition/apprentissage-federe

[123] This is the case for *ADLIN Science*, a start-up spun off from École Polytechnique and based at the *Genopole* in Evry, as well as the *Alliance Santé IA* project led by the University Hospital of Montpellier, both of which aim to develop new uses of AI in research and healthcare organizations.

[124] Tim Wu: Network Neutrality, Broadband Discrimination (Journal of Telecommunications and High Technology Law, Vol. 2, p. 141, Jun 2003)
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=388863

[125] Communication from the Commission to the Council, the European Parliament, the European Economic and Social committee and the Committee of the Regions - Towards a Global Partnership in the Information Society: the Contribution of the European Union to the Second Phase of the World Summit on the Information Society (WSIS)
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52005DC0234

[126] "The internet got better and faster by moving data closer to users. Now the same must happen with computing power" The Economist Technology Quaterly Jan 2024
https://www.economist.com/technology-quarterly/2024/01/29/the-internet-got-better-and-faster-by-moving-data-closer-to-users

processing and storage offerings, more "frugal" AI systems, requiring fewer computational resources and less data, could eventually be developed.

IX. AI AND INTELLECTUAL PROPERTY NEW LEGAL FRONTIERS

A. Licensing and Regulation of AI Code

Initially, AI libraries used for deep learning were predominantly published as open source, with many developed by European computer scientists (e.g., *Scikit-Learn*, *Keras*, and *SpaCy*). The rise of large language models *(LLMs)* has raised new questions for AI stakeholders regarding the economic implications of open formats for these models. Some companies have opted to maintain open-source licenses for their models, while others have imposed proprietary restrictions on their more advanced models. However, beyond the code itself, the data required to train these models presents additional challenges.

Lawrence Lessig, a world-renowned expert in intellectual property law and a long-time advocate of open-source software, takes a nuanced position on the openness of AI models. He notes:

> *AI is more a category than a technology. Like the category "weapon", it ranges from the relatively harmless to the potentially catastrophic. No one would believe that the access we allow to pea-shooters should be the same for stinger missiles. Neither should we believe that the software norms developed for operating systems or media players must apply in the same way to highly capable AI systems with the potential to cause immense harm. Private companies alone, in fierce competition with each other, do not have sufficient incentives to avoid catastrophic risk. Neither would simply banning open-source AI avoid the risk of great harm. Instead, we need to develop the regulatory capacity to ensure an environment within which safe AI can be*

> *developed, and the regulatory judgment to determine when the public risk from any AI deployment is too great. Today, these risks are imposed upon all of us by private actors with little public oversight. That formula has not worked with dangerous technologies in the past. It will not work with the AI systems of the future.*[127]

The analogy between AI systems and military technologies has led some experts to consider imposing restrictions on the architecture of the graphics processing units (*GPUs*) that power AI models. In addition to the export bans already enacted by the U.S. government, new forms of technological restrictions could also be proposed.

In his article for *Wired* magazine,[128] Will Knight explores the concept of limiting high-risk uses of AI by integrating control mechanisms directly into the *GPUs*. This approach would involve embedding rules within the architecture of these chips to limit the capabilities of AI algorithms, thereby preventing the covert development of potentially dangerous AI systems by hostile nations or unethical corporations. This proposal, put forward by the U.S. think tank *Center for a New American Security (CNAS),* is akin to the issuance of licenses by a regulatory body to control access to the most advanced computing powers required to train the most powerful AIs. This idea is reminiscent of the Clinton administration's failed attempt to introduce a backdoor into personal computer chips (known as *Clipper Chip*[129]*).* The proposal suggests that secure components, such as the cryptographic modules already integrated into certain *Nvidia* and *Intel* chips, could be used to restrict unauthorized access to specific AI models.

[127] Not all AI models should be freely available, argues a legal scholar (Lawrence Lessig in The Economist Jul 29, 2024) https://www.economist.com/by-invitation/2024/07/29/not-all-ai-models-should-be-freely-available-argues-a-legal-scholar

[128] Will Knight. Etching AI Controls Into Silicon Could Keep Doomsday at Bay (Wired Jan 25, 2024) https://www.wired.com/story/fast-forward-ai-silicon-doomsday/

[129] The Short Life and Humiliating Death of the Clipper Chip (Gizmodo Apr 7, 2023) https://gizmodo.com/life-and-death-of-clipper-chip-encryption-backdoors-att-1850177832

B. AI' s Impact on Authors and Creators

Generative AI, and especially the ease with which non-specialists can create content (text, images or videos) by writing scripts (or prompts), represent a major risk to authors and creators, and will require new regulatory measures in addition to existing legislation. According to Jean-Marie Cavada and Colette Bouckaert in the *Cahiers de la Documentation Française*:

> *The emergence of generative AI in 2022, a branch of AI focused on creating models capable of generating new data, has been a seismic event for all cultural sectors. These sectors saw their content massively and illegally exploited to feed the training datasets of this AI. This can be described as a shock, one that necessitates a transformation of the economy of the cultural industries.*[130]

Indeed, the speed with which generative AI has developed has taken lawmakers by surprise. The recent *AI Act*, for example, does not take copyright aspects into account, and yet this aspect is even more necessary today given that tech players view authors' claims as major obstacles to the development of business models for new generations of services. An illustration of this very relative concern for legality and respect for intellectual property was recently given by former *Google* CEO Eric Schmidt. At a recent conference at *Stanford*, he advised students to take the following approach to AI projects:

> *"Make me a copy of TikTok, steal all the users, steal all the music, put my preferences in it, produce this program in the next 30 seconds, release it and in one hour, if it's not viral, do something different along the same lines. So in the example that I gave of the TikTok competitor, and by the way, I was not arguing that you should illegally steal everybody's music. What you would do if you're a Silicon Valley entrepreneur, which hopefully all of you will be, is if it took off,*

[130] Jean-Marie Cavada and Colette Bouckaert: Intelligence articifielle et propriété intellectuelle: quels progrès ? (Cahiers français, September-October 2024 No. 441)

> *then you'd hire a whole bunch of lawyers to go clean the mess up, right? But if nobody uses your product, it doesn't matter that you stole all the content. And do not quote me…"* [sic][131]

The motto "*move fast and break things*" was coined by Mark Zuckerberg to explain to developers, but also to investors, the methods behind *Facebook' s* success.[132] Since then, this same motto has been adopted by many companies such as *Uber* and *Amazon.* These companies have brought in lobbyists alongside their many lawyers to influence regulations that could be detrimental to their interests. In the case of *Facebook*, *The Guardian* revealed in 2019 that these lobbying campaigns could even include intimidation or threats against European elected representatives.[133]

C. Generative AI: the New Legal Frontiers

From a legal standpoint, generative AI raises a number of new questions regarding the use of emerging forms of tech-based creation. Apart from concerns about the sources used to train these AI systems and the need for transparency and negotiations with rights holders, situations will arise in which the outputs of generative AI could constitute an infringement of intellectual property. Scarlett Johansson's lawsuit against OpenAI serves as a reminder that, beyond celebrity

[131] Ex-Google CEO Schmidt advised students to steal TikTok's IP and 'clean up the mess' later (Fortune Aug 16, 2024) https://fortune.com/2024/08/15/ex-google-ceo-eric-schmidt-stanford-ai-advice-steal-ip-hire-lawyers/

[133] Revealed: Facebook's global lobbying against data privacy laws (The Guardian, Mar 2, 2019) https://www.theguardian.com/technology/2019/mar/02/facebook-global-lobbying-campaign-against-data-privacy-laws-investment

cases, it will become increasingly difficult to resolve disputes in which a voice, appearance, or even body language has been created by generative AI.[134]

Another aspect of intellectual property issues arising from the development of generative AI concerns the field of inventions. In addition to its implications for artistic creation, many inventors could use generative AI to assist in designing or formalizing their work. According to *The Wall Street Journal*, one concern could be the economic consequences of developments in patent law, particularly if it becomes impossible to protect inventions partially created using AI: "The patent office threatens to deny any patent application in which it is determined that the inventor used AI, if it is found that the invention did not involve a 'significant contribution' by a 'natural person.' Here, too, the term 'significant' is used subjectively and is never clearly defined. Even if it were defined, it is questionable whether examiners could accurately assess the significance of the human contribution."[135]

Two years ago, the French newspaper *Le Monde* raised the question of whether the Nobel Prize could one day be awarded to an AI. AI is now deeply integrated into research across all scientific disciplines. While it may be premature to consider awarding the prize to an AGI, the 2024 Nobel Prize in Physics was awarded to J. Hopfield and G. Hinton, pioneers in the development of artificial neural networks, and the Nobel Prize in Chemistry was awarded to D. Hassabis and J. Jumper, the creators of *AlphaFold*, the AI tool developed by *Google DeepMind* that predicts protein conformation.[136]

[134] OpenAI v. Scarlett Johansson? Georgetown Law Professor Answers Legal Questions on AI-Generated Content (Georgetown University, Jun 4, 2024) https://www.georgetown.edu/news/ask-a-professor-openai-v-scarlett-johansson/

[135] New Patent Guidance on AI Could Quash Innovation (Wall Street Journal, Jul 11, 2024) https://www.wsj.com/articles/new-patent-guidance-on-ai-could-quash-innovation-dd848ea4

[136] AI comes to the Nobels: double win sparks debate about scientific fields (Nature, Oct 10, 2024) https://www.nature.com/articles/d41586-024-03310-8

X. THE ETHICAL CHALLENGES OF AI

A. What Human and Political Values for AI?

In the space of just a few years, the question of trust in AI has become a key issue for our societies. As these technologies become associated with essential functions across all human activities, it becomes more and more urgent to render them both intelligible and predictable. One of the first questions raised about how AIs work concerns their unpredictability, including for their designers, as AI academic Melanie Mitchell points out: *"Even the humans who train deep networks generally cannot look under the hood and provide explanations for the decisions their networks make. MIT Technology Review magazine called this impenetrability 'the dark secret at the heart of AI'. The fear is that if we don't understand how AI systems work, we can't really trust them or predict the circumstances under which they will make errors."*[137]

> ***"Even the humans who train deep networks generally cannot look under the hood and provide explanations for the decisions their networks make. MIT's Technology Review magazine called this impenetrability 'the dark secret at the heart of AI'...***
>
> **Melanie Mitchell**

The unpredictability of the results of AI models makes their use particularly tricky in situations where the safety or freedom of individuals is at stake. For Brian Christian, the need to integrate societal norms and values into AI is all the greater

[137] Melanie Mitchell: Artificial Intelligence: A Guide for Thinking Humans (Pelican Books 2019)

since "*machine-learning systems like this not only demonstrate bias but may silently, subtly perpetuate it*".[138]

B. The Difficulties of AI Alignment

The moral or social dilemmas facing users, and especially the designers of these AIs, are such that leading AI experts like Stuart Russell consider that the problem is connected with the expression of our goals when we question AIs.

> *Finding a solution to the AI control problem is an important task; it may be, in the words of philosopher Nick Bostrom, "the essential task of our age." Up to now, AI research has focused on systems that are better at making decisions, but this is not the same as making better decisions if human and machine objectives diverge. This problem requires a change in the definition of AI itself: from a field concerned with a unary notion of intelligence as the optimization of a given objective to a field concerned with a binary notion of machines that are provably beneficial for humans. Taking the problem seriously seems likely to yield new ways of thinking about AI, its purpose, and our relationship with it.*[139]

Users who assume that AI will incorporate moral considerations or even practices commonly accepted by humans are met with responses that do not address these concerns. This problem leads AIs to answer questions in unexpected ways, following a 'logic' that remains fundamentally alien to that of human beings. Stuart Russell used the following metaphor to describe the considerable gap that can exist between the goals formulated by humans and the results provided by AI:

> *We call this the King Midas problem. King Midas specified his objective: I want everything I touch turned to gold. He got exactly what he asked for.*

[138] Brian Christian. The Alignment Problem (W. W. Norton & Company 2020)

[139] Stuart Russell: If We Succeed in Daedalus "AI & Society" Volume 151, Number 2; Spring by the American Academy of Arts & Sciences 2022) https://www.amacad.org/sites/default/files/daedalus/downloads/Daedalus_Sp22_AI-and-Society.pdf

> *Unfortunately, that included his food and his drink and his family members, and he dies in misery and starvation. Many cultures have the same story. The genie grants you three wishes. Always the third wish is "please undo the first two wishes" because I ruined the world. And unfortunately, with systems that are more intelligent and therefore more powerful than we are, you don't necessarily get a second and third wish. So the problem comes from increasing capabilities, coupled with our inability to specify objectives completely and correctly. Can we restore our carbon dioxide to historical levels so that we get the climate back in balance? Sounds like a great objective. Well, the easiest way to do that is to get rid of all those things that are producing carbon dioxide, which happen to be humans. You want to cure cancer as quickly as possible. Sounds great, right? But the quickest way to do it is to run medical trials in parallel with millions of human subjects or billions of human subjects. So you give everyone cancer and then you see what treatments work.*[140]

The temptation to assume that an AI system, because it is based on data sourced from humans, will inherently integrate their moral or ethical concerns could have dangerous consequences for users and, more broadly, for societies. The goal of enabling AI to structure its responses based on moral or ethical considerations remains, in fact, one of the key unresolved dilemmas for AI designers.

C. Trust: A Strategic Goal for AI Designers

One of the strategic aspects for the development of AI specialists' activities will surround the social and cultural acceptability of these technologies. In this respect, user confidence will be a decisive factor of their industrial growth trajectory. In their 2014 book *"The Age of Context",* Scoble and Israel already emphasized the importance of trust for the development of online services: *"We believe the most*

[140] AI could be a disaster for humanity. A top computer scientist thinks he has the solution. Stuart Russell interview by Kelsey Piper Oct 26, 2019 https://www.vox.com/future-perfect/2019/10/26/20932289/ai-stuart-russell-human-compatible

trustworthy companies will thrive in the Age of Context, and those found to be short on candor will end up short on customers. Transparency and trustworthiness will be the differentiating factors by which customers will make an increasing number of choices."[141]

AI companies will need to rely on strict rules to ensure that these technologies do not pose social, cultural or security risks over the long term. These companies can only become architects of trust in the eyes of their users by demonstrating their commitment to these rules.

After his recommendations on the "qualified transparency" of algorithms outlined in *The Black Box Society*,[142] Frank Pasquale's book on the new laws of robotics sets out the principles that, following the example of the rules proposed by Isaac Asimov, should structure the functioning of AIs in our societies.

Extract from *New Laws of Robotics* by Frank Pasquale[143]

- "Robotic systems and AI should complement professionals, not replace them."
- "Robotic systems and AI should not counterfeit humanity."
- "Robotic systems and AI should not intensify zero-sum arms races."
- "Robotic systems and AI must always indicate the identity of their creator(s), controller(s), and owner(s)"

[141] R. Scoble & S. Israel; Age of Context: Mobile, Sensors, Data and the Future of Privacy (CreateSpace Independent Publishing Sep 2014)

[142] "Black Box Society - The Secret Algorithms That Control Money and Information" by Frank Pasquale (Harvard University Press 2015)

[143] Frank Pasquale. New Laws of Robotics (The Belknap Press of Harvard University Press 2020)

> Regulators will need to require responsibility-by-design (to complement extant models of *security by design* and *privacy by design*). That may involve requiring certain hard-coded audit logs, or licensing practices that explicitly contemplate problematic outcomes. Such initiatives will not simply regulate robotics and AI post hoc but will influence systems development by foreclosing some design options and encouraging others.

D. AI integrated into Social, Cultural, and Political Life

Thanks to the mass collection of supposedly 'harmless' data on individuals, it becomes possible to infer characteristics of a person that they themselves are unaware of or do not wish to be obvious. This makes it virtually impossible to separate 'sensitive' data from other personal data. Consumption patterns, for example, have been shown to relate to political choices: In 1999, at the time of the President Clinton impeachment crisis, advertising executive Mark DiMassimo determined that *"84% of Campbell Soup partisans favored Bill Clinton's impeachment."*[144]

AIs will also be increasingly involved in decisions crucial to people's lives, safety and even freedom. One of the most sensitive areas in the use of personal data regards crime statistics and their use to predict possible repeat offenses. Here again, 'toxic' feedback loops have already been described in the use of these AIs in judicial systems that have adopted these technologies:

[144] Darren Bridger and David Lewis: Soul of the New Consumer: Authenticity What We Buy and Why in the New Economy (Nicholas Brealey; Updated edition 2001)

> *This gap, between what we intend for our tool to measure and what the data actually captures, should worry conservatives and progressives alike. Criminals who successfully evade arrest get treated by the system as "low-risk"—prompting recommendations for the release of other similar criminals. And the overpoliced, and wrongfully convicted, become part of the alleged ground-truth profile of "high-risk" individuals—prompting the system to recommend detention for others like them. This is particularly worrisome in the context of predictive policing, where this training data is used to determine the very police activity that, in turn, generates arrest data—setting up a potential long-term feedback loop.*[145]

The potential negative consequences of AI on our institutions and societies are no longer a concern limited to techno-pessimists who fear that AI will surpass human capabilities and seek to dominate them. Daron Acemoglu, a Nobel Prize-winning economist and professor at the Massachusetts Institute of Technology[146], has even called for the establishment of a *precautionary principle* to guide the development and integration of AI into our societies.

> *Although I don't believe superintelligence and evil AI pose major threats, I often think about how the current risks might be perceived looking back 50 years from now. The risk that our children or grandchildren in 2074 accuse us of moving too slowly in 2024 at the expense of growth seems far lower than the risk that we end up moving too quickly and destroy institutions, democracy, and beyond in the process. So, the costs of the mistakes that we risk making are much more asymmetric on the downside. That's why it's important to resist the hype and take a somewhat cautious approach, which may include better regulatory tools, as AI technologies continue to evolve.*[147]

[145] Brian Christian. The Alignment Problem, Machine Learning and Human Values (p. 76) (W. W. Norton & Company 2020)

[146] Three Win Nobel in Economics for Research on Global Inequality (New York Times, Oct 14, 2024) https://www.nytimes.com/2024/10/14/business/nobel-economics.html

[147] Interview with Daron Acemoglu Professor of Economics at Massachusetts Institute of Technology in Gen AI: Too Much Spend, Too Little Benefit? (Goldman Sachs Global Macro Research Jun 25, 2024) https://www.goldmansachs.com/intelligence/pages/gs-research/gen-ai-too-much-spend-too-little-benefit/report.pdf

Another platform regulation tool concerns the analysis of the algorithms that process personal data. Transparency with regard to the '*Code*' of these algorithms could soon become mandatory for democratic societies. Regulation could apply in particular to the design of algorithms for connected devices that will impact people's safety. This is the case for the AIs behind the functioning of autonomous vehicles, for example. The *Media Lab* of the *MIT (Massachusetts Institute of Technology)* set up its *Moral Machine* project to analyze users' ethical and moral choices in the event of accidents involving driverless cars. The project's international survey, conducted on people from 233 different countries and territories, showed not only the points of consensus but also the differences in users' ethical choices according to their country or culture of origin:[148] "*The researchers found that countries' preferences differ widely, but they also correlate highly with culture and economics. For example, participants from collectivist cultures like China and Japan are less likely to spare the young over the old—perhaps, the researchers hypothesized, because of a greater emphasis on respecting the elderly.*"[149]

The prevailing choices of AIs with consequences for people's lives and safety must be the subject of democratic debate. Indeed, if these choices were imposed on us by non-European players, they could have social or political consequences that could run counter to the principles and values of European citizens. Transparency should therefore apply not only to the code of these AIs but also to their training methods to ensure the 'alignment' of these applications with the choices democratically approved by European citizens.

[148] The Moral Machine Experiment (E. Awad, S. Dsouza, R. Kim, J. Schulz, J. Henrich, A. Shariff, J-F. Bonnefon & I. Rahwan - Nature, Oct 24, 2018) www.americaninno.com/wp-content/uploads/2017/05/The-MM-Experiment.pdf

[149] Should a self-driving car kill the baby or the grandma? Depends on where you're from (MIT Technology Review, Oct 24, 2018)
https://www.technologyreview.com/2018/10/24/139313/a-global-ethics-study-aims-to-help-ai-solve-the-self-driving-trolley-problem/

XI. GEOPOLITICAL CHALLENGES OF AI

Beyond the economic, social, and cultural consequences of AI, these technologies now represent a significant geopolitical issue tied to the economic and military rivalry between the United States and China. In the era of hybrid warfare, AI could play a pivotal role in redefining the global balance of power. The ease with which AI technologies can be harnessed for decision-making or military action makes them a key geopolitical concern, particularly in the context of the ongoing rivalry between the U.S. and China. AI technologies now facilitate real-time analysis of conflict zones, missile guidance, coordination of drone swarms, and the development of fully autonomous 'smart' weapons.[150] The dual-use (civil and military) nature of AI technologies could have far-reaching implications for the balance of power among nations in the years ahead, prompting analysts at the Rand Corporation to assert: *"Although technology has often influenced geopolitics, the prospect of AI means that the technology itself could become a geopolitical actor."*[151]

A. From Wars *for* AI… to Wars *via* AI

In addition to their contribution to new forms of economic dominance, artificial intelligence technologies are already having a significant impact on weapons design and the conduct of military operations, including more immersive simulations, optimized radar and sonar data processing, and cyber-defense

[150] Jonah M. Kessel, Natalie Reneau and Melissa Chan, "A.I. Is Making It Easier to Kill (You). Here's How", The New York Times, Dec 13, 2019; https://www.nytimes.com/video/technology/100000006082083/lethal-autonomous-weapons.html.

[151] AI and Geopolitics: How Might AI Affect the Rise and Fall of Nations? (Barry Pavel, Ivana Ke, Michael Spirtas, James Ryseff, Lea Sabbag, Gregory Smith, Keller Scholl, Domenique Lumpkin (Rand Corporation, Nov 3, 2023)
https://www.rand.org/pubs/perspectives/PEA3034-1.html

coordination. The war in Ukraine has demonstrated that these technologies have become indispensable to the execution of operations in armed conflict.[152] Another geopolitical dimension of AI pertains to technologies that can be used to reinforce ideological and political control over populations. From the early detection of dissent on social media to the real-time monitoring and control of individuals' activities via facial recognition, AI has become a critical tool for ensuring the durability of authoritarian regimes. Simultaneously, AI has facilitated the 'democratization' of large-scale disinformation campaigns and interference operations aimed at destabilizing democracies. These technologies enable fully automated disinformation and propaganda campaigns that can affect millions of individuals,[153] often without them realizing that the messages they receive are AI-generated. As Henry Kissinger, Eric Schmidt, and Daniel Huttenlocher observed in their book *The Age of AI*, open democracies—whose economic and political systems are heavily reliant on digital technologies—are prime targets for hostile foreign actors or terrorist organizations.[154]

> ***“Although technology has often influenced geopolitics, the prospect of AI means that the technology itself could become a geopolitical actor...***
>
> ***Barry Pavel***

[152] Jeremy Wagstaff, “New Model Army”, International Monetary Fund, December 2023; https://www.imf.org/en/Publications/fandd/issues/2023/12/Case-Studies-New-model-army-Jeremy-Wagstaff.

[153] M. J. Banias, “Inside Countercloud: A Fully Autonomous AI Disinformation System”, The Debrief, Aug 16, 2023; thedebrief.org/countercloud-ai-disinformation.

[154] Henry Kissinger, Eric Schmidt and Daniel Huttenlocher, The Age of AI: And Our Human Future (Little, Brown and Company 2021)

B. Slowing Down but Not Choking Chinese AI Technologies

Artificial intelligence has become a significant factor in the geopolitical rivalry between the U.S. and China, with some commentators referring to this conflict as "Cold War 2.0." However, beyond the economic tensions, there is growing concern about the potential for direct military confrontation, particularly over Taiwan, which could serve as the flashpoint. Taiwan is home to the vast majority of high-performance chips essential for running AI systems, primarily produced by *Taiwan Semiconductor Manufacturing Company (TSMC)*.

For the U.S. administration, the concern about Chinese supremacy in AI is twofold: first, the fear that AI technologies could enable the Chinese military to catch up with American military capabilities; and second, the worry that AI will help China accelerate its economic development to the point of surpassing the U.S.. In response, the U.S. initially focused its efforts in October 2022 on restricting American investments in Chinese AI companies.[155] According to the *Center for Strategic and International Studies (CSIS),* the Biden administration essentially communicated to China: *"If your policy is military-civil fusion, then the only realistic way of implementing our policy of no military end use is to end all sales to China, and we are now willing to take that step."*[156]

As a result, the U.S. has effectively imposed a *de facto* embargo on China regarding critical technologies for AI design and development, including high-performance microprocessors, supercomputers, and AI models. In retaliation, China imposed export restrictions on certain critical metals necessary for the production of technological components. The U.S. embargo on advanced chips represents a

[155] "Biden bans range of US high tech investments in China citing national security risk", The Guardian, Aug 9, 2023; https://www.theguardian.com/world/2023/aug/09/biden-executive-order-us-investment-chinese-technology.

[156] Gregory C. Allen, "Choking off China's Access to the Future of AI", Center for Strategic and International Studies, 11 Oct 2022
www.csis.org/analysis/choking-chinas-access-future-ai.

strategic challenge to AI development in China and underscores the genuine fear of Chinese dominance in AI technologies.

American efforts to prevent China from acquiring the most advanced chips are viewed by Chinese authorities as unacceptable attacks on the country's economic progress. In response to the risks of economic stagnation in AI, China's ambassador to the United States, Xie Feng, framed the situation in sporting terms: *"This is like restricting the other side to wear outdated swimwear in a swimming contest while you yourself are wearing a speedo..."*[157]

It is worth noting that even before the inauguration of the elected president, Donald Trump, the Biden administration imposed a new set of sanctions against China on December 2, 2024.[158] These sanctions aim to restrict Chinese manufacturers' access to technologies essential for the development of new artificial intelligence models, which could also have military applications.[159] The measures affect both access to microprocessors and the software and hardware technologies involved in the design of these chips. In response, China announced new export restrictions on minerals critical for chip manufacturing and military equipment, such as gallium, germanium, and antimony.[160] The rise to power of Donald Trump could once again be accompanied by additional sanctions targeting Chinese "dual-use" technologies, which are used for both civilian and military purposes.

[157] "China warns of retaliation to US curbs on investment and chips" (Financial Times, Jul 19, 2023) www.ft.com/content/ad1350b9-0e4b-40a4-bb70-4c236e513e7a.

[158] Commerce Strengthens Export Controls to Restrict China's Capability to Produce Advanced Semiconductors for Military Applications (Bureau of Industry and Security, 2 Dec 2024) https://www.bis.gov/press-release/commerce-strengthens-export-controls-restrict-chinas-capability-produce-advanced

[159] Chip war ramps up with new US semiconductor restrictions on China (The Guardian 3 Dec 2024) https://www.theguardian.com/us-news/2024/dec/03/joe-biden-china-microchip-export-restrictions-law-changes

[160] China Bans Rare Mineral Exports to the U.S. (New York Times 3 Dec 2024) https://www.nytimes.com/2024/12/03/world/asia/china-minerals-semiconductors.html

C. AI: Increasingly Energy-Intensive Yet Crucial for Sustainable Development

AI technologies are particularly resource-intensive in terms of computing power (both during the training phases of AI models and when executing queries), and they also place significant demands on costly raw materials to produce high-performance graphics processing units (GPUs).[161] As a result, securing the energy required to run AI systems has become a strategic priority for major technology companies. Accordingly, *Microsoft, Google,* and *Amazon* are now investing in nuclear energy sectors to power their AI infrastructures.[162] Their objectives are twofold: to control energy costs and to reduce the carbon footprint of their data centers. However, the development of new generations of "small modular reactors" (SMRs) requires substantial investments, and these reactors are unlikely to be operational within the next decade.[163]

Paradoxically, these AI technologies are also a key component of the 'low carbon' transition for other industrial sectors. For instance, AI is expected to contribute to the detection of deposits of critical raw materials needed for renewable energy production. As Jef Caers, Professor of Earth Sciences at Stanford University, points out, AI could also play a pivotal role in optimizing the energy efficiency of processing these metals.[164] Further evidence of the versatility of AI models is found in the fact that AI technologies initially developed for military applications, particularly in counterterrorism efforts, have been repurposed for civilian use, enabling the detection of rare metal deposits.

[161] AI Power Consumption: Rapidly Becoming Mission-Critical (Forbes, Jun 20, 2024) https://www.forbes.com/sites/bethkindig/2024/06/20/ai-power-consumption-rapidly-becoming-mission-critical/

[162] Nuclear-Powered AI: Big Tech's Bold Solution or a Pipedream? (Wall Street Journal, Oct 22, 2024) https://www.wsj.com/business/energy-oil/nuclear-power-artificial-intelligence-tech-bb673012

[163] Google Backs New Nuclear Plants to Power AI (Wall Street Journal, Oct 14, 2024) https://www.wsj.com/business/energy-oil/google-nuclear-power-artificial-intelligence-87966624

[164] "Can AI Help Us Go Green?" Interview with Jef Caers, Techsequences, Jul 19, 2023; www.techsequences.org/podcasts/2023/07/can-ai-help-us-go-green.

The increasing dependence of global economies on critical raw materials has led the Secretary-General of the *Organization for Economic Cooperation and Development (OECD),* Mathias Cormann, to warn about the ecological consequences of restricting these exports: *"The challenge of achieving net zero CO2 emissions will require a significant scaling up of production and international trade in critical raw materials. Policy makers must closely scrutinize how the concentration of production and trade coupled with the increasing use of export restrictions are affecting international markets for critical raw materials. We must ensure that materials shortfalls do not prevent us from meeting our climate change commitments."*[165]

D. "Friendshoring": the End of Globalization?

The embargo measures imposed by the United States are not universally supported regarding their long-term effects. *ASML* CEO Peter Wennink is reported to have stated that cutting off China would essentially drive innovation within the country, allowing China to become highly competitive at the expense of others. Speaking to the media, he said, *"There are 1.4 billion Chinese, many of them smart. They come up with solutions that we have not yet thought of. You force them to become very innovative."*[166] This risk is also acknowledged by Ben Buchanan, White House Special Advisor on AI:

> *Controls on chips would be risky. There is a possibility that strong export controls might provide an advantage in the short term but lead to downsides in the long term. Cutting off China's access to chips would create an immediate incentive for China to invest even more in building a domestic industry. Given that Chinese buyers would in such circumstances have no access to foreign*

[165] "Raw materials critical for the green transition", OECD, Apr 11, 2023; www.oecd.org/fr/presse/approvisionnements-en-matieres-premieres-critiques-les-risques-pour-la-transition-verte.htm.

[166] Rajeswari Pillai Rajagopalan, "AI Chips for China Face Additional US Restrictions", The Diplomat, Apr 5, 2024; https://thediplomat.com/2024/04/ai-chips-for-china-face-additional-us-restrictions/.

> *chips, the industry that China has been trying (and largely failing) to build just might finally flourish. American and allied chip companies would also see massive hits to their revenues, potentially constraining them from investing in the R&D that maintains their technological edge.*[167]

It should be noted that the European Union and the United States do not share the same strategic objectives vis-à-vis China. For the U.S., the primary objective is to limit the ambitions of the Chinese authorities in the military and economic spheres, ambitions that are increasingly supported by advances in AI. For EU countries, the focus is more on rebalancing trade with China, as trade practices have often been viewed as unfair in the past. According to Matthew Eitel of the Center for European Policy Analysis, *"Although the European Union is exploring ways to strengthen its export controls regime, its emphasis on "de-risking" supply chains from China is increasingly focused on China's "unfair" trade practices, not on limiting China's technological advancement."*[168]

Another consequence of the war in Ukraine is that economic relations with countries now considered new geopolitical risks for the U.S. are being reevaluated. This shift in geopolitical realities has led to profound changes in the rules that have governed global trade since the collapse of the Soviet bloc. These developments are now challenging the trend toward economic globalization that began after World War II with the creation of the *General Agreement on Tariffs and Trade (GATT)* in 1947, which later evolved into the *World Trade Organization (WTO)* in 1995. U.S. Treasury Secretary Janet Yellen refers to this reconfiguration of global trade as "*friend-shoring*" (or *"ally-shoring"),* as opposed to *"offshoring"*: *"[...] friend-shoring means—and you've seen this in action—that we have a group of countries that have strong adherence to a set of norms and values about how to [...] run*

[167] Buchanan, Ben; Imbrie, Andrew. The New Fire: War, Peace, and Democracy in the Age of AI (p. 246). MIT Press 2022

[168] US China Tech Controls Face Problematic Diagnosis; Matthew Eitel (Center for European Policy Analysis May 13, 2024)
https://cepa.org/article/us-china-tech-controls-face-problematic-diagnosis/

the global economic system, and we need to deepen our ties with those partners and to work together to make sure that we can supply our needs of critical materials."[169]

Another consequence of the war in Ukraine is that economic trade with countries that now constitute new geopolitical risks for the U.S. is being called into question. This new geopolitical reality has resulted in profound changes to the rules that have prevailed since the collapse of the Soviet bloc. These factors are now calling into question the trend towards economic globalization that began in the post-war period with the creation of the *General Agreement on Tariffs and Trade (GATT)* in 1947, which in 1995 went on to become the *World Trade Organization (WTO)* .U.S. Treasury Secretary Janet Yellen refers to this reconfiguration of world trade as "*friend-shoring*" (or *"ally-shoring"),* as opposed to *"offshoring"*: *"[...] friend-shoring means—and you've seen this in action—that we have a group of countries that have strong adherence to a set of norms and values about how to [...] run the global economic system, and we need to deepen our ties with those partners and to work together to make sure that we can supply our needs of critical materials."*[170]

The re-emergence of two antagonistic economic blocs is raising concerns among economists, who warn of a potential Sino-American confrontation that senior Chinese officials are already referring to as the "silicon curtain." As Solveig Godeluck reminds us in *Les Échos*, *"The division of the commercial world into two blocs—friends and non-friends—would come at a significant economic cost. According to the Center for Strategic and International Studies (CSIS), international trade would be penalized 'to the extent that non-aligned developing countries will be excluded from the friend-shoring orbit.'" The World Trade Organization (WTO) has also estimated that trade between blocs could reduce global GDP by approximately 5% in the long term.*[171]

[169] "US Treasury Secretary Janet Yellen on the next steps for Russia sanctions and "friend-shoring" supply chains", Atlantic Council, Apr 13, 2022; www.atlanticcouncil.org/news/transcripts/transcript-us-treasury-secretary-janet-yellen-on-the-next-steps-for-russia-sanctions-and-friend-shoring-supply-chains.

[170] "US Treasury Secretary Janet Yellen on the next steps for Russia sanctions and "friend-shoring" supply chains", Atlantic Council, Apr 13, 2022; www.atlanticcouncil.org/news/transcripts/transcript-us-treasury-secretary-janet-yellen-on-the-next-steps-for-russia-sanctions-and-friend-shoring-supply-chains.

[171] Solveig Godeluck, "Washington is promoting a new world trade order with "friend-shoring", Les Échos, Feb 7, 2023; www.lesechos.fr/monde/etats-unis/washington-promeut-un-nouvel-ordre-commercial-mondial-avec-le-friend-shoring-1904348.

E. A Clash between EU, U.S., and China Technological and Legal Models

According to Anu Bradford, Professor of international law at Columbia University, "some commentators describe Europe as becoming a *"casualty in the U.S.–China tech war" or "a colony caught between the U.S. and China,"*[172] *with "less bargaining power to determine its own digital fate," and thus forced to "make a choice" between the US and China."*[173] But another war has begun at the same time: that of the 'horizontal' influence of the technological models promoted by China and the U.S.. The confrontation between these two countries could ultimately hinder the development of AI in many countries. Democratic nations therefore prefer regulation based on European frameworks to the adoption of technologies under intense state control by China or the major American platforms, which are fraught with the potential for abuse. Most democratic countries have witnessed the failure of self-regulation by the major technological platforms, which have been embroiled in numerous scandals that have tarnished their reputation in the eyes of democracies. For instance, the uncontrolled use of personal data in the *Cambridge Analytica* scandal, the use of these platforms by extremist groups for insurrectionary purposes during the storming of the U.S *Capitol* on January 6, 2021, and the accusations of complicity in ethnic cleansing in Myanmar brought against *Facebook* have profoundly altered political leaders' perceptions of these platforms.

The European influence, which Anu Bradford terms "*the Brussels Effect*", reflects the impact of European tech regulations beyond the EU's borders. European standards have a recognized leverage effect on the digital laws developed elsewhere in the world. In her book *Digital Empires,* Anu Bradford describes the obligation, including for the U.S., to draw inspiration from European forms of

[172] On this point, see the report by Catherine Morin-Desailly, "L'Union européenne, colonie du monde numérique ?"submitted to the Senate on March 20, 2013; https://www.senat.fr/rap/r12-443/r12-443.html.
[173] Anu Bradford, Digital Empires The Global Battle to Regulate Technology, Oxford University Press, 2023.

regulation, particularly for AI, rather than continuing to take the risks associated with the lack of regulation for major platforms. European legislation appears to be more protective of freedoms than regulations drafted in other parts of the world. This was especially evident for the *General Data Protection Regulation (GDPR),* the principles of which have been adopted in more than 17 countries and territories worldwide: from California to Brazil, India and even China. China's *Personal Information Protection Law (PIPL),* enacted in August 2021, applies to individuals, companies and even Chinese government departments. However, there are notable exceptions if the constraints associated with the processing of Chinese citizens' personal data *"is necessary to fulfill statutory duties".*[174]

Conversely, the political leaders of authoritarian regimes are more likely to adopt Chinese technologies and regulations, giving them ever-tighter control over their populations. By leveraging on Chinese technologies, such as the *Social Credit System*, authoritarian regimes can benefit from the world's largest ongoing social engineering experiment. The *Social Credit System* uses AI technologies to perform facial recognition, analyze individual activities and monitor individuals by rating their economic, social and political behavior within Chinese society. The system thus exploits all the data collected on people's activities and could potentially expand to include genetic profiles in the future.[175]

F. AIs Even More Suited to Mass Surveillance

New generations of connected devices, such as connected cars, smart city sensors, and health-monitoring technologies, are further diversifying the channels through which personal information can be 'extracted'. However, these technological advancements do more than simply increase the amount of data collected on

[174] Jamie P. Horsley, "How will China's privacy law apply to the Chinese state?" (New America, Jan 26, 2021) www.newamerica.org/cybersecurity-initiative/digichina/blog/how-will-chinas-privacy-law-apply-to-the-chinese-state.

[175] Sui-Lee Wee, "China is Collecting DNA From Tens of Millions of Men and Boys, Using U.S. Equipment" (The New York Times, Jun 17, 2020) https://www.nytimes.com/2020/06/17/world/asia/China-DNA-surveillance.html.

individuals; they have fundamentally altered the nature of surveillance practices. These tools have introduced shorter cycles between the collection of information and its use for interference, manipulation, or modification of opinions and behavior. Beyond combating criminal activity, terrorism, or espionage, it is now possible to influence individuals' behavior and even shape their beliefs without their awareness. These hyper-targeted political messages were used by Cambridge Analytica, in collaboration with Russian intelligence agencies,[176] to sway U.S. voters during the 2016 presidential election and to influence British citizens during the Brexit campaign.

Artificial intelligence has enabled surveillance technologies to extend beyond the scope of traditional systems, which required significant human and logistical resources to collect and analyze data on entire populations. For authoritarian governments and dictatorships, these technologies have become essential for ensuring the continuity of their regimes and, in some cases, extending their influence to unprecedented levels. Michael Kanaan, head of artificial intelligence at the U.S. Department of Defense, commented on the concern raised by the U.S. military following Vladimir Putin's 2017 speech in which he stated, *"Artificial intelligence is the future, not only for Russia, but for all humankind. It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world.*"[177] For the U.S. military, this geopolitical warning was interpreted as a direct challenge to all technology players: *"Now, for all to hear, Putin had just declared everything at stake. Without any room for misunderstanding, he equated AI superiority to global supremacy, to a strength akin to economic or even nuclear domination. He said it for public consumption, but it was rife with political purpose. Whoever becomes the leader in this sphere will become the ruler of the world."*[178]

[176] Whistleblower: Cambridge Analytica shared data with Russia (Euractiv, May 17, 2018) https://www.euractiv.com/section/global-europe/news/whistleblower-cambridge-analytica-shared-data-with-russia/

[177] Putin says the nation that leads in AI 'will be the ruler of the world' (The Verge, Sep 4, 2017.) https://www.theverge.com/2017/9/4/16251226/russia-ai-putin-rule-the-world

[178] Michael Kanaan T-Minus AI: Humanity's Countdown to Artificial Intelligence and the New Pursuit of Global Power p17 (BenBella Books 2020)

In addition to traditional military technologies, Putin's remarks also referenced the potential for political and ideological control that AI technologies now enable. Ben Buchanan describes a "Sputnik moment" for China when, in 2015, DeepMind's AlphaGo defeated the world's top Chinese Go players. He draws a parallel between the U.S. space race that followed the Soviet Union's 1957 launch of Sputnik and China's response to AlphaGo's victory.[179] He argues that AlphaGo had the same effect on China that Sputnik had on the United States: it marked AI as a new domain of geopolitical competition, with profound implications for both warfare and diplomacy.

> *AlphaGo, the argument went, had done for China what Sputnik did for the United States. It established AI as a new terrain of geopolitical competition, one with obvious implications for war and peace. It accelerated Chinese government investment in AI to the tune of billions of dollars each year and drove more students into the relevant fields in math and science.*[180]

Another example illustrates the strategic significance of AI technologies for China. In 2021, a large-scale cyberattack on Microsoft Exchange servers was attributed to China.[181] Unlike traditional cyberattacks, which typically aim to gather intelligence, steal intellectual property, or obtain state secrets, the goal of this hack was to use the data from hundreds of thousands of Western companies to train Chinese AI models. This was a 'win-win' situation for China, as it not only gained deeper insights into Western individuals and companies but also enhanced its AI systems. These AI models, in turn, will enable more sophisticated data analysis, providing the Chinese government with powerful tools for population control.

[179] China's 'New Generation Artificial Intelligence Development Plan' (2017) (translation by DigiChina - Stanford University, Aug 1, 2017) https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/

[180] Buchanan, Ben; Imbrie, Andrew. The New Fire: War, Peace, and Democracy in the Age of AI (p. 51) (MIT Press 2022)

[181] China's Microsoft Hack May Have Had A Bigger Purpose Than Just Spying (NPR Aug 26, 2021) https://www.npr.org/2021/08/26/1013501080/chinas-microsoft-hack-may-have-had-a-bigger-purpose-than-just-spying

Looking ahead, economic, political, and military objectives could become increasingly intertwined within government AI strategies. This is especially true for China, which views AI as a central tool in its broader strategy to achieve global political and economic leadership

Mastering AI: a threefold strategic objective for China

1. Economic Development: To advance the growth of its technological sector and, more broadly, its entire economy, particularly in the context of its ongoing competition with the United States for global leadership.
2. Ideological and Political Control: To maintain ideological and political dominance over its population, ensuring the stability and continuity of its regime.
3. Military Superiority: To develop next-generation smart weapons systems, the strategic importance of which has been highlighted by recent conflicts, particularly the war in Ukraine.

G. AI at the Center of the China-Taiwan Conflict

An example of the growing importance of AI technologies in military operations is provided by George Takach in his book *Cold War 2.0*, which explores new forms of AI-based conflict. Takach describes how the initial moments of a potential Chinese war against Taiwan might unfold and highlights the strategic significance of AI technologies in such a conflict:

> *Chances are, on that fateful day when China launches the opening shower of missiles and drones on Taiwan, the AI will quickly be set to automatic mode. Why? For the simple reason that there will be insufficient time to permit Taiwanese and American military personnel to effectively insert themselves into target selection and approval activities. Human eyes and ears cannot meaningfully absorb sufficient data points in this battle scenario quickly enough, and the human brain cannot digest and process information fast enough to know rationally what to do in the case of hundreds of missiles and drones coming at them, most of which will hit their intended targets a few minutes after launch. Ironically, the previous handicap of the military leader, the so-called fog of war (where the leader knew too little of what was going on in the battle space), has been replaced in this Taiwan attack scenario by the equally problematic "overwhelming clarity of war" (the leader has all the data points, just doesn't know what to do with them all in a timely manner). The only chance of success here for the democracies is that the AI underpinning all the different defensive ADS weapons does its job spectacularly well, autonomously.*[182]

Recent conflicts, particularly in Ukraine, have demonstrated that AI's use on the battlefield offers crucial advantages. These technologies could thus play a decisive role in future conflicts. In this context, the training of AI specialists and the development of industrial sectors related to AI are becoming critical strategic concerns for all nations. To support his argument, Takach draws on Winston Churchill's metaphor regarding the *Royal Air Force*[183], replacing Spitfire pilots with AI technology specialists as the figures to whom democratic nations owe their freedom.

> *Victory in what history will call the Battle of the Taiwan Strait will be determined by the degree of each side's mastery over four critical technologies: artificial intelligence, semiconductor chips, quantum computers, and biotechnology. In turn, the success (or failure) of these four technologies will*

[182] George S. Takach. Cold War 2.0_ Artificial Intelligence in the New Battle between China, Russia, and America by George S. Takach (p. 15) (Pegasus Books 2024)

[183] 'Never in the field of human conflict was so much owed by so many to so few' (UK Parliament, Aug 20, 1940) https://www.parliament.uk/about/living-heritage/transformingsociety/private-lives/yourcountry/collections/churchillexhibition/churchill-the-orator/human-conflict/

> *depend on two factors: first, the quality of the science, technology, engineering, and mathematics (STEM) graduates that have been coming out of master's and PhD programs in the autocracies and the democracies, respectively, over the twenty-five years prior to the battle, and second, how well the military-industrial complex in both camps has integrated these innovations into their respective weapons systems and defense ecosystems. Fusing civilian and military innovation is very hard work. Never before will so much depend on so few college grads.*[184]

In addition to China's longstanding claims over its 'rebel province', Taiwan is also home to the world's leading high-level chip manufacturer, *TSMC (Taiwan Semiconductor Manufacturing Company)*. Despite recent efforts to replicate *TSMC's* production in factories in the U.S. and Europe, manufacturing in Taiwan remains essential, particularly for the chips required to power AI systems. A successful Chinese assault on Taiwan would thus represent a triple victory for the Chinese regime:

- Political victory: The forced reunification, long championed by the regime, would strengthen its narrative of nationalism and the centrality of Chinese military power.
- Economic victory: China would gain the opportunity to exploit the technological restrictions imposed by the United States, slowing down Western tech companies while positioning China to accelerate its global leadership in AI technologies.
- Geostrategic victory: The takeover would validate the Chinese authorities' broader strategy of regional control, including strategic shipping lanes and the expansion of their influence over the islands of the South China Sea.

[184] George S. Takach. Cold War 2.0_ Artificial Intelligence in the New Battle between China, Russia, and America (p. 17). (Pegasus Books 2024)

XII. AI INDUSTRIAL POLICY AS AN INSTRUMENT TO PROTECT HUMAN RIGHTS

A. Towards a 'Third Way' for European AIs

In its effort to regulate high-risk practices within the AI sector, the *Artificial Intelligence Act (AI Act)* aims to prevent a drift toward authoritarian control over populations, such as the regime established by China's *Social Credit System*. The *AI Act* also promotes the development of ethical competencies within companies and public administrations, as well as fostering economic activities linked to the creation of ethical AI technologies in the long term. According to Mark Coeckelbergh, a philosopher of technology specializing in AI ethics, the introduction of ethical regulations from the very design stage of AI models could provide a competitive edge for European technologies.

> *That being said, generally speaking AI ethics is not necessarily about banning things (Boddington 2017). Another barrier to getting AI ethics to work in practice is that many actors in the AI field such as companies and technical researchers still think of ethics as a constraint, as something negative. This idea is not totally misguided: often ethics has to constrain, has to limit, has to say that something is unacceptable. And if we take AI ethics seriously and implement its recommendations, we might face some trade-offs, in particular in the short term. Ethics may have a cost: in terms of money, time, and energy. However, by reducing risks, ethics and responsible innovation support the long-term, sustainable development of businesses and of society. It is still a challenge to convince all the actors in the AI field, including policymakers, that this is indeed the case. Note also that policy and regulation are not only about banning*

> *things or making things more difficult; they can also be supportive, offering incentives, for example.*[185]

> **“ *“There is not a single key technology behind the iPhone that has not been state funded. This includes the wireless networks, the Internet, GPS, a touch-screen display, and … the voice-activated personal assistant Siri...***
>
> **Mariana Mazzucato**

Another key aspect of industrial policy in the technological sector will involve strengthening the links between civil and military research. Indeed, many foundational technologies behind the Internet were originally developed within the military sector. As economist Mariana Mazzucato points out: *“There is not a single key technology behind the iPhone that has not been state funded. This includes the wireless networks, “the Internet, GPS, a touch-screen display, and… the voice-activated personal assistant Siri.”*[186]

Looking ahead, EU industrial policies may also be shaped by the perceived need to bolster political and military power in the face of potentially hostile blocs, much as the U.S. responded to Vladimir Putin's warning regarding the geopolitical significance of AI.

B. EU Responses to Techno (or Data) Colonialism

> *When the British, Russians and Japanese made their bids for hegemony in the nineteenth and twentieth centuries, they relied on steamships, locomotives and machine guns. In the 21st century, to dominate a colony, you no longer need to send in the gunboats. You need to take out the data.*

[185] Mark Coeckelbergh. AI Ethics (pp. 174-176) (MIT Press 2020)

[186] Mariana Mazzucato, quoted by Amanda Schaffer, “Tech’s Enduring Great-Man Myth”, MIT Technology Review, Aug 4, 2015; www.technologyreview.com/2015/08/04/166593/techs-enduring-great-man-myth.

> *A few corporations or governments harvesting the world's data could transform the rest of the globe into data colonies – territories they control not with overt military force but with information.*
>
> Yuval Noah Harari (2024)[187]

Faced with the reality of Europe potentially becoming a "techno-colony"[188] of the U.S. and China, proposed industrial policy responses are beginning to take shape. Europe's dependency on American tech giants, particularly in cloud computing, is a key concern. *Amazon (AWS), Microsoft Azure,* and *Google Cloud* together account for nearly 75% of the European cloud market, making the development of independent AI strategies for European states even more urgent.

In his report on European competitiveness, Mario Draghi, former President of the European Central Bank (ECB), offered a bleak assessment of the EU's technological dependency and lag: *"We rely on a handful of suppliers for critical raw materials and import over 80% of our digital technology. We are severely lagging behind in new technologies: only four of the world's top 50 tech companies are European."*[189]

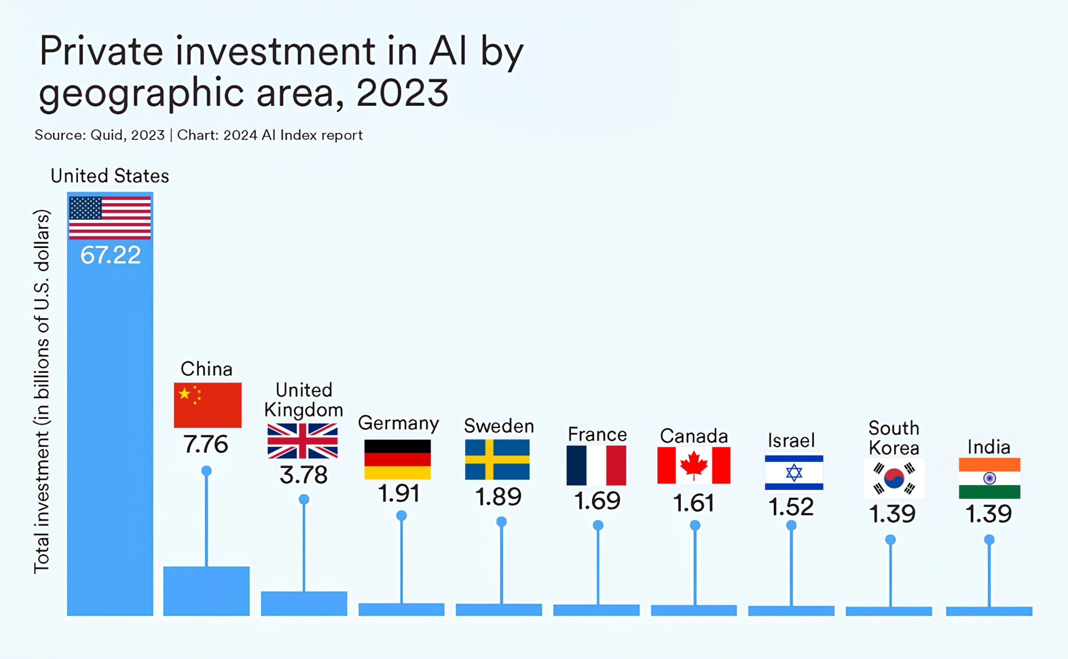
Draghi emphasized the urgency for the EU to collectively invest in strategic technologies, particularly AI, to avoid remaining on the defensive against both the U.S. and China. To achieve this, he advocates for joint action on an unprecedented scale. Until recently, the idea of common debt, often referred to as "*Eurobonds*," faced systematic opposition from many European partners, particularly Germany. However, the crises triggered by the *COVID-19* pandemic and, more recently, the war in Ukraine, have shifted this stance. In 2020, the EU launched its recovery

[187] Harari, Yuval Noah. Nexus (p. 370). Vintage Publishing 2024

[188] The Rise of Techno-Colonialism (Hermann Hauser and Hazem Danny Nakib, Project Syndicate Aug 7, 2024)
https://www.project-syndicate.org/commentary/techno-colonialism-defines-us-china-rivalry-by-hermann-hauser-and-hazem-danny-nakib-2024-08

[189] Address by Mr. Draghi – Presentation of the report on the Future of European competitiveness – European Parliament – Strasbourg – Sep 17, 2024
https://commission.europa.eu/document/download/fcbc7ada-213b-4679-83f7-69a4c2127a25_en

plan, amounting to an initial €750 billion.[190] Originally intended to address the pandemic (hence the term "*Coronabonds*"), the plan also aimed to finance the energy and digital transitions across EU countries.



Private investment in Al by geographic area, 2023 - AI Index
Source: HAI Human-Centered Artificial Intelligence (Stanford University)[191]

Building on these exceptional debt measures, Draghi's report calls for long-term European investment to support the development of strategic technologies, with AI as the foremost priority. The investment gap between the U.S., China, and the EU in AI continues to grow. Since 2013, the U.S. has invested $335.2 billion in private AI funding, followed by China with $103.7 billion, and the UK at $22.3 billion. In 2023 alone, American private investment in AI reached

[190] Recovery plan for Europe European Commission (2020) https://commission.europa.eu/strategy-and-policy/recovery-plan-europe_en#the-largest-stimulus-package-ever
[191] https://hai.stanford.edu/news/ai-index-state-ai-13-charts

$67.22 billion, while Germany and France invested respectively $1.91 billion and $1.69 billion.

According to Mario Draghi, for the EU to catch up with the U.S. and China, it needs to make “massive and unprecedented” investments of between €750 and €800 billion per year, or approximately 5% of its GDP.

To achieve this, new financing methods must be employed to support technologies of strategic importance for the EU. Furthermore, the only way to convince all EU countries to commit to such joint investment will be through robust political advocacy and campaigns.[192]

However, beyond European control over technologies vital for economic growth, the protection of Europe’s social and democratic model is also at stake. Legal measures alone regulating AI technologies will not be sufficient to prevent the misuse of technologies that are increasingly pervasive across social, economic, and political life.

In addition to regulating existing AI platforms, European industrial actors must be able to develop technologies that uphold the principles and values of European citizens. This would involve creating alternatives to current services that rely heavily on the mass extraction of user data for advertising purposes.

For Europe to contribute fully to the development of sustainable AI technologies that protect freedoms, it must implement an industrial policy of unprecedented scale. Beyond regulatory texts, only a coordinated investment policy will allow Europe to develop future generations of ethical AI technologies. This will require efforts far beyond those made during the *COVID-19* pandemic. The urgency is heightened by the geopolitical crises already unfolding in Europe, as well as the

[192] For the former ECB president, Europe needs to make €800 billion of additional investments per year in the future – the equivalent of 5% of the European Union’s (EU) GDP – or around three times the Marshall Plan (between 1% and 2% of GDP in annual investments in the post-war period).
Thomas Piketti: “Europe must invest: Draghi is right” (Le Monde Sep 17, 2024)
https://www.lemonde.fr/blog/piketty/2024/09/17/europe-must-invest-draghi-is-right/

potential for other global conflicts. These factors make industrial proactivity, particularly in AI, even more essential.

Without European alternatives for strategic technologies, the EU's regulatory efforts on AI will be insufficient to control the political and social risks associated with non-European platforms. Therefore, the defense of human rights must go hand in hand with an industrial policy that enables the design and development of AIs that integrate European principles and values from the outset. In addition to the introduction of ethical guidelines for technologies, an industrial policy for the AI boom must also be developed. Too often, the EU focuses on the risks of these technologies without providing the political or economic means to develop alternatives to non-European technologies and services.

Thierry Breton, former European Commissioner for Internal Market, has stated that it is time for Europe to end the naivety that has characterized its approach to industrial policy in the technology sector. This naivety primarily concerns the rebalancing of industrial and commercial trade with the U.S. and China, whose markets remain largely closed to European companies. Breton's remarks, made in 2020, are particularly relevant today considering the re-election of Donald Trump, who plans to retaliate against EU industrialists with new trade barriers. Trump has even compared the EU to a "***mini-China.***"[193]

For Breton, another dimension of European naivety lies in the political destabilization and disinformation campaigns carried out by hostile entities and nations within Europe. He remarked, ***"The era of a conciliatory or naïve Europe that solely relies on the virtue of its soft power is behind us. We are now witnessing the dawn of a Europe determined to defend its strategic interests."***[194]

[193] Trump's trade tariffs: how protectionist US policies will hit German carmakers (The Guardian, Nov 9, 2024) https://www.theguardian.com/business/2024/nov/09/trumps-trade-tariffs-how-protectionist-us-policies-will-hit-german-carmakers

[194] Thierry Breton, 'Europe: The End of 'Naïvety" (LinkedIn Sep 10, 2020) https://www.linkedin.com/pulse/europe-end-naïvety-thierry-breton

Numerous examples have demonstrated the current difficulty, if not the impossibility, for the European Union to compel large platforms such as *X* (formerly *Twitter*), *Facebook*, or especially *TikTok* to comply with European law. Whether it concerns Elon Musk's interference in European electoral processes[195], the lack of algorithmic transparency on *TikTok*, or issues related to the absence of personal data protection in the case of Facebook with the Cambridge Analytica scandal or more recently Mark Zuckerberg's changing stance on *Facebook's* content regulation following the election of Donald Trump, these cases highlight ongoing challenges. In the past, antitrust sanctions imposed by the EU did not result in changes to their industrial or ethical behavior.

This same issue arises when examining the fiscal practices of these companies, their respect for labor rights, and the protection of users' personal data. In the absence of European tech players of international stature, especially in the AI field, European regulatory measures will continue to face political and industrial obstacles. The supremacy of the three major American cloud computing players—*Amazon, Microsoft, and Google*—is particularly evident in Europe. Given their widespread use by Europe's largest companies, addressing these issues becomes even more complex, as demonstrated by the case of *TikTok* in the U.S.

In April 2024, the U.S. Congress passed a law aimed at forcing *ByteDance*, the parent company of *TikTok*, to sell its American branch to a U.S. company or face a ban.[196] This attempt at a ban was subject to political arbitration at the highest levels of the U.S. government. Thus, Donald Trump, who was in favor of banning *TikTok* in 2020, confirmed that he would oppose such a ban upon his inauguration on January 20, 2025. In doing so, he echoed the arguments of leaders of the

[195] EU Commission urged to act over Elon Musk's 'interference' in elections (The Guardian 8 Jan 2025) https://www.theguardian.com/world/2025/jan/08/eu-commission-urged-to-act-elon-musk-interference-elections

[196] Support for a U.S. TikTok ban continues to decline, and half of adults doubt it will happen (Pew Research Center - Sep 5, 2024) https://www.pewresearch.org/short-reads/2024/09/05/support-for-a-us-tiktok-ban-continues-to-decline-and-half-of-adults-doubt-it-will-happen/

prominent Chinese tech platform, citing the *"staggering impact on the free speech of U.S. users..."* [sic].[197]

C. A Synergy between Tech Policy and Tech Regulation

In the antitrust field, the EU has historically prioritized the fight against intra-European monopolies, particularly the risks of price increases for consumers. This approach has, however, relegated concerns about the stifling of innovation due to monopolistic practices to the background. Nevertheless, the EU may soon need to reconsider its stance to better support European tech players in the face of competition from the U.S. and China.

In this context, the EU will once again have to make high-level decisions about balancing the risks posed by non-European monopolies and those associated with the emergence of European industrial giants. This decision hinges on whether European companies, which are more likely to comply with regulatory measures, should be favored over American or Chinese tech giants. The latter not only enjoy technological supremacy but also present significant risks in terms of political abuse, violations of freedoms, and political interference. Europe's experience has shown that these non-European companies not only resist regulatory measures but also find ways to circumvent them, maintaining their dominance despite sanctions imposed under EU antitrust laws.

To safeguard the rights of European citizens, European industrial policies and competition regulations must be coordinated and work in synergy to help develop European tech players of international stature that adhere to EU principles and values.

[197] TikTok says US ban would have 'staggering' impact on free speech (BBC Sep 24, 2024) https://www.bbc.com/news/articles/c5y3y79llndo

Ben Buchanan, White House Special Advisor on AI, highlights the crucial role of American public authorities in fostering AI industrial development. This includes facilitating small and medium-sized enterprises (*SMEs*) access to public procurement through dedicated administrations (such as the Small Business Administration), providing favorable access to the immense computing power required to operate AI technologies, and supporting the aggregation of vast data resources that companies can leverage to train AI systems.

> *Democratic governments should make computing power more accessible to people with big ideas. Governments could buy cloud computing credits from leading companies at a large scale and at low rates. They could then grant access to those computing resources for academic labs or startups with broad social benefits, much as the National Science Foundation and Small Business Administration offer monetary grants to researchers and entrepreneurs. In addition to increasing innovation, making computational resources more accessible is essential for the age of AI, especially since many of the privacy-preserving or data-independent algorithms described above require substantial computational power to work.*[198]

To uphold European values and principles, EU has established a corpus of regulatory texts that govern the activity of major Internet platforms: *GDPR (General Data Protection Regulation)*, *DSA (Digital Services Act) DMA (Digital Markets Act)* and *DGA (Data Governance Act)*. The purpose of all these texts is to limit the possible abuses linked to the activity of these platforms (personal data protection, fight against abuse of a dominant position, moderation of harmful content, interoperability and algorithmic transparency). These texts represent pioneering initiatives and have been emulated in many other countries beyond the EU's borders, more recently the *AI Act*, constitute the latest responses to the risk of abuse by tech platforms.

However, the absence of alternative European solutions for citizens and companies hinders the implementation of stronger policy measures. As

[198] Buchanan, Ben; Imbrie, Andrew. The New Fire: War, Peace, and Democracy in the Age of AI (p. 239). MIT Press 2022

demonstrated by the difficulty in establishing a ban on *TikTok* in the U.S. or Europe, despite the risks it represents to democratic processes in terms of manipulation or interference.

Barack Obama summed up the perception of a solely defensive Europe buttressed by regulations in the face of American tech players in 2015: *"We have owned the internet. Our companies have created it, expanded it, perfected it in ways that they can't compete. And oftentimes what is portrayed as high-minded positions on issues sometimes is just designed to carve out some of their commercial interests."*[199]

Similarly, Jack Ma, CEO of China's heavyweight *Alibaba*, gave his view at *VivaTech Paris* in 2019, expounding not without irony his doubts regarding the rules put in place by Europeans in the digital field: *"I worry about the worries of Europe. You have such perfect rules and laws. Then everything they do it's "Let's follow the rules and laws." When they worry, they make rules and laws. When we have problems [in China], we start to solve the problems."*[200]

Just months later, Jack Ma went as far as to publicly criticize Xi Jinping's finance policy, at a time when he was the richest man in China and his company was preparing the largest IPO of all time.[201] Ironically, it was the absence of rule of law in China (and particularly the absence of judicial independence) that resulted in his being placed under house arrest and subsequently ousted from *Alibaba*, the company he founded, then forced into exile after undergoing ideological re-education to advocate the Chinese regime...

It is worth noting that each time the Chinese regime has been faced with a choice between the economic development of Chinese tech companies and compliance with the regime's ideological line, it has always come down on the side of ideological orthodoxy. To the point of even throttling the entire tech sector in

[199] Obama accuses EU of attacking American tech companies because it 'can't compete' (The Verge Feb 15, 2015)
https://www.theverge.com/2015/2/17/8050691/obama-our-companies-created-the-internet

[200] Jack Ma at VivaTech, Paris May 23, 2019
https://x.com/AlibabaGroup/status/1131673464773914624

[201] Jack Ma's Costliest Business Lesson: China Has Only One Leader (Wall Street Journal, Aug 20, 2021)
https://www.wsj.com/articles/jack-mas-costliest-business-lesson-china-has-only-one-leader-11629473637

China, according to the *Financial Times*.[202] As the *Wall Street Journal* underscores, the development of Chinese AI is now required to mold to the regime's political and ideological discourse.

> *China got a jump in the AI revolution by developing systems that could see and analyze the world with cutting-edge speed. The area of AI known as computer vision, which enables tracking and surveillance, aligns with Chinese leader Xi Jinping's emphasis on political control. Despite that early success, the country was caught flat-footed by the public debut of OpenAI's ChatGPT in late 2022 and the generative AI craze it unleashed. Generative AI's large language models, which are used to produce content at speed, can be difficult to predict and are much more likely to undermine that control.*[203]

This political alignment with the *Chinese Communist Party*'s ideological line is non-negotiable for the regime: *"These large models need to implement core socialist values," Arcesati said. "There's this challenge of political alignment that generative AI developers need to come to terms with."*[204]

D. New Industrial Policy Strategies for AI

At the outbreak of the *COVID-19* pandemic, even the highly influential *MIT Technology Review* acknowledged Silicon Valley's failure to create useful innovations for citizens. A failing partly due to these companies' advertising-based economic model which aims to collect (or extract) as much data as possible on individuals, as quickly as possible to qualify advertising targets using micro-targeting technologies often associated with AI:

[202] How China has 'throttled' its private sector (Financial Times Sep 12, 2024) https://www.ft.com/content/1e9e7544-974c-4662-a901-d30c4ab56eb7

[203] China Puts Power of State Behind AI — and Risks Strangling It. Liza Lin (Wall Street Journal, Jul 16, 2024) https://www.wsj.com/tech/china-puts-power-of-state-behind-aiand-risks-strangling-it-f045e11d

[204] Chat Xi PT? China's Chatbot Makes Sure It's a Good Comrade (Wall Street Journal, May 24, 2024) https://www.wsj.com/tech/ai/chat-xi-pt-chinas-chatbot-makes-sure-its-a-good-comrade-bdcf575c

> *The pandemic has made clear this festering problem: the US is no longer very good at coming up with new ideas and technologies relevant to our most basic needs. We're great at devising shiny, mainly software-driven bling that makes our lives more convenient in many ways. But we're far less accomplished at reinventing health care, rethinking education, making food production and distribution more efficient, and, in general, turning our technical know-how loose on the largest sectors of the economy.*[205]

In addition to existing financing mechanisms for technologies, such as subsidies awarded through technology-focused calls for proposals, innovative financing models are needed to support the development of technological innovations. These new mechanisms should be specifically designed to foster the advancement of public-interest technologies in sectors that are not adequately covered by current funding options or that face challenges due to misaligned timelines. Furthermore, there is a pressing need to introduce new support measures for AI technologies with the greatest potential for social impact. According to the recommendations of *Rand Corporation* experts:

> *Data and computing power are widely available to companies large and small, and no single entity can reliably predict from where the next revolutionary AI advance might originate. Consequently, governments should consider expanding their toolboxes beyond traditional regulatory techniques. Two creative mechanisms could be for governments to invest in establishing robust, publicly owned data sets for AI research or issue challenge grants that encourage socially beneficial uses for AI.*[206]

Private initiatives have also been devised to develop financing for tech ventures, such as the *XPRIZE Foundation*, a non-profit organization founded in 1995 which

[205] Covid-19 has blown apart the myth of Silicon Valley innovation (MIT Tech Review 25 Apr 2020) www.technologyreview.com/2020/04/25/1000563/covid-19-has-killed-the-myth-of-silicon-valley-innovation/

[206] AI and Geopolitics: How Might AI Affect the Rise and Fall of Nations? (Barry Pavel, Ivana Ke, Michael Spirtas, James Ryseff, Lea Sabbag, Gregory Smith, Keller Scholl, Domenique Lumpkin (Rand Corporation, Nov 3, 2023)
https://www.rand.org/pubs/perspectives/PEA3034-1.html

was used to launch the *IBM Watson AI Xprize* challenge dedicated to improving human-AI collaboration.[207] To further support the development of AI technologies in strategic sectors, Europe should consider establishing finance platforms dedicated to addressing specific technological challenges. One such example is the Challenge.gov platform, created in the United States in 2010 under the Obama Administration. Unlike thematic calls for proposals, such as those in the *Horizon Europe* program, *Challenge.gov* aimed to fund technologies that meet public interest criteria, guided by strict standards established by expert panels.[208]

The 3 Strategic Goals of an AI Industrial Policy in Europe

- Reducing dependence on non-European technologies by strengthening the European AI ecosystem, while enforcing strict regulations on competition, transparency, and platform interoperability.

- Promoting the development of ethical AI technologies that avoid replicating autocratic or repressive tendencies seen in non-European systems, and countering harmful economic models based on large-scale personal data extraction.

- Advancing the European AI ecosystem, particularly for AI technologies crucial to democratic processes and the management of sensitive, strategic data. This includes supporting European AI companies by improving their access to public procurement opportunities.

[207] https://www.xprize.org/prizes/artificial-intelligence

[208] Challenge.gov: Two Years and 200 Prizes Later (Obama White House Archive Sep 5, 2012) obamawhitehouse.archives.gov/blog/2012/09/05/challengegov-two-years-and-200-prizes-later

XIII. RECOMMENDATIONS

Recommendation 1

Create a large-scale joint funding initiative dedicated to AI technologies across the European Union. Through a common investment procedure, this initiative will aim to support the development of businesses, key sectors, and AI technologies in Europe. One component of this initiative will focus on funding the infrastructure and technologies necessary for the design and operation of AI systems in Europe.

Recommendation 2

Appoint a State Technology Coordinator. Establish the role of State Technology Coordinator at both the national and European levels, like the position of Chief Technology Officer (CTO) within the White House. The role of this coordinator will be to promote and implement French and European strategies related to technologies developed or used by public administrations, and to raise awareness among public procurement officials about implementing public contracts in the field of AI technologies. The coordinator will also oversee a forward-thinking task force focused on the impact of AI on Europe's strategic autonomy, as well as the consequences of AI on employment and employability by professional sector.

Recommendation 3

Create a Small Business Act and a Small Business Administration in France and Europe. Establish a Small Business Act in France and Europe to reserve a significant share of public procurement for innovative SMEs in the field of AI, enabling these businesses to grow through secure, solvable contracts.

Recommendation 4

Create a European Buy Act for EU tech companies, particularly for strategic AI technologies and systems for managing sensitive data. At the same time, there should be mechanisms to block potentially hostile takeovers in the AI sector, especially for technologies that ensure the EU's strategic independence.

Recommendation 5

Develop national and European programs to provide computing power and data storage at reduced costs for research organizations and innovative SMEs in the field of AI. These programs will also aim to promote the development of innovative architectures necessary for the design and operation of AI systems from European companies and institutions.

Recommendation 6

Include AI technologies in the EU's "*carbon tax*." Incorporate AI technologies into the planned 2026 revision of the EU's *"Carbon Border Adjustment Mechanism" (CBAM),* also known as the carbon tax. This should include industrial sectors related to AI technologies, to encourage the transition of EU partner countries towards low-carbon AI technologies, considering the energy and raw materials needed for their operation (such as microprocessors and cloud infrastructures).

Recommendation 7

Develop AI ethics education within schools and universities, as well as through ongoing training for public sector employees. Create awareness programs on the ethical aspects of AI for students and learners. Establish competitions to develop public interest projects focused on combating disinformation and addressing the political misuse of AI.

Recommendation 8

Develop training and support programs on the state's expertise in AI for its traditional roles, particularly for its sovereign functions, as well as on the impact of AI on democracy and the economy.

Recommendation 9

Promote the development of decentralized “resident” AI architectures and enhance European research programs focused on the decentralized operation of European AI systems. Concurrently, develop European R&D programs on federated learning for AI and homomorphic encryption for data security used by AI systems operating on remote structures. Establish a public program for the development of technologies necessary for the long-term storage and processing of public data.

Recommendation 10

Design innovative funding mechanisms for public interest AI projects in Europe. Develop innovative funding models in France and Europe that bring together public and private stakeholders to advance AI for public good in areas such as healthcare, energy management, and transportation. These mechanisms will be based on the principle of “technology challenges” and should accelerate the development of technologies that would not be available through the sole intervention of existing industrial actors.

XIV. PRESENTATION OF THE ISN AND IDFRIGHTS

This report was produced by the *Institute of Digital Sovereignty (ISN)* in partnership with *iDFRights*. It was coordinated by Bernard Benhamou, General Secretary of the ISN, with the help of Jean-Marie Cavada, President, and Colette Bouckaert, General Secretary of *iDFRights*.

Bernard Benhamou, General Secretary of the ISN, served as the interministerial delegate on Internet usage at the French Ministry of Research and Ministry of the Digital Economy. He coordinated the first European Union Ministerial Conference on the Internet of Things and Digital Inclusion during the French Presidency of the European Union in 2008, prior to which he was Advisor to the French Delegation at the United Nations World Summit on the Information Society. He currently teaches Internet governance at Paris 1 Panthéon-Sorbonne University after having introduced the first teaching programs on Internet and administrations at ENA and Sciences Po Paris.

INSTITUTE OF DIGITAL SOVEREIGNTY (ISN)

The Institute of Digital Sovereignty (ISN) is a non-profit association whose mission is to bring together digital and economic actors to create synergy in the challenges posed by European digital sovereignty. Since its foundation in 2015, the ISN has been committed to educating and mobilizing citizens and their representatives on digital sovereignty challenges. The ISN considers that our cyberspace should be protected in the same way as our land, sea and air spaces.

The ISN recommends technological, legal and political actions and measures enabling digital sovereignty to be asserted over all of our digital resources and in particular over our data. The ISN's work with European industrial representatives, institutions (local authorities, administrations, ministries, parliament) covers the traditional subjects of digital sovereignty as well as prospects for technological developments and their impact on society. Whether with regard to the protection of personal or industrial data, the sovereignty of critical digital infrastructure, or the social, political and cultural transformations brought about by technological advances. The ISN's goal is also to enable European tech players to better incorporate the notions of digital sovereignty within the development of their projects in order to generate a competitive advantage in terms of trust and security for their users. Lastly, the ISN seeks to contribute to the digital transformation of the French government to ensure protection of our sovereignty and to preserve our individual and collective freedoms at the same time.

iDFRights: Institute for Digital Fundamental Rights

The past 20 years has been a period of global expansion for digital technologies, which have taken on a central role in both our individual and collective lives, as well as within corporations. Through the innumerable data provided by individuals and economic actors, humans have built up a new stratum capable of steering our lives and organizations. This layer of data, often fed without us realizing it, has fueled the world's digital giants: predominantly American, followed by China, with Europe yet to carve out a seat at the table. Their business practices flout the competition laws in force in democratic countries, bringing condemnation from courts and parliaments. No reasonable person stands against digital technology and its opportunities for development, nor are they against artificial intelligence or quantum possibilities. But the future will depend on one drastic choice: do we want humanity, on the grounds of the boundless services digital technology offers, to continue to dominate the machines it has invented, or do we accept that we will become subjugated - in part or entirely - to the machines we have built? Do you want to become a slave to your smartphone, or be its master? Organizations are springing up throughout our democratic countries to protest against digital capitalism, universities are studying the question, lawsuits of all types are calling on courts to uphold our freedoms, parliaments are launching inquiries and seeking to lay down regulatory laws, and even governments are arming themselves to instill order and extricate digital technology from under American or Chinese control and bring it under the banner of our democracies, these days plagued by these unlawful conducts. That is why, after many years spent fighting to regulate this sector at the European Parliament (GDPR, "Copyright" Directive for authors, "Related rights" for press publishers), we decided to create this Institute. Alongside legal professionals, international universities, digital experts and moral civil society figures, we strive to educate on risks, draw up doctrines based on real-life cases, promote norms and standards to overcome the problems encountered by companies and organizations, and support legislators and national and EU leaders, with modernized legal materials.

XV. ACKNOWLEDGMENTS

The *Institute of Digital Sovereignty (ISN)* and Institute for Digital Fundamental *Rights (iDFRights)* would like to express their special thanks to the following people for their contribution to this report:

Catherine Morin-Desailly
Senator for Seine-Maritime, member of the European Affairs Committee. Chair of the Special Commission on the Bill to Secure and Regulate the Digital Space.

Philippe Latombe
Member of Parliament for Vendée. Secretary of the Constitutional Acts, Legislation and General Administration of the Republic.

Luc Julia
French-American engineer and computer scientist specializing in AI. Co-designer of the *Siri* voice assistant. Former Vice President of Innovation for *Samsung Electronics* and Chief Scientific Officer for *Renault* since 2021.

Michael Nelson
Director, Technology and International Affairs,
Carnegie Endowment for International Peace. Former Special Assistant for Information Technology to United States Vice President Al Gore

Desiree Milosevic
Co-Chair *RIPE* (Réseaux IP Européens Network Coordination Centre) Cooperation Working Group. Former special Advisor to the Chair Advisory Group *United Nations - Internet Governance Forum (IGF).*

Olena Kushakovska
President of *SAP Labs* France, Co-Director of the *Industrial Council of Artificial Intelligence Research (ICAIR).* Graduate of Kiev University in Applied Mathematics and *École Polytechnique.*

Laurence Devillers
Professor of AI and Ethics at Sorbonne University and a researcher at CNRS in the Interdisciplinary Laboratory of Digital Sciences (LISN) in Saclay. Her research focuses on human-machine interaction, emotion detection, spoken dialogue, and socio-affective robotics.

Eric Saund

Research scientist in Cognitive Science and AI
Formerly at the Xerox Palo Alto Research Center (PARC) and AAAS (American Association for the Advancement of Science) Congressional Science & Technology Policy Fellow. Currently Chief AI Officer at a Silicon Valley startup.

XVI. PARTNER ORGANIZATIONS

The Institute of Digital Sovereignty and *iDFRights* would like to thank the following organizations that made this report possible.

PLATINUM PARTNER

clever cloud

Since 2010, *Clever Cloud* provides a *Platform as a Service* service based in Europe. The PaaS helps development teams to put digital applications and services into production on a reliable infrastructure, with automatic scalability and transparent pricing. *Clever Cloud* is convinced that industrialized hosting will enable companies to work faster, be more agile in their markets, focus on their added value and stop worrying about their hosting technology. *Clever Cloud* is an outspoken advocate of data sovereignty, to ensure a free and autonomous digital space for all. *Clever Cloud* creates and contributes to many European Open Source projects and prefers to work with French and European technology partners. *Clever Cloud* is thus committed to an ecosystem of partners working every day to defend and improve our European digital sovereignty, notably through its commitment as chairman of the *Open Internet Project*.

SILVER PARTNERS

claranet
Make modern happen®

Founded in 1996, *Claranet* has 3,500 employees worldwide and operates managed services for over 10,000 customers in all business sectors. *Claranet* provides network, hosting and managed application services in the UK, France, Germany, the Netherlands, Portugal, Spain, Italy and Brazil. Today, *Claranet*'s customers include half of the companies in France's CAC 40 index. *Claranet* supports digital transformation with the advent of technologies and approaches that are bringing about far-reaching changes: Public Clouds, Big Data, DevOps, Agile, Machine Learning, Artificial Intelligence, Blockchain, micro-services, modern workplace. This transformation creates not only opportunities to seize (new economic models, new dematerialized uses, new markets, reduced time-to-market) but also brings with it complexity and threats (cyber-attacks, confidentiality of sensitive data, strengthened regulations, new entrants). More than just a service provider entrusted with a specific operational mission, *Claranet* is a partner that guides, advises and accompanies its customers with their strategy and the implementation of these transformations.

BeTomorrow

BeTomorrow is a full-service agency specializing in digital applications and transformation that has made products used by millions of users. For over 20 years, our cross-disciplinary teams have been shaping smart solutions with a thorough and pragmatic approach. From start-ups to big groups, *BeTomorrow* has assisted several hundred customers in the design, development and deployment of solutions: mobile, web, cloud/data apps & AI/IoT & 3D. From audit to market launch strategy, from design to development and from training to coaching, we adapt our support to the specific needs of our customers at each stage of their evolution. Our three major areas of expertise are *digital factory*, *digital transformation* and *digital innovation*. Our constant quest for technological expertise and innovation defines our identity, and our unique, agile operating method allows us to handle the most complex digital shifts.

PROMETHEUS-X

Prometheus-X is a nonprofit organization dedicated to the development of a sovereign infrastructure for data exchange in Europe, in line with the European data strategy. It brings together over 500 organizations across the EU for the creation of data spaces. Through the development of sovereign data spaces, *Prometheus-X* enables organizations and individuals to securely share and access data in an interoperable manner, fully compliant with European regulations, particularly in relation to data protection and digital sovereignty.

Prometheus-X approach is based on the creation of digital commons, which are open-source technological building blocks allowing any entity to create and operate a data space without being dependent on a single provider. These commons ensure an open and ethical infrastructure for data sharing across various sectors such as education, media, healthcare, law, and tourism. By consolidating a European ecosystem of sovereign, ethical, and secure data exchange, *Prometheus-X* aims to play a crucial role in ensuring Europe's technological independence.

QUICKTEXT

By harnessing AI and big data in the hotel sector, *Quicktext* provides a unique proposal to structure and share hotels' data, interact automatically with customers (thanks to its *Velma* virtual assistant), optimize sales, content, and marketing, and supply data for business intelligence. *Quicktext* is regularly ranked #1 in the field of AI in the hospitality industry by the Hotel Tech Report. *Quicktext*'s technological advance has been notably recognized at the Global Tech Award in Austin, TX (USA), Premium Travel Awards in Shenzen (China) and the Rencontres Internationales du Tourisme in Paris (France). *Quicktext* is used by the world's largest hotel groups: Accor (FRA), Emaar (UAE), Hyatt (USA), Highgate (USA), Warwick (USA), Eurostars (SPA), H-Hotels (GER), Jaz Hotels (EGY) Palladium (SPA), Precise (DEU), Santa Fe (MEX).

INSTITUTE OF
DIGITAL
SOVEREIGNTY

iDF RIGHTS
INSTITUTE

This report was produced by the Institute of Digital Sovereignty in partnership with iDFRights and coordinated by Bernard Benhamou, General Secretary of the ISN (bernard.benhamou@souverainetenumerique.fr). Unless otherwise mentioned, the translations contained in this report are those of the authors.