CÉDRIC VILLANI

Mathematician and
Member of the French Parliament

FOR A MEANINGFUL ARTIFICIAL INTELLIGENCE

TOWARDS A FRENCH AND EUROPEAN STRATEGY

Composition of the mission

Marc Schoenauer Principal Senior Researcher with INRIA • **Yann Bonnet** General secretary to the French Digital Council • **Charly Berthet** Head of legal and institutional affairs at the French Digital Council • **Anne-Charlotte Cornut** Rapporteur of the French Digital Council • **François Levin** Head of economic and social affairs at the French Digital Council • **Bertrand Rondepierre** Engineer for the French defense procurement agency

Mission assigned by the Prime Minister Édouard Philippe

A parliamentary mission from 8th September 2017 to 8th March 2018

Assisted by **Anne-Lise Meurier**, **Zineb Ghafoor**, **Candice Foehrenbach**, **Stella Biabiany-Rosier**, **Camille Hartmann**, **Judith Herzog**, **Marylou le Roy**, **Jan Krewer**, **Lofred Madzou** and **Ruben Narzul**.

Cédric Villani

Like so many teenage science lovers in the 1980s, I first discovered artificial intelligence—AI—by reading the captivating books by Douglas Hofstadter, who popularised science and portrayed Alan Turing with an enthusiasm that was irresistible.

However, when I began my career as a mathematician in the 1990s, like many of my peers, I deeply underestimated the impact of artificial intelligence, which yielded very few tangible results at the time. What a surprise it was to see the unbelievable progress achieved in the 2010s… Having decided to try my own hand at popularising scientific concepts for a general audience, I began to expound on AI frequently in my public lectures and in my discussions with the corporate world. And it was no less surprising for me to see my optimal transport research cited in recent articles about AI. It was almost as if I couldn't avoid coming across this multifaceted subject! As a matter of fact, over the past few years, no one has been able to avoid AI given its omnipresence in economic and social debate.

So I was not terribly surprised when the Prime Minister asked me to head up a task force on the artificial intelligence strategy for France and Europe. This was a challenging assignment, but my enthusiasm ran high. To lay out the initial guidelines, I benefited from the full support of Mounir Mahjoubi, the Minister of State with responsibility for Digital Affairs, and from the expertise of my colleagues specialised in AI, especially my former research associate Yann Ollivier. With their help and support from government institutions, I set up a "dream team" of seven highly competent individuals of diverse backgrounds, dedicated full time to the task force. This was a crucial stage because—as everyone knows—human resources are the first key to any project's success.

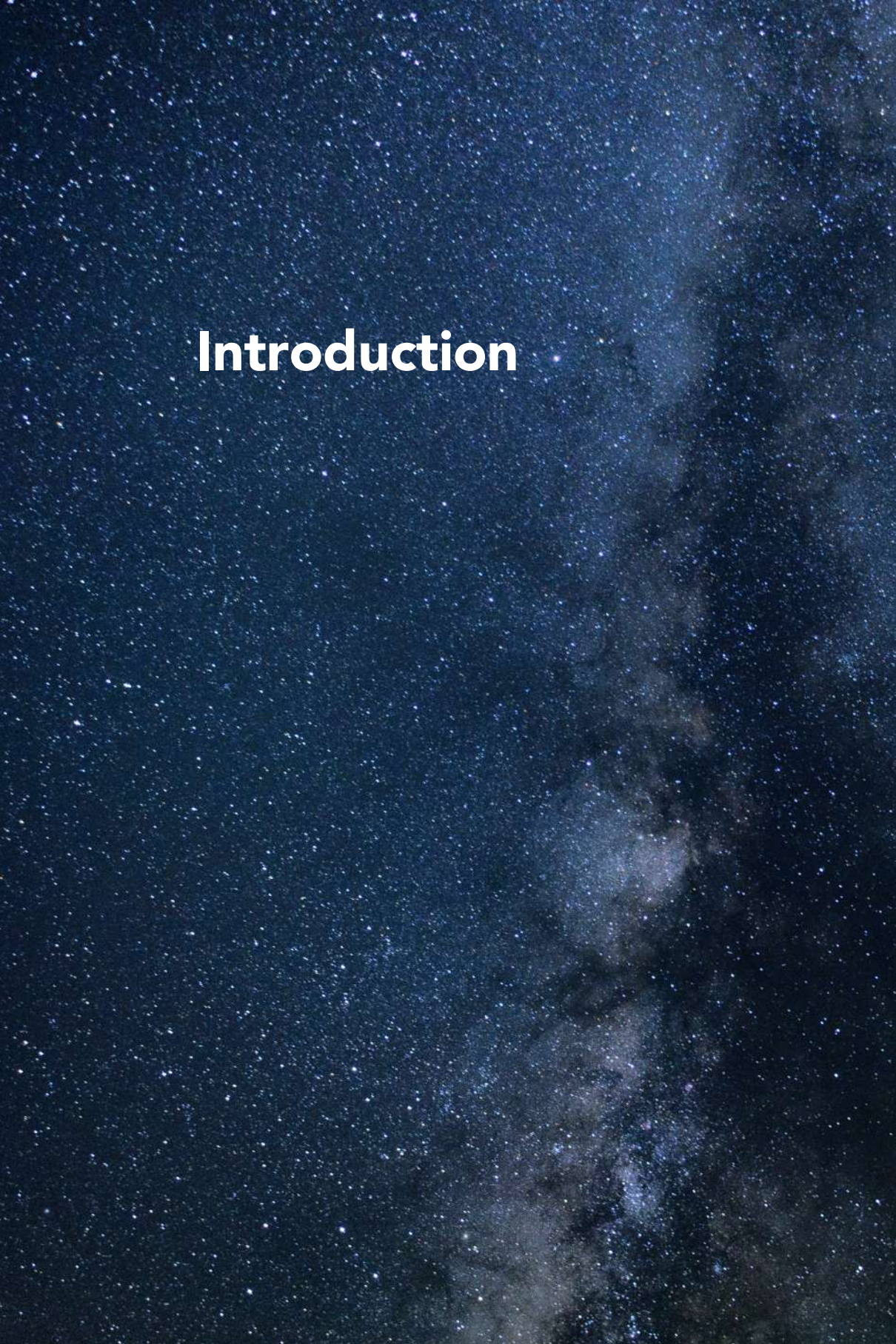
To kick-start the task force, we could rely on excellent sources, including the "France IA" report, spearheaded by Axelle Lemaire; the report by the Parliamentary Office

for the Evaluation of Scientific and Technological Choices (OPECST), sponsored by my fellow MPs Claude de Ganay and Dominique Gillot; not to mention the CNIL's outstanding works on the ethics of algorithms, along with the reports by the Employment Advisory Council (Conseil d'orientation pour l'emploi, COE). France Stratégie also provided input. The contributions grew in number, and soon there was a considerable amount of material to process! But working together, we were able to gather and summarise the masses of information provided by the hundreds of experts and thousands of members of the general public who contributed their ideas. I would like to extend my sincere thanks to Parlement & Citoyens, the non-partisan non-profit organisation that launched in record time the online platform to collect these contributions!

We cannot conceive AI in a purely national framework, so this task force was also an opportunity for a series of brief, intense visits to the stimulating places driving AI internationally: Palo Alto, Beijing, Berlin, Regensburg, London, Zürich, Bologna, Lisbon, Tel Aviv and Haifa. I would like to thank the many efficient institutional bodies involved in organising the logistics behind these visits. It goes without saying that we also visited the most stimulating AI sites in France, including The Camp, near Aix-en-Provence, which deserves special mention for hosting our task force for a few days.

This task force was a fascinating experience thanks to the wide variety of topics we studied. It was also a chance to work collaboratively for six whole months with all of society's stakeholders—from the hard sciences and humanities to government administrations, not to mention entrepreneurs, journalists and talented science fiction authors. Special thanks to Anne-Caroline Paucot and Olivier Paquet, two such writers who kindly agreed to let us include short stories in our report. After confronting these many viewpoints head on, we realised that AI is a universal subject. While it breaks down into countless variations, it must be tackled systemically. We are convinced that France—and Europe as a whole—must act synergistically, with confidence and determination, to become part of the emerging AI revolution.

Table of content



Introduction

Defining artificial intelligence is no easy matter. Since the mid-20th century when it was first recognized as a specific field of research, AI has always been envisioned as an evolving boundary, rather than a settled research field. Fundamentally, it refers to a programme whose ambitious objective is to understand and reproduce human cognition; creating cognitive processes comparable to those found in human beings.

Therefore, we are naturally dealing with a wide scope here, both in terms of the technical procedures that can be employed and the various disciplines that can be called upon: mathematics, information technology, cognitive sciences, etc. There is a great variety of approaches when it comes to AI: ontological, reinforcement learning, adversarial learning and neural networks, to name just a few. Most of them have been known for decades and many of the algorithms used today were developed in the '60s and '70s.

Since the 1956 Dartmouth conference, artificial intelligence has alternated between periods of great enthusiasm and disillusionment, impressive progress and frustrating failures. Yet, it has relentlessly pushed back the limits of what was only thought to be achievable by human beings. Along the way, AI research has achieved significant successes: outperforming human beings in complex games (chess, Go), understanding natural language, etc. It has also played a critical role in the history of mathematics and information technology. Consider how many softwares that we now take for granted once represented a major breakthrough in AI: chess game apps, online translation programmes, etc.

In recent years, AI has entered a new era, which gives rise to many hopes

Its visionary nature makes AI one of the most fascinating scientific endeavors of our time; and as such its development has always been accompanied by the wildest, most alarming and far-fetched fantasies that have deeply colored the general population's ideas about AI and the way researchers themselves relate to their own discipline. (Science) fiction, fantasy and mass projections have accompanied the development of artificial intelligence and sometimes influence its long-term objectives: evidence of this can be seen in the wealth of works of fiction on the subject, from 2001: A Space Odyssey to Her, Blade Runner and a significant proportion of literary science fiction. Finally, it is probably this relationship between fictional projections and scientific research which constitutes the essence of what is known as AI. Fantasies—often ethnocentric and based on underlying political ideologies—thus play a major role, albeit frequently disregarded, in the direction this discipline is evolving in.

In recent years, artificial intelligence has entered a new era, which gives rise to many hopes. Most notably, this has been tied to the recent success of machine learning. Thanks to complex algorithms, increased computing power and the exponential growth of human and machine-generated data, various applications have been developed in translation, transport (driverless cars), health (cancer detection), etc. It is worth noting that progress in AI is taking place in a technological context marked by the datafication of the world which affects all sectors of our society and economy, the development robotics and the blockchain (the distributed ledger technology which enables transactions between two, or more, agents without the presence of a trusted third party or institution which most notably underlines cryptocurrencies such

as bitcoin). The future of artificial intelligence surely depends on its exposure to these different technological developments.

These new applications fuel new narratives and fears based on, amongst other concepts, the omnipotence of artificial intelligence, the myth of Singularity and transhumanism. In recent years, these views have been largely endorsed and promoted by some of the most prominent actors in the AI landscape. Indeed, Silicon Valley is still the epicenter for the politics and economics of artificial intelligence, and it is held up as a model for anything that Europe regards as innovative. For many public and private stakeholders, it is more than a unique ecosystem; it is a mindset that must be adopted. California still dominates in word and in thought and encourages the concept of a single way, technological deterministic approach. If the development of artificial intelligence is fully shaped by private stakeholders, based abroad France and Europe will have no other choice than to their vision. This is already happening in the public sector. Think of the agreement signed between Microsoft and the Ministry of Education during the previous five-year term and the DGSI's[1] use of software provided by Palantir—a startup with links to the CIA. This is equally true in the private sector. Across Europe, businesses convinced that they have already lost the battle frequently succumb to the persuasive powers of the U. S tech giants, sometimes at the expense of our own digital "nuggets".

Unlike the fads of previous years regarding AI research, the subject now belongs not just to the scientific sphere but is on everyone's lips. Extraordinary amounts of money are invested in its research and industry, particularly in China. Politicians all over the world address it in their general statements of policy as a key means of leverage: Barack Obama's iconic interview with Wired in October 2016 illustrated how much he was aware that American progress in artificial intelligence could be a formidable tool for soft power. The Russian president, Vladimir Putin, himself asserted that "whoever became the leader in the field would rule the world", comparing artificial intelligence to nuclear technology. Even if he most likely felt the need to compensate for Russia's having lagged behind with artificial intelligence by making a powerful speech on the subject, his assertion reveals the geostrategic importance acquired by this technology. In the sense that value chains, particularly in the digital sector, are now global, countries that become leaders in the field of artificial intelligence will not only capture much of the value of the systems that they transform, but also control these same systems, calling into question the independence of other countries.

From now on, AI will play a much more important role than it has done so far

The point is that from now on, artificial intelligence will play a much more important role than it has done so far. It is no longer merely a research field confined to laboratories or to a specific application. It will become one of the keys to the future. Indeed, we are living in an ever more completely digital world. A world of data. This data is central to the functioning of artificial intelligence as we know it today. In a digital world, which is now our own, this technology represents much more than a research field: it determines our capacity to organize knowledge and give it meaning, it increases our decision-making capabilities and our control over these

1. Direction générale de la sécurité intérieure (French internal security directorate).

systems and, most notably, it enables us to capitalize on the value of data. Therefore, artificial intelligence is one of the keys to power in tomorrow's digital world.

Because of this, collectively addressing this issue is in the general interest; France and Europe need to ensure that their voices are heard and must do their utmost to remain independent. But there is a lot of competition: The United States and China are at the forefront of this technology and their investments far exceed those made in Europe. Canada, the United Kingdom and, especially, Israel hold key positions in this emerging ecosystem. Considering that France and Europe can already be regarded as "cybercolonies"[2] in many aspects, it is essential that they resist all forms of determinism by proposing a coordinated response at European level.

This is why the role of the State must be reaffirmed: market forces alone are proving an inadequate guarantee of true political independence. In addition, the rules governing international exchanges and the opening up of internal markets do not always serve the economic interests of European states, who too frequently apply them in one direction only. Now more than ever, we have to provide a meaning to the AI revolution. This is the aim of this report.

A meaningful AI implies that we know the way forward.

A meaningful AI implies that we know the way forward. This is the objective of the industrial policy presented in part 1 and structured around four strategic sectors: health, ecology, transport/mobility and defense/security. These sectors have several characteristics in common: they serve the general interest and the major challenges of our time, they may constitute a comparative advantage for France and for Europe and they all require State intervention for their structuring. These sectors will be developed via precise and specific innovation awards which will establish key objectives and also by means of an aggressive policy concerning data. The benefits of data, which are central to developments in AI, are currently enjoyed by a set of a few major stakeholders who tend to limit their capacities for innovation to their ever more powerful enterprises. It will only be possible to redress the balance of power by extending the circulation of this data; this would benefit not just public authorities but also the smallest of stakeholders in the economy.

France plays a decisive role in AI research: French researchers have been involved have been involved in a major breakthrough in AI and French schools of mathematics and information technology enjoy international acclaim. Nevertheless, there is an ever-greater outflow: each week, researchers are recruited by private and frequently foreign enterprises and leave the state laboratories. It is therefore essential to provide public research with more resources to enable it to achieve its ambitions within a system ranging from training to transfer and innovation.

Finally, the economic development of the artificial intelligence sector needs to make ecology its first priority. This is crucial for the sector, as mentioned above: innovations in AI could be used to optimize energy consumption and recycling and achieve a better understanding of the effects of human activity on the environment.

2. This expression was used in a report by Catherine MORIN-DESAILLY for the Committee for European Affairs in 2013 (*L'Union européenne, colonie du monde numérique ?*).

But we need to ensure that the artificial intelligence being developed makes the most economical use of energy and resources possible.

A meaningful AI is another way to say that it is not an end in itself. Its development should take several considerations into account. First, the need to formulate ways in which humans and intelligent systems can work together. Whether at an individual or a collective level, this complementarity may take different forms and could be as alienating as it is liberating. The need to establish an enabling complementarity should lie at the heart of the development of AI, inasmuch as it would allow the de-automation of human tasks. To encourage the movement of tasks and professions in this direction, experiments should be set up across all communities, focusing particularly on the populations most affected by automation.

In a world marked by inequality, artificial intelligence should not end up reinforcing the problems of exclusion and the concentration of wealth and resources. With regards to AI, a policy of inclusion should thus fulfill a dual objective: ensuring that the development of this technology does not contribute to an increase in social and economic inequality; and using AI to help genuinely reduce these problems. Rather than undermining our individual paths in life and our welfare systems, AI's first priority should be to help promote our fundamental human rights, enhance social relations and reinforce solidarity. Diversity should also figure within these priorities. In this respect, the situation in the digital sector is alarming, with women very poorly represented. Their under-representation may lead to the spread of nurture gender-biased algorithms.

Our digital society cannot be governed by black box algorithms

Finally, our digital society could not be governed by black box algorithms: artificial intelligence is going to play a decisive role in critical domains for human flourishing (health, banking, housing, etc) and there is currently a high risk of embedding existing discrimination into AI algorithms or creating new areas where it might occur. Further, we also run the risk that normalization may spread attitudes that could lead to the general development of algorithms within artificial intelligence. It should be possible to open these black boxes, but equally to think ahead about the ethical issues that may be raised by algorithms within artificial intelligence.

A meaningful AI finally implies that AI should be explainable: explaining this technology to the public so as to demystify it—and the role of the media is vital from this point of view—but also explaining artificial intelligence by extending research into explicability itself. AI specialists themselves frequently maintain that significant advances could be made on this subject.

More generally, there is a need for collective debate on the subject of this technology: the constant acceleration in the patterns of its deployment should not stand in the way of political discussions on the purpose and validity of our objectives.

Executive summary

Part 1 — Building a Data-Focused Economic Policy

In this area AI heavyweights, such as China and the US, and emerging AI powers, such as the UK Canada and Israel, are developing extremely different approaches. Thus, France and Europe will not necessarily take their place on the world AI stage by creating a "European Google", instead they must design their own tailored model.

European Data Ecosystem

A whole range of uses and applications rely on the availability of data, so this is usually the starting point for any AI-based strategy. Yet data currently mostly benefit just a handful of very large operators, so greater data access and circulation will be required to restore a more even balance of power by extending these benefits to government authorities, as well as smaller economic actors and public research.

For this to happen, the public authorities must introduce new ways of producing, sharing and governing data by making data a common good[1]. This should involve encouraging economic players to share and pool their data, with the State acting as a trusted third party. In some circumstances, public authorities could impose openness on certain data of public interest. Meanwhile in Europe, a number of reforms currently underway must provide for greater access and wider circulation of data. The forthcoming revision to the directive on the re-use of public sector information must provide an opportunity to speed up the opening of public data and outline the terms and conditions for access to personal data on public interest grounds. The current reform of EU copyright rules should at last authorize text and data mining and enable our public research to be more competitive.

This data policy must be designed with the aim of safeguarding sovereignty: it is vital for France and Europe to maintain a firm stance on data transfer outside the European Union. The AI strategy must also capitalize on the high protection standards enshrined in the incoming European General Data Protection Regulation (GDPR). Recent laws on individuals' rights to data portability[2] could therefore be part of a broader citizen-based rationale, to enable the State and local authorities to recover data with the aim of developing AI-based uses for public policy purposes.

Raising Visibility for AI Players

France has all the required assets to take its rightful place on the international arena, yet our companies and academic networks suffer from a lack of visibility both in Europe and overseas. Large companies sometimes opt to rely on dominant world actors in the sector, rather than entrusting their data to our home-grown talent, either because they are not aware of this wealth of skills within the country or because they prefer to adopt a very cautious approach. Our mission therefore suggests bringing together French AI actors under a unique and strong banner, which would include certifications and "innovation in the

1. Common goods refer to resources where use and governance are defined by a community.

2. Users' ability to receive their personal data for their own use or to transmit to another data controller.

field" awards aimed at singling out the most innovative AI solutions and attracting potential buyers.

This approach must also be set alongside a more organized approach to demand for AI, which could involve the creation of an information one-stop shop aimed at helping potential AI buyers outline their requirements more effectively and ascertain the companies that could best address their needs.

A Clear Policy to Focus on Four Strategic Sectors

It is vital to take advantage of our economy's comparative advantages and its areas of excellence in order to bolster the French and European artificial intelligence ecosystem. In this respect, our task force recommends avoiding spreading efforts too thinly, but rather focus on four key sectors: healthcare, environment, transport-mobility and defense-security. These sectors are all crucial from a public interest standpoint, all require strong impetus from the State, and they can all be the focus of interest and ongoing involvement from public and private stakeholders.

The business strategy for each of these sectors must allow for the creation and organization of ecosystems based on the different major sectoral challenges. Artificial intelligence should not be developed as an objective or an end in itself, but rather it must be a way to channel this energy to develop practical applications and uses that help improve our economic performances while contributing to the public interest i.e. early detection of diseases, the 4 Ps of healthcare[3], elimination of medical deserts, emission-free urban transport, etc. These various business policy issues and challenges, each specific to its own sector, go beyond the boundaries of AI, but could help provide a ripe breeding ground for its development.

The second key point of this strategy involves setting up shared sector platforms, which must provide secure and tailored access for the various participants in these different ecosystems (researchers, companies, public authorities) to useful data for the development of AI, as well as to software resources and extensive computing infrastructure. In a public-private continuum, these platforms must enable the various stakeholders to develop new functionalities that are tailored to the individual features of each sector.

Lastly, it is vital to streamline the AI innovation track with the implementation of innovation sandboxes, involving three key features: a temporary easing in certain regulatory restrictions in order to give free rein to innovation, support for participants as they address their obligations and lastly resources for use in field testing.

The State Both Transforms and Shows the Way

It is vital for the State to be a key driver in these various areas of transformation. Public authorities must ensure that they adopt the necessary material and human resources to factor AI into the way they address public policy, with the aim of both pursuing modernization and acting as an example to be followed.

3. Personalized, preventive, predictive and participatory healthcare.

This transformation will obviously take time and the various ministries and government bodies display varying degrees of progress in the field of AI. An inter-ministerial coordinator role should therefore be created, devoted to implementing this strategy, with support from a shared specialist center consisting of around thirty staff tasked with acting in an advisory capacity for the different government bodies.

Meanwhile, public procurement needs to be reviewed: this budget is estimated at close to 70 billion euros for the State, public authorities and local bodies each year and it is insufficiently oriented towards innovation. Our task force recommends a number of measures aimed at using public procurement to support European industries and at breathing fresh momentum into innovative public spending.

Part 2 — Promoting Agile and Enabling Research

The French academic research is at the forefront of worldwide exploration on mathematics and artificial intelligence, but the country's scientific progress does not always translate into concrete industrial and economic applications. The country is hit by the brain drain towards US heavyweights, and training capabilities on AI and data science fall well short of requirements.

Bringing Academics Together Within Interdisciplinary Research Institutes on Artificial Intelligence

It is key to bolster our position worldwide on AI research by setting up a network of independent but coordinating Interdisciplinary Institutes for Artificial Intelligence within defined number of public higher education institutions. These bodies would house researchers, engineers and students, and should be located all across the country, each one devoted to specific aspects of AI, and with a very strong focus on an interdisciplinary approach, notably by including social scientists.

First and foremost, it will be crucial to attract French and international academics, and these institutes will therefore have to create an attractive working environment in order to effectively address competition from "Big Tech". They should therefore be set up as AI "free zones", with a considerable reduction in administrative formalities across the board, hefty salary top-ups, and support in improving quality of living. These institutes could offer full-time positions as well as intermediary affiliate status for researchers who remain in founding establishments.

It will also be important to attract private partners, such as large groups, SMEs and start-ups, which can deliver brand new AI solutions, by enabling them to train their own engineers, recruit premium quality engineers, and make or consolidate technological breakthroughs. A range of options could be provided to enable participants to get involved on a tailored basis, based on personalized framework contracts that provide for a simple fast-track cooperation process.

These institutes should heavily invest to increase the supply of attractive and diversified AI training programmes. The presence of internationally renowned academics with the support of premium teams,

the opportunity to interact with world-class corporations via internships and innovation competitions, multi-disciplinary training programmes with joint degrees, and scholarships for Masters' degree and Ph.D. students should help significantly boost the number of students taking AI training at these institutes.

Lastly, it is essential to take a nation-wide approach to coordinate this interdisciplinary institute network from both scientific and administrative standpoints, in order to ensure that they are run efficiently and transparently. From a scientific standpoint, this involves the coordination of seminars, pooling training resources, coordination of internships and consolidation of their results. Meanwhile, in administrative terms, this will involve assessing the red-tape fast-track provisions granted to all institutes and ensuring that each one benefits from this set-up, while keeping procedures streamlined and ensuring that each institute can operate independently.

Research Computing Resources

AI research institutes need to have the computing resources required to compete with the virtually unlimited resources of private dominant actors. To do so, our task force therefore suggests setting up a supercomputer designed specifically for AI usage and devoted to researchers and their economic partners during their shared initiatives.

This supercomputer is vital but should also be rounded out by an access package to a private cloud set-up, developed European-wide and tailored to meet the specific features of AI in terms of computing time and data storage space.

Make Public Research Careers More Attractive

It is unrealistic to try to compete with GAFAM's salary scale, but the gap is currently so wide that it tends to discourage young graduates, even those who are extremely interested in public research and contributing to the common good to join public research institutions. Doubling salaries in the early stages of their careers at the very least is a vital starting point, otherwise the pool of young graduates interested in higher education and academic research will definitively dry up. It is also important to make France more attractive to expatriate or foreign talents, with financial incentives for example.

Part 3 — Assessing the Effects of AI on the Future of Work and the Labor Market, and Experiment Adequate Policy Responses

The labor market is undergoing vast changes, but it is not yet fully equipped to address it. There are considerable uncertainties on the effects of the development of artificial intelligence, automation and robotics, particularly on job creation and destruction. However, it looks increasingly certain that most sectors and companies will be widely reshaped. We are entering a new era of major technological transition and history shows us that previous periods of transition did not always run smoothly. Indeed, they sometimes involved drastic political

readjustment, which often hit the most fragile portions of the population the hardest. So it is important to face this issue head-on and take resolute action, while not giving in to panic or fatalism.

This firstly involves looking into the complementarity between humans and artificial intelligence: if we are to assume that, for most jobs, individuals will have to work with a machine, then it is vital to find a complementarity set-up that does not alienate staff but instead allows for the development of truly human capabilities, such as creativity, manual dexterity, problem-solving abilities, etc. This can take several forms. Firstly, it might involve a shift in labor relations to fully integrate digital challenges and develop a 'positive complementarity index'. More broadly speaking, legislation could be implemented to deal with working conditions at a time of increasing automation in order to factor in new risks. Lastly, formal education and lifelong learning should be overhauled in order to promote experimental teaching methods that can help graduates and staff develop the creative skills that are becoming increasingly vital.

Setting up a Public Lab for Labor Transformations

The top priority is to ensure that the ability to anticipate is sustainable, continuous and above all articulated with public policies. The publication of studies on the future of the labor market often sparks off fascinating collective debate, but does not always result in concrete actions, with public policy being only slightly adapted without fully taking into account the results of these forecasting exercises yet. Transformation can be extremely fast, while public policy implementation procedures are complex and difficult to steer. For example, professional training is worth 32 billion euros per year, with a vast array of funding channels and a whole range of different stakeholders involved.

It is therefore crucial to create a space where both prospective capacities, macroeconomic forecasts and analysis of changes in uses can be linked to concrete experimentation capacities articulated with actions aimed at certain categories of workers. A permanent structure could therefore be created to spearhead these subjects within labor and professional training public policy, with a twofold role: to anticipate and experiment.

This experimental approach can then be used to initiate logics different from those currently in force in vocational training, i.e. it is now broadly left up to employees, who take personal responsibility for their own training. Yet in light of the potentially swift or even exponential speed of transformation, it is difficult for current general programmes to incorporate all possible situations and take on board both the requirements of the entire population and the need for a fast but targeted approach. Furthermore, staff do not all react in the same way to the transformation of their jobs and do not all have the same ability to build a new career path.

In this respect, trials could be carried out to design programmes that target certain groups, whose jobs are deemed to be more at risk from automation and who would have more difficulty addressing their professional development without guidance. This approach involves moving somewhat away from the current strategy whereby employees

alone are responsible for their own career development.

Trying out New Professional Training Funding Methods to Successfully Deal with Value Transfer

Funding for staff training is calculated on the basis of a company's total payroll, yet the development of AI further promotes the transformation in value chains and reduce the link between those funding professional training and those who derive the value-added from it. Companies with a very small payroll can therefore create a large portion of the value-added in an overall value chain that they are responsible for extensively changing, e.g. by developing software for self-driving cars. Yet for the moment, they do not take part in funding the career transition of staff employed by other companies that operate across the value chain.

We therefore propose initiating dialogue with industrial partners on how value-added is shared across the entire value chain. This type of negotiation cannot be based on the usual formats for social dialogue, which mostly operate nationwide with a vocational branch approach. Trials could be organized by the International Labor Organization or sector social dialogue committees focused on products and value chains that are particularly affected by these value questions.

Training Talents in AI at Each and Every Degree Level

One clear target must be set: triple the number of people trained in artificial intelligence in France in the next three years, by ensuring that existing training programmes focus more on AI on the one hand, but also by setting up new programmes and new courses on AI on the other e.g. law-AI joint degrees, general modules, etc. All degree courses should be involved, i.e. 2-year, 3-year, Masters, Ph. D, etc.

Part 4 — Artificial intelligence Working for a More Ecological Economy

Carving out a meaningful role for artificial intelligence also means addressing its sustainability, especially from an ecological standpoint. This does not just mean considering the application of AI in our ecological transition, but rather designing natively ecological AI and using it to tackle the impact of human action on the environment. This is an urgent matter as world data storage requirements, inherently correlated to the development of digital technology and AI, could exceed available worldwide silicon production out to 2040.

First and foremost, France and Europe can spearhead this smart ecological transition by raising awareness on the international arena. The primary task is to consider both the impact of AI on achievement of the UN's sustainable development goals, how it puts pressure on certain goals and how it can accelerate others. AI must be included in initiatives emerging as part of the Paris Climate agreement and the Global Pact for the Environment.

Players in both digital and ecological transition must join forces, which require setting up a devoted space for AI research and energy resource optimization research to meet, and

promoting projects at the crossroads of life sciences and ecology, climate and weather research.

Consumers must also play a part in making these technologies greener. Our task force therefore proposes the creation of a platform devoted to assessing the environmental impact of smart digital solutions. This platform should also include a simple calculator to enable all citizens to gain greater awareness of these impacts and compare the environmental footprint of the various products, services, software and hardware.

Fostering Greener AI

It is also important to tackle breakthrough innovation in the semiconductor sector, one of the physical building blocks of AI. In this respect, neuromorphic[4] technology can allow for considerable energy savings, and France is already a pioneer in this area.

Public authorities must also act to make the value chain greener and support the European cloud industry to promote its ecological transition. Some market participants already provide excellent examples of energy optimization and these best practices now need to be extended to the entire sector. A certification process could also be set up to reward the most outstanding solutions.

Lastly, making the AI value chain greener will clearly require open hardware and open software, which are not only a confidence indicator but can also lead to significant energy savings and provide inspiration for initiatives currently underway in Europe.

4. Neuromorphic chips are based on the workings of the human brain.

Dissemination of Ecological Data

The development of green AI is only feasible if ecological data can be open. So it is vital to make currently available public data open to all, both researchers and European companies alike, out to 2019 in order to develop AI solutions to promote ecological transition i.e. data on weather, agriculture, transport, energy, biodiversity, climate, waste, land registry and energy performance assessments. Access to more sensitive data could be managed on the basis of more specific situations, e.g. to address sector challenges. It is also important to open privately-owned data where necessary.

Part 5 — Ethical Considerations of AI

Recent AI-led progress across a number of sectors (self-driving cars, image recognition, virtual assistants) and its increasing influence on our lives are driving public debate on the issue. This debate included extensive analysis of the ethical challenges raised by the development of artificial intelligence technologies and more broadly speaking by algorithms. Far from the speculative considerations on the existential threats of AI for humanity, the debate seems to focus on algorithms that are already present in our daily lives and that can have a major impact on our day-to-day existence.

If we want to develop AI technologies that comply with our values and social

norms, then it is vital to act now to rally round the scientific community, public authorities, industry, business owners and civil society organizations. Our mission has endeavored to put forward some humble suggestions that could lay the foundations for the ethical development of AI and promote debate on this issue within society at large.

Opening the Black Box

A large proportion of ethical considerations are raised by the lack of transparency of these technologies. AI provides spectacular results for reasons that researchers sometimes have difficulty to explain: this is known as the black box phenomenon, where we can see input data and output data for algorithm-based systems, but we do not really understand what exactly happens in between. AI can reproduce bias and discrimination and is becoming increasingly present in our social and economic environments, so opening the black box is a key democratic issue.

Explaining machine-learning algorithms has become a very urgent matter and is now actually a separate field of research, which must be supported by public authorities. Three areas in particular require an extra focus: obviously the production of more explicable models, but also the production of more intelligible user interfaces and an understanding of the cognitive mechanisms used to produce a satisfactory explanation.

Transparency is clearly key, but looking beyond this issue, it is also vital to facilitate audits of AI systems. This could involve the creation of a group of certified public experts who can conduct audits of algorithms and databases and carry out testing using any methods required. These experts could be called on in the event of legal proceedings, during an investigation undertaken by an independent administrative authority or on request by the Defender of Rights (*Défenseur des Droits*).

Implementing Ethics by Design

Research staff, engineers and business owners who contribute to designing, developing and marketing AI systems play a decisive role in tomorrow's digital society, so it is vital that they act responsibly and factor in the socio-economic effects of their actions. With this in mind, it is important to make them aware of the ethical issues involved in the development of digital technologies right from the start of their training. This aspect is lacking in today's courses at engineering school and in universities' IT programmes, yet the extent and complexity of ethical issues these future graduates will face continue to grow.

Looking beyond engineer training, ethical considerations must be fully factored into the development of artificial intelligence algorithms. A discrimination impact assessment could be introduced, similar to the privacy impact assessments already made compulsory by General Data Protection Regulation for some data processing. The overarching aim here is very simple: have AI developers consider the right questions at the right time.

More broadly speaking, the increasing use of AI in some sensitive areas such as policing, banking, insurance, the courts and in Defense (with the question of autonomous weapons) raises a real society-wide debate and implies an analysis of the issue of human responsibility. We must also consider the role of automation in human decisions: are there areas where human judgement,

fallible though it is, must not be replaced by a machine?

Setting Up an AI Ethics Committee

Our mission recommends the creation of a digital technology and AI ethics committee that is open to society. This body would be in charge of leading public discussion in a transparent way, and organized and governed by law. It should work alongside sector committees and combine short-term considerations, such as economic and industrial impacts, with the ability to take a step back and take the long view.

Recommendations from the committee, which would operate entirely independently, could help inform researchers', economic players', industry's and the State's technological decisions. Its recommendations could act as a benchmark for resolving ethical matters (e.g. on self-driving vehicles) and hence provide a standard for AI developments.

Part 6 — Inclusive and Diverse AI

Artificial intelligence must not become a new way of excluding parts of the population. At a time when these technologies are becoming the keys to opening the world of the future, this is a democratic requirement. AI creates vast opportunities for value creation and the development of our societies and individuals, but these opportunities must benefit everyone across the board.

Parity and Diversity: Acting to Promote Equality

Despite the slow but steady feminization of scientific and technical sectors, digital technologies remain something of an exception, with gender balance still very far off. As digital technologies and, in the very near future, artificial intelligence become widely present in our lives, this lack of diversity can lead algorithms to reproduce often unconscious cognitive bias in programme design, data analysis and the interpretation of results. One of the major challenges of AI is ensuring greater representation within our societies.

Educational efforts on equality and digital technology are obviously vital, but greater diversity could also be achieved with an incentive policy aimed at achieving 40% of female students in digital subject areas in universities, business schools and their preparatory classes out to 2020.

All moves to promote diversity in digital companies could be further fostered by a nation-wide approach to promote diversity in technology via a national database aimed at documenting gender inequality in the workplace and the provision of funds devoted to supporting diversity in AI.

Developing Digital Mediation and Social Innovation to Ensure AI Benefits All

Given the extent of future AI-led transformation, we have a collective responsibility to ensure that no-one gets left behind. For everyone to truly benefit from breakthroughs made in AI, our procedures for access to rights must change and our mediation capabilities must also be considerably bolstered. So our mission puts forward a proposal to set up an

automated system to help manage administrative formalities, aimed at improving public awareness of administrative regulations and how they apply to each individual's personal situation. In addition, fresh mediation capabilities must be developed to support those who require help, in cooperation with care networks already present nation-wide.

Lastly, it is crucial that public authorities support the development of AI-based initiatives in the social arena. AI-led innovation capabilities remain very focused within a small number of companies. Setting aside healthcare, social fields receive only a tiny portion of private investment. This set-up for the AI-led innovation ecosystem has consequences on the speed of progress made in social matters. In order to redistribute these innovation capabilities, public authorities could embark on specific programmes to support AI innovation in the social arena and provide the necessary systems for the various parties in the sector so that they can benefit from AI-related progress.

Part 1 — An Economic Policy Based on Data

The worldwide artificial intelligence race has escalated in recent years. In July 2017, China unveiled its roadmap[1] for the creation of an industry which will be worth $150bn by 2030. This is the Chinese response to its principal rival, the United States, which has been investing massively in AI for a number of years[2]. Considering such a duopoly, is there any room for France or for Europe?

The latter have considerable assets for muscling in on the world stage. France can rely on the excellence of its research and training, a pool of specialized start-ups, very large data sets and a worldwide industrial network; Europe can offer a market of almost 500 million consumers, cutting-edge research, world economic leaders and a financial power which might, despite its obvious fragmentation, stand up to the industry's giants. It is also structured both around a system of common values and around a legal framework that is in the process of alignment; from this point of view, it is on a par with the current leaders.

It is important to realize that the current colossi of artificial intelligence—the United States and China—and the emerging economies in that field (Israel, Canada and the United Kingdom in particular) have sometimes developed or are still developing in radically different ways. France and Europe will not necessarily need to launch their own 'European-style Google' to secure a place on the international stage.

A policy aimed at promoting data access, as well as their circulation and sharing. Data is the raw material of AI and the emergence of new uses and applications depends on it.

In this context, our mission recommends a three-pronged strategy.

Firstly, an aggressive policy aimed at promoting data access, as well as their circulation and sharing. Data is the raw material of AI and the emergence of new uses and applications depends on it. At the outset, it will be crucial to accelerate and flesh out the policy for making data publicly available (open data), in particular with regard to data which is critical for AI applications. For several years now, the open data process has been the subject of a proactive policy, mainly under the impetus of the Law for a Digital Republic[3]: these huge efforts need to be carried on. In addition, the authorities need to initiate new methods of data production, collaboration and governance through the provision of 'data commons'[4]; they need to take responsibility for providing incentives for economic stakeholders to share and pool some of their data and even, in certain cases, enforce them to make it public. Last but not least, such a policy must be consistent with the idea of sovereignty and should capitalize on European

1. Document available at the following address: http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757016/c5960820/content.html
2. As a rough estimate, the American digital giants represent a value of $2.2 trillion when the whole of the CAC40 only amounts to $1.5 trillion...
3. Law 2016-1321 of 7 October 2016 for a Digital Republic.
4. *Commons,* or common goods, describe a resource whose use and governance are common to everyone.

standards of protection. In recent years, the European Union has been committed to consolidating the European market (Digital Single Market) and that is also the purpose of the following propositions.

Secondly, efforts made through industrial policy need to be focused on four key areas in the development of AI: health, transport/mobility, environment and defense/security. The suggested measures are particularly aimed at structuring support for innovation around the major challenges of our time, uniting the various ecosystems around sector-specific pooling platforms and making space for experimentation. Here, the role of the State consists in laying the foundation for innovation and providing stakeholders with the means and the resources for breaking new ground, without actually steering the movement in any way.

Finally, this is about initiating profound changes in the State, which needs to be a driving force in these transformations. The authorities need to provide themselves with the financial and human resources that will be required in order to incorporate AI into the delivery of its public policies, as much with a view to modernization as to setting an example. This implies making progress in a number of areas, from public procurement to State policy relating to human resources and skills; but it also concerns its approach to innovation itself.

This section is the longest, not because it is more important than the others—all these priorities deserve the same amount of attention!—but because the recommendations it contains, particularly those which deal with data, are designed to bolster the others.

1. Reinforcing the European Data Ecosystem

The techniques of machine learning signal a break with conventional algorithms, especially because they mark the gradual transition from a programming approach to one that involves learning. This is what led the magazine *Wired* to predict 'the end of the code' in June 2016; in the future, we will no longer programme computers, we will train them instead. The functioning of a machine learning algorithm can be compared to the cognitive development of a child who learns by observing the world around him, by analyzing the way in which individuals interact and by reproducing implicit nonverbal rules. Roughly speaking, machine learning follows the same pattern: algorithms are now trained to learn by themselves without actually being programmed. Rather than programming a car so that it can drive by itself, the manufacturers will for example present it with an infinite number of driving scenarios so that it will be able to take action even in the most unlikely situations[5]. Data clearly forms the basis for this type of learning.

Even though machine learning is not the only expression of artificial intelligence (far from it), it is currently the one which is both the most used, the fastest developing and the most subject to global competition.

5. "At our test site in California, people throw themselves down flat in front of the cars and then curl themselves into a ball": Chris Urmson, director of the division Google Car (https://www.lesechos.fr/14/03/2016/lesechos.fr/021765692246_comment-la-google-car-utilise-le—deep-learning–.htm)

The point of departure for most artificial intelligence strategies thus lies in the accumulation of a large corpus of data. Many of its uses and applications depend directly upon the availability of data; it is, for example, the reason why the automatic processing of the French language is not as advanced as the processing of the English language. It is also the reason why translating from French into English works much better than translating from French into Thai, the corpus of Franco-Thai texts being in shorter supply.

The point of departure for most AI strategies lies in the accumulation of a large corpus of data

While raw data is essential, then its value is tenfold when it is structured and annotated[6] in such a way that it can convey information that is recoverable by AI techniques. The enhancing and the annotation of datasets are particularly important for machine learning, but this represents a difficult, time-consuming and very costly process in terms of both human and financial resources. This is why, in many fields, crowdsourcing (mass outsourcing) is used to collect and especially to annotate this information (particularly through the use of micro- task platforms such as Amazon Mechanical Turk). AI packaged applications generally rely on large bodies of data in the public domain (for example, multilingual texts produced by international organizations are used to improve automatic translation tools); but when it comes to the industrial domain, the onerous tasks of collecting and annotating become a strategic issue.

Data constitutes a major competitive advantage in the global competition for AI; from this point of view, it is undeniable that the tech giants have a considerable advantage. However, the volume of data is not everything: smaller datasets (small data) may provide significant results if they are coupled with relevant models.

Access to data nevertheless remains an essential condition for the emergence of a French and European AI industry. In an increasingly automated world, not only does public policy and performance of our research depend on this access, but also our collective capacity to determine the way forward for artificial intelligence and the outline of our automated society.

However, the current situation in AI is characterized by a critical imbalance between the major stakeholders (the GAFAM[7]: Google, Amazon, Facebook, Apple and Microsoft, and the BATX: Baidu, Alibaba, Tencent and Xiaomi—whose pre-eminence is entirely due to data collection and recovery) and the rest—businesses and administrations—whose long-term survival is threatened. Associated with this primary imbalance is the secondary, critical one that exists between Europe and the United States. For evidence of this, we only need to look at the flow of data between these huge geographical areas: in France alone, almost 80% of visits to the 25 most popular sites over one month are picked up by the major American platforms[8]. From this point of view, Europe can be regarded as an exception: both Russia and China, for example, manage to pick up the majority of their users' data. This is largely due

6. The annotation refers to the addition of information to data describing its content.
7. The acronym varies depending on whether Microsoft and Intel are included, but it still describes a very small number of companies.
8. A study by Cyberstratégie's Castex Chair: http://www.cyberstrategie.org/?q=fr/flux-donnees

to the proactive policy of their governments, which are working to promote the emergence of their own digital leaders[9].

For France and the European Union, data policy which matches the requirements of artificial intelligence therefore needs to be structured around the goals of sovereignty and strategic autonomy. At the outset, it should be stated that this balance is fragile, and this objective requires vision. It is, nonetheless, a prerequisite for the development of artificial intelligence in France and in Europe so that they can avoid becoming just 'digital colonies' of the Chinese and American giants. In the same way, it is possible to develop artificial intelligence without renouncing our strongly-defended legal and political traditions of protecting the individual. Moreover, one of the main points of our mission is to consider these high standards as strategic opportunities, even distinguishing elements, in the global artificial intelligence race.

The current debate on artificial intelligence coincides with the impending application of the General Data Protection Regulation (GDPR). Welcomed by some, scorned by others—for a multitude of reasons in both cases—the GDPR nonetheless remains one of the most ambitious pieces of European legislation in recent decades. It is also a rare example of the European Parliament playing a major role, mainly thanks to the initiative of Jan Philipp Albrecht, the German MEP. In many respects, this text constitutes a minor legislative breakthrough, not so much in terms of its contents (in France and elsewhere, algorithms and data processing have already been regulated for forty years) but for the message it sends out to public and private stakeholders as well as to the rest of the world. Europe has chosen to impose high standards of data protection: all businesses that are intending to process data belonging to Europeans are required to comply with the GDPR (the principle of extraterritoriality) or face record fines (2 to 4% of global turnover). The GDPR is, in addition, a powerful tool for consolidating the European digital ecosystem. If this legislation had existed 20 years ago, it is probable that Facebook, Amazon and Google would not have been able to penetrate the European market as easily and competition would have been established on a more equitable basis. The time required for them to adapt to the regulations could have made it possible for European businesses to develop competitive services.

Artificial intelligence within the context of the GDPR

The GDPR assists in the regulation of the usage of personal data, which means any information relating to directly or indirectly identified or identifiable natural persons. Obviously, the GDPR is relevant to AI on several counts.

Firstly, because it assists in the regulation of the conditions relative to the collection and storage of data of a strictly personal nature which could be used by artificial intelligence, as well as in the exercise of their rights by data subjects (the right of information, the right to object, the right of access, and the right to rectification).

In addition, the GDPR assists in the affirmation of the rights of the individual to data portability: Article 20 stipulates that 'the data subject shall have the right to receive

9. The implementation of an aggressive trade policy, the systematic leverage of public procurement, ongoing direct support and investments, etc.

the personal data concerning him or her, which he or she has provided to a controller'.

The GDPR also provides that the data subject shall have the right to obtain from the controller information about the operation of algorithms (Article 15.1 of the GDPR).

Encouraging Economic Stakeholders to Pool Their Data

In the digital sphere, innovation very frequently relies on open-door approaches and AI is no exception. Data itself is inherently conducive to free access and to sharing due to its uncompetitive nature and its low cost of production. Data as such is frequently of little value, but this increases when it is contextualized and cross-referenced. The person who collects the data is frequently not the only one to benefit from it, or the best placed to capitalize on it; hence the need to promote its circulation so as to maximize its economic and social utility. The Internet giants understand this perfectly; in addition to their remarkable sense and instinct when it comes to communications, the strengths of these huge platforms essentially lie in their capacity to capitalize on this inclusiveness and build whole ecosystems with themselves at the center (see inset).

The APIsation of the economy

If data is the fuel of the digital economy, then APIs (application programming interfaces) are its driving force. APIs relate to interfaces made available by platforms to allow third-party stakeholders to break new ground using their resources. Facebook used one of its APIs to introduce the button like online and thereby dominate the recommendation market. In the same way, the thousands of programmers who use Netflix's APIs are responsible for its success. According to its director, employing them as in-house programmers would have cost him almost a billion dollars per year. The dominance of these platforms is largely due to their capacity to aggregate ecosystems around themselves and then occupy the centers. The APIs are clearly at the heart of these ecosystems.

On the basis of this analysis, a growing number of considerations can be seen to characterize the data as a new infrastructure. This observation applies, for example, to an OECD report from 2015 relating to innovation and big data[10]. According to the organization, this justifies the pursuit of more ambitious policies of open access to public data, the promotion of data sharing between stakeholders and also the revision of the framework for legislative intervention in cases of monopoly. For many economic stakeholders, however, open access is still too frequently the exception to the rule (see inset).

10. OECD, *Data-Driven Innovation: big data for Growth and Well-Being* (2015).

For many private stakeholders, the figures show that open access remains the exception

In 2017, a study financed by the European Union established that around 90% of businesses questioned declared that they did not share their data with other businesses (Hofheinz & Osimo, 2017). Even within organizations, data silos constitute barriers to the reuse of data by different departments. As early as 2012, a survey carried out by the Economist Intelligence Unit came to a similar conclusion: 60% of businesses declared that corporate silos constituted the principal curb on the use of data for big data.

The fact remains that this movement towards open access represents a groundswell for the digital economy. In the private sector, we can see numerous spontaneous initiatives working towards varying degrees of free access to data. These may consist of 'vertical' exchanges between businesses within the context of bilateral partnerships, for example between main contractors and sub-contractors. They may include businesses allowing access to data on an occasional basis, frequently within the context of an initiative aimed at stimulating creativity on the subject of possible uses for this data ('hackathons', for example). As we have seen, businesses may still choose to make certain sets of data available via an API, free of charge or for a fee, in order to generate new openings and, ultimately, provide added value. Free access may equally be useful in education and training initiatives (this is mainly in evidence in Canada; it is virtually unknown in France). Finally, certain platforms have a completely open policy, a crowdsourcing approach, when it comes to data (eg OpenStreetMap).

Following in the footsteps of Waze, the American giant Uber—whose hybrid bike riders navigate almost a billion kilometers worldwide every month—recently embarked on the huge undertaking of promoting its data by making it available to local authorities. The company is sitting on one of the largest and most specific databases concerning urban traffic worldwide, far larger than many specialist agencies and municipal services. Although until now Uber maintained a tight control over its data in order to optimize the services it provides, today it is making some of this data available as open data via the Uber Movement, an initiative which has involved the city of Paris since October 2017; this data will make it possible to take a very detailed and proactive look at the flow of traffic in the Île-de-France region. Access to new data could equally provide full access to speeds registered on main traffic routes, for example, and make it possible to locate junctions where drivers are obliged to brake suddenly. With the same objective of winning over local authorities, Airbnb, the platform that allows individuals to rent out accommodation, has also launched its DataVille portal which gives access to certain statistics concerning the use of its services. Although these are obviously strategic moves on the part of the companies in question—certainly in terms of their image, as they actually remain in control of the data made available—they are nonetheless indicative of the forces at work.

Free access to and sharing of data generated by the private sector may therefore contribute to an increase in the mass of available data and thus contribute to the

development of artificial intelligence. The first offensive in the 'AI war' focused on data of a personal nature; this battle was won by the major platforms. The second offensive will focus on sector-specific data: this is where France and Europe can make their mark. For French and European stakeholders, the objective is primarily a strategic one since it is a means by which businesses in the same sector can compete with the world leaders in the field.

In certain cases, the sharing of data also needs to be encouraged in the interests of security, where solutions using artificial intelligence are concerned. In the example of the driverless car, today each manufacturer develops his own learning models. To ensure the reliability of their prototypes and achieve an acceptable level of risk, they are obliged to envisage the maximum number of possibilities: for example, they need to collect a year's worth of data relating to the running of the car so as to be able to address variations in weather conditions. In addition, references for the scenarios are only valid for the region concerned; roads and driving techniques in Paris are quite different to those in Mumbai, New York and Hong Kong. All these variables make it impossible for even the most experienced manufacturer to anticipate all the possible scenarios by himself. So although the American giants have gained a relative advantage in this field, they are still far from achieving an acceptable level of reliability[11]. Sharing data and references for autonomous driving scenarios (at least in part) therefore amounts to ensuring that, in the event of litigation, the vehicle concerned has a state-of-the-art validation plan and not one specific to a particular manufacturer.

Governments should therefore promote another data production and governance pattern, focusing on reciprocity, collaboration and exchange

Governments should therefore promote another data production and governance pattern, focusing on reciprocity, collaboration and exchange in order to foster the sharing of data between stakeholders in any given sector. Consequently, several countries pursue policies based on incentives for sharing private data, such as in the case of the United Kingdom where, for several years now, the Open Data Institute[12] has encouraged full access to private data so as to stimulate economic growth; for example, the ODI highlights the case of the company Thomson Reuters, which is developing a collaborative platform with the aim of making its data available to everyone. This approach is aimed at improving not just its customer relations but also the quality of its data, its products and its services[13]. In the United States, the Bureau of Transportation Statistics (BTS) operates a programme which lets airlines exchange certain sets of data concerning the take-up of domestic flights. Data collected in this way is aggregated and then its statistics are processed before it is made available to the transporters by the BTS to assist them in planning their own strategies.

11. The consensus on reliability in a driverless car is fixed at 10-8/hour, i.e. the probability of a serious malfunction occurring at any given time must be less than 0.00000001. This factor is 10 times lower than the European average for regulating faulty goods.
12. Created in 2012 with support from the Technology Strategy Board, which provides it with finance amounting to £10m over five years.
13. https://theodi.org/open-data-means-business

Government incentives for the sharing and pooling of data may rely on private initiatives or, alternatively, foster their development. These initiatives exist within many sectors; they would be worth supporting and promoting (see inset).

Regarding the sectors that the mission considers should be given priority in the development of AI (see the suggestions below): mechanisms for pooling data could be built into the recommended sector-specific platforms.

When it comes to sharing data, many initiatives are worth promoting

Founded in 2015, the French start-up Dawex aims to launch a stock exchange for data by centralizing exchanges between economic stakeholders. Unlike data brokers who buy, format and resell data, Dawex assists businesses with the contractual side of their data exchanges (licensing agreements, time span, territory, uses, sub-licensing capacity, etc) and makes sure that they abide by the legislation (in particular the GDPR) in force in the country where the data is being produced and processed. This start-up equally makes it possible for economic stakeholders to share data privately with corporate partners. This enterprise won the Digital Innovation Contest and has joined the Bpifrance Hub, following its funding by the Caisse des Dépôts.

Mention should also be made of the emergence of new services which are offering to aggregate public and private data: in the field of transport and mobility, for example, the French company Transdev has recently announced the launch of a platform which aspires to become the international 'Wikipedia' of open data, Catalogue (www.catalogue.global). The company is therefore endeavoring to collect and compile this data, to clean it and put it in an open format. Their objective is to reduce the barriers to the creation of innovative services—particularly for AI—in the fields of transport and mobility.

Still on the subject of transport, La Fabrique des Mobilités (The Manufacturers of Mobility) seems to be one of the most successful initiatives. This is the first European accelerator to be devoted to the mobility ecosystem. La Fabrique brings together all the stakeholders and projects, and capitalizes on feedback and errors to foster the emergence of a common culture of innovation. It is aimed at start-ups, industrial projects and regions which are developing new transport options. La Fabrique gives them preferential access to data resources whilst safeguarding the principle of reciprocity: to have access to this pool of data, a contribution must be made to it. This virtuous logic results in all the stakeholders involved in the project benefitting from this development of resources. The platform's appeal lies equally in the fact that it is able to offer different types of access to different stakeholders, depending on their nature and on their contributions.

Organizing sector-specific events to raise awareness and provide incentives for sharing and pooling data

This point is crucial: it is the role of public authorities to promote meetings between businesses that hold data—very often these are large private, public and semi-public groups—and start-ups and other stakeholders in the digital economy who might be interested in getting their hands on it and exploiting it within the context of AI

solutions. These events also need to promote the paradigm shift at work in the digital economy and highlight the advantages of free access in the development of AI.

These meetings could take the form of a 'Data forum': a platform for dialogue, ideation and, for some, acculturation. The aim would be to encourage data sharing 'by example', highlighting various initiatives in which other stakeholders could take part or which might inspire them to suggest a strategy that would really help them get the most from their data, through an approach based on sharing and 'coopetition'.

Supporting and advising businesses in their contractual arrangements bearing on exchanges of data

Lastly, the State could play a mediating role between businesses that wish to free up their data but do not know how to go about it. In conjunction with the CNIL (the French Data Protection Authority), the Direction générale des entreprises (General Directorate for Enterprises) could support these private stakeholders and provide a guide to best practice as well as standard contracts.

The aim is simple: to reduce friction and reservations, cultural or organizational, when it comes to the sharing and pooling of data, since its supervision is regularly neglected. To alleviate these difficulties, public authorities could recommend the creation of private charitable trusts aimed at the long-term structuring of data relations between economic stakeholders, voluntary organizations and sector-specific ecosystems. Certain bodies of data could be coproduced, using an approach based on common values and reciprocity which would be managed by such trusts.

Organizing Access to Certain Data Held by Private Entities on a Case-By-Case Basis

A review of the Law for a Digital Republic has allowed the emergence of a new concept: the data of public interest. This is a form of 'private open data' and applies to data which is of particular relevance in the efficient operation of the market and in public policy of public interest. The legislation brought in by Axelle Lemaire has already opened up this opportunity for public service concessionary companies, companies that run State-owned natural gas and electricity networks and also for statistical purposes. A similar obligation has been brought in that relates to certain data contained in what is known as the 'Macron law'[14] and the so-called energy transition law[15]. This is all about going one step further in the development of uses for artificial intelligence.

The findings of the mission carried out by Laurent Cytermann concerning data of general interest[16] expressed reservations regarding the possible creation of a

14. The Law 2015-990 of 6 August 2015 for growth, activity and equal economic opportunities and Law 2015-992 of 17 August 2015 relating to energy transition for growth.
15. Law 2015-992 of 17 August 2015 relating to energy transition for green growth.
16. A report from the IGF, the French Conseil d'État and the Conseil général de l'économie, de l'industrie, de l'énergie et des technologies sur les données d'intérêt général (the General

general 'data of public interest' status. The issues at stake were the impossibility of including public interest criteria which would apply to all sectors and the crucial balance that needed to be maintained in order to avoid infringing on freedom of enterprise; reading between the lines, there were also the risks of compromising the emergence of new services and undermining the equilibrium of emerging ecosystems and the risk that this access would mainly be of benefit to the Internet's major stakeholders. Our mission is aware of these reservations, which are all the more relevant now that the debate about AI is tending very rapidly to become divided. In the field of AI, there is no such thing as a standard approach; the development of AI depends on multiple sector-specific approaches and all the expertise, issues and data associated with them. For all these reasons, a general regime of free access to private data seems neither entirely possible nor wholly desirable. This approach could nevertheless contribute to the avoidance of the Balkanization of sectoral regimes, particularly in view of the various barriers and the resistance within the spheres under consideration.

Nonetheless, most of the stakeholders interviewed for this mission remained positive about the gradual opening-up of access to certain sets of data—on a case by case basis and depending on the different sectors—on grounds of general interest. This opening-up could take one of two forms: access to this data by public authorities alone, in order to feed into a public data platform, for example; or wider access (open data) which would be open to other economic stakeholders. The extent to which the data is made available will need to depend on all the factors being taken into account, in particular the economic, financial and competitive impact on businesses concerned. Legislation would need to ensure that these provisions would not dissuade businesses from undertaking the collection of this data or from inventing new business models. It is equally important to anticipate the cost of this access—following the opening of an API, for example, or the essential anonymization of personal data. The next review of the directive on the re-use of public sector information, which has been announced by the European Commission, will be an opportunity to accelerate the movement for access to public data and to define the terms and conditions of access to private data for reasons of public interest.

Possible uses for data of public interest

	Examples of data	Interest for AI
Health	Data relating to general well-being generated by devices connected to the Internet	Pre-diagnosis, assistance in getting patients appropriate treatment, etc
Transport	Data generated by motorway CCTV	Training driverless car prototypes, etc
	Transport data generated by hybrid bikes; geolocation of drivers and traffic speeds	The development of an intelligent and dynamic system for regulating traffic, congestion prediction, etc

Council for the Economy, for Industry, Energy and Technology concerning Data of General Interest) (September 2015).

Environment	Data from Linky electricity meters (individual energy consumption)	Optimization of individual energy consumption, more accurate estimates and the evening-out of peaks and troughs, etc.
	Data that concerns air pollution	Warnings, assistance with decision-making, controlling urban policy, etc
	Data relating to rainfall (e.g. the Montana coefficient, etc) and sunshine	Automated thermal auditing, etc

The Urgent Need to Promote the Practices of Text and Data Mining (TDM)

The European legislative framework needs to promote new uses for data. To this end, the current reform of the legislative framework relating to copyright and the protection of databases is an opportunity to achieve a balance which is more conducive to the flow of data and to allowing certain types of user's access to this data. Among the various elements of this reform, there is one that is of particular interest in terms of the development of AI at a national and European level: exception from the rules of copyright and the rights of producers of databases for the purposes of text and data mining.

'Text and data mining' describes a whole range of computer processes that involve extracting knowledge from texts or databases according to criteria of novelty or similarity. For example, it makes it possible to search for 'weak signals' that are difficult to grasp on a cursory reading, and to locate and analyze accounts of failed experiments. Text and data mining has enormous potential for scientific discovery and the development of new expertise.

Today, the duplication of databases essential to the setting-up of mining systems requires the explicit agreement of the owner of the work or the licensee of the databases concerned—even when access to this data is lawful, for example when a researcher has paid for rights of access so as to be able to read articles in a database belonging to a publisher of scientific articles. Ireland, the United Kingdom, the United States, China, Japan and, more recently, Germany and Estonia have therefore adopted legislation which allows researchers to digitally duplicate databases from a legitimate source. In the absence of a clear legal framework, Europe is lagging a long way behind in the competitiveness of its research and, therefore, in its capacity for innovation. Alongside the new learning methods of artificial intelligence, authorized access to data thus represents great potential for a number of scientific projects, in particular within the context of interdisciplinary research.

The question as to whether such an exception should be limited to scientific contexts and non-commercial purposes is still to be resolved; our mission advocates wider dialogue on this question. In fact, many stakeholders—journalists, associations and businesses—could benefit greatly from this exception, especially for the automated processing of information accessible online. Today, an investigative journalist wishing to use text and data mining techniques to analyze site contents—to which

he nevertheless has lawful access—must comply with the access licensing required by each individual site or negotiate separately with each site for consent to carry out this automated process.

The Law for a Digital Republic has already granted such an exception to public research. This legislation is still waiting for an implementation decree. Researchers need to be able to benefit from this exception without further hindrance, especially since once it comes into force, the European texts in question will have to be incorporated into national law, which could cause further delay.

Implementing Citizens' Rights to Portability

The General Data Protection Regulation (GDPR) recognizes data subjects' rights to portability concerning the personal data that they have provided to a service provider. The Law for a Digital Republic goes further, allowing the retrieval of all data linked to a user's account (see inset).

Data portability in the GDPR and the law for a Digital Republic

In Article 20, the GDPR stipulates that 'the data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller'. Article 48 of the Law for a Digital Republic incorporates the right of the consumer at all times to retrieve all of their data. This text gives individuals rights with a broader scope than the rights to retrieval recognized by the GDPR, in the sense that it includes all data and not just personal data. Service providers (only the largest in the context of the Law for a Digital Republic) should therefore offer a free facility that allows the retrieval of all files that have been posted online as well as 'all data resulting from the consumer's use of his/her user account and which are consultable online by the latter, with the exception of those that have been significantly enriched by the provider in question'. Provision is also made for the consumer to retrieve other data linked to a user account, the perimeters of which have been defined by decree.

The law concerning the portability of data is one of the major innovations in recent French and European texts. In practice, all citizens may exercise this right in order to migrate from one service ecosystem to another without relinquishing their digital history.

The exercise of this right could be declined in the case of 'civic' AI applications: it is conceivable, in the medium term, that citizens might decide to retrieve their data from various services so as to make them available to a public stakeholder or a stakeholder in scientific research for the benefit of missions of general interest. The possibilities are numerous and exciting: in terms of health, for example, patient communities might respond to a call from a research institute that is committed to developing artificial intelligence that will make it possible to improve the detection and treatment of certain pathologies. A mayor might appeal to his constituents for the data that they have retrieved from transport applications for the purpose of optimizing traffic in his/her municipality. Using appeals for blood donation as an example, it is possible to imagine significant campaigns at municipal, regional and

national level for the establishment of the databases required for the development of artificial intelligence geared towards public service missions.

The merits of such a process are threefold:

- It would allow the creation of new databases for the use of public services.
- It would help give new meaning to the right of portability by allowing an improved data flow which would be under citizens' exclusive control.
- It could be implemented from the moment the European regulations come into force, without the need to impose new constraints on private stakeholders.

In order to ensure that the right to portability is truly effective, users will need to have all the appropriate tools at their disposal. This is the reason for the emergence of new services, which are volunteering to manage technical relations and the transfer of data from one service to another; the initiative personaldata.io, for example, takes the form of a *chatbot*, a virtual agent which handles the applications to service providers in the assertion of users' rights (the right of access to personal data, the right to rectification and erasure and to portability, etc). In a similar initiative, personal information management systems (PIMS) offer their users a dashboard, a 360° view of their digital life and the data being held by different services, with the possibility of controlling the various means of access to them. Although these initiatives are mounted by start-ups and associations of activists and are still at an embryonic stage, this movement should be able to take full advantage of the future regulations when they come into force and should be encouraged.

Reforming the International Framework Applicable to Data Transfers

Although in France and in Europe it is crucial to create genuine ecosystems around the data needed for the development of AI, this situation should not, however, result in facilitating the transfer of data outside the European Union. This concerns the principle known as the free flow of data at an international level. Through large-scale lobbying, the tech giants have long called for a policy to be established; they see that this has strategic value in terms of the current imbalance in the flow of data.

Such legislation, incorporated into free-trade agreements, would be a serious setback for Europe in terms of sovereignty, competitiveness and consumer protection. It would leave Europe with no room for maneuver in terms of the possibility of restricting the flow of data in the future.

This would not be improved by the fact that in practice, the free flow of data is made possible through international agreements—notably the '*Privacy Shield*' agreement which is responsible for a large proportion of the transfer of data between the European Union and the United States. This agreement, the follow-up to its predecessor which was invalidated following the revelations of Edward Snowden, still includes a great number of grey areas and does not provide sufficient guarantees for the protection of the personal data of Europeans. For this reason, it should only be seen as a transitional arrangement.

It is vital to get on with negotiations for an agreement which would be more robust from a legal point of view, in order to guarantee the protection of personal data belonging to all Europeans; the framework for this would need to be sufficiently stable for our businesses. We also need to be fully aware of the existing imbalance

in terms of the flow of data between the United States and the European Union. Enforcement of the GDPR coming next May and alignment of national legislation should be an opportunity for negotiations based on a firmer footing.

2. Consolidating and Raising the Profile of the French AI Ecosystem

Europe and France have a high-quality industrial and academic network at their disposal. They are in a position to occupy a central role on the international stage; however, our stakeholders suffer from a genuine lack of visibility. On the other hand, the Asian and North American giants (BATX and GAFAM) guarantee an international reputation for the whole of the ecosystem that they play host to. In the absence of such powerful unifying forces, we must reinforce the connections between stakeholders in our ecosystem with two goals in mind: raising their profile and reinforcing competition in both the domestic and export markets.

This lack of visibility is also attributable to the fact that stakeholders in the ecosystem are less able to promote and communicate about their capabilities and their successes. To take an example: whilst the exploits of the major platforms are frequently covered by the mass media, the French company Therapixel's victory in an international competition in the field of medical imaging was given very little coverage in France. There need to be changes in the rhetoric and in attitudes in order to be able to promote the national ecosystem more widely.

Creating a 'One-Stop Shop' for Information Relating to AI

At least two problems are encountered by potential purchasers of AI solutions: those of formalizing their requirements and identifying the stakeholders who could provide them with a solution. The most common situation goes as follows: a business has a wealth of data history; it would like to make use of this to improve its systems and generate new applications and opportunities. By default, this business tends towards convenience. Understandably, it approaches the international leaders in the field, with their finely-honed rhetoric and their gift for communication, who offers, in exchange for this data, to take on the triple role of providing advice and help with design and development. This is where the problem lies: for the majority of these applications, there are frequently smaller-scale stakeholders who are able to meet their needs with more effective and sometimes less costly solutions. These stakeholders would benefit from being better identified and identifiable, allowing businesses to make a more informed choice.

To support future purchasers of AI solutions, it would be advisable to create a 'one-stop shop' which could give them advice concerning the nature of their requirements and the stakeholders that it would be appropriate for them to approach.

Creating Labels for the Purpose of Raising the Profile of Domestic Uses for AI

Establishing a label affiliated with French tech and awards for innovation

The visibility of our ecosystem is therefore a key factor in its success. We should look to the example of *French Tech* for inspiration here: in the few years of its existence, this label has been instrumental in its unification, establishing networks and raising the international profile of the French digital ecosystem; the example it provides could usefully be complemented by a specialization in the field of AI. Such a label, affiliated with *French Tech*, could specifically identify French stakeholders in AI: academic laboratories, manufacturers and communities of interest. It could serve as a basis for organizing events and for specific communication on the subject of research and innovation in AI (meet-ups, conferences, business communications, etc), thus contributing to the consolidation of our ecosystem.

The visibility of our ecosystem is a key factor in its success

Establishing 'homegrown innovation' awards

Beyond these issues of visibility and transparency, there still remains the question of perceived risk. In developing an innovative AI solution, a start-up will too frequently have to brave the resistance of large companies and public authorities who are reluctant to adopt solutions that are considered, sometimes wrongly, to be too risky. The establishment of 'home-grown innovation' awards for AI solutions could help to secure these potential customers. They could identify and reward businesses which have supplied operational solutions that have met their customers' requirements. These awards could initially be deployed within the context of public procurement before perhaps being extended to larger companies.

The aim would be to create a showcase for the public which would promote businesses whose solutions have been tried and tested and at the same time reassure future customers about the extent to which these solutions can meet their requirements on a permanent basis.

This label and these awards would need to be accompanied by the creation of a public information portal in order to contribute to the goal of greater visibility, and the 'one-stop information shop' mentioned above would need to give wide publicity to these labels as a mean of supporting them.

Consolidating Customer Sectors

The establishment and development of the French and European AI ecosystem should result in a wide and sophisticated range of options. Consequently, there is a need for it to be at the forefront of clear and well-structured demands, which are currently lacking in potential users of AI. The reason? Although they are aware of the great potential of this technology, traditional stakeholders remain a little ignorant on the subject. Preoccupied, on the whole, with distinguishing genuine innovation from the buzz surrounding AI, these businesses are still often only at the thinking stage when it comes to changes in their professions and in their business models or, in more advanced cases, that of experimentation—which does not always result in fully

operational developments. We should, however, be aware that we are not starting from a situation involving unconditional support: to change people's minds and convince them of the advantages of AI, we must first dispel the fears that are associated with the subject.

As such, although a critical proportion of the market will come from Europe, it is vital that French economic stakeholders are strongly convinced about AI. Identification and understanding of the industrial sectors' issues need to improve, and individual requirements and strategies for change need to be fostered. The aim would be to structure the domestic market and limit purchases outside of Europe as far as possible when there are better alternatives.

Initiating strategies for change at the level of economic stakeholders who are users of AI

Support for the provision of AI should therefore go hand in hand with the structuring of demands from its users. Traditional economic stakeholders should put themselves in a position to invest in AI but, in order to do that, they should not overlook the need to reflect on their strategy for internal change (business and financial models, and technical aspects) and their requirements and expectations.

On a national level, various entities within the General Directorate for Enterprise are designed to help instigate such reflections and initiate strategies for dialogue and change: the French National Advisory Council for Industry and its strategic sectoral committees, the French National Services Commission or alternatively the French National Commission for Cooperation and Commerce.

Since this technology may be deployed throughout a whole value chain, these dialogues will obviously need to involve all the stakeholders in the chain: large companies, integrators, start-ups, small and medium-sized businesses and major platforms, without whom they would only have a partial picture of the ecosystem.

These reflections need to go hand in hand with training initiatives designed for small and medium-sized businesses; they could be part of the overall scheme for the digitalization of small and medium-sized enterprises—which has, for the most part, been taken on by the regions—in which AI should be a cross-cutting theme.

Facilitating dialogue between AI's stakeholders and regulators

Certain sectors need to inform themselves well in advance about the specific regulations relating to the development of AI solutions, such as: the sector-specific regulations which apply to markets and financial stakeholders which fall under the control of the ACPR (Autorité de contrôle prudentiel et de résolution —French Authority for Prudential Supervision and Resolution) or the AMF (Autorité des marchés financiers—French Financial Markets Authority); the regulations concerning the security of information systems which fall under control of the ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information —French National Cybersecurity Agency); and the regulations relating to the use of personal data operated by the CNIL (Commission nationale de l'informatique et des libertés — French Data Protection Authority).

Sources of innovations in AI may be faced with uncertainties concerning the compatibility of their business models with the legal framework and the attendant

risk of penalties being imposed, especially when they are the sources of disruptive innovations: in this case, penalties are extremely substantial (under the GDPR, fines may be as much as 4% of a business's global turnover).

Dialogue with sectoral authorities should therefore be encouraged by making the necessary technical and human resources available. Certain sectoral authorities have in fact already set up teams devoted to studying and supporting innovative projects: in June 2016, for example, with the backing of the Banque de France, the Autorité de contrôle prudentiel et de résolution and the Autorité de contrôle des banques et des assurances (the authority in charge of the control of banks and insurance) set up the FinTech Innovation center.

Added to the complaints about lack of support, there is the problem of the response time deadlines set by certain sectoral authorities. In view of the ever-dwindling innovation cycles and the uncertain growth of start-up companies, this is a crucial issue. To overcome these difficulties, a 'one-stop shop' to support stakeholders in innovation could be set up, and a 3-month limit could be imposed on response time deadlines. Finally, the possibility of recourse to an ombudsman could be guaranteed in order to resolve certain individual cases, such as when regulatory authorities appear to contradict themselves.

French Tech Central

French Tech Central is a stopping-off point for information and a meeting place for French start-ups everywhere; it is located on the Station F start-up campus. 30 public services are available on-site, on a permanent or part-time basis, whose mission is to provide advice and guidance to businesses; the latter may make a request for a private meeting with a representative from one or more public services via a special online platform or attend information workshops run by administrations specializing in the problems encountered by start-ups.

The proximity of these administrations to the start-up ecosystem is intended to allow the testing of new deals from public services. For these trials constitute the initial stages of a project which is designed to involve the whole country, and in particular the 13 French Tech Cities, as of the first half of 2018.

Assisting in the development of stakeholders capable of delivering AI solutions for industry

AI solutions are the equivalent of one unique component destined to be incorporated into much more complex systems. These systems may be difficult to comprehend, especially for small organizations which are attempting to enter European or even global markets. Two scenarios present themselves: either a business is by its very nature in a position to go one step further and target the French, European and world markets directly, with the capacity to stand alone; or this is not the case and it needs to join a large group of other businesses in the role of a 'building block'. This second scenario requires industrial 'building block' integrators capable of supplying the various specialist markets with their specific demands. The diversity of the European industrial AI landscape has prompted this observation: the only way to exist across a fairly large section of business verticals

seems to be to organize ecosystems which include major stakeholders and those with varying degrees of involvement, from start-ups to integrators.

We need to provide incentives for the creation of ecosystems centered around sectors that use AI, that are organized by large companies and integrators from the industry by means of business partnerships. In practice, leverage can be brought to bear in three ways: through dialogue between user sectors as mentioned above; through using incentives to create business consortia within the context of public procurement; and through the creation of an information center to assist businesses in finding partners within the context of responding to State or private-sector calls for tender.

However, one difficulty persists in the creation and organization of these ecosystems. The integrators' objective is to provide solutions for the industry which can be reproduced and which are competitive and economically viable, using the 'building blocks' made available by stakeholders in innovation. For each of these 'blocks', limited maturity means greater integration whilst a more mature product may be carried independently by an integrator. In addition, in the interests of controlling a product, an integrator may wish to have full access to the 'building blocks' that he is integrating and this may be perceived as a risk in terms of intellectual property. In order to regulate these risks and these levels of involvement, provision of model contracts for ecosystems and guides to best practice would make for the establishment of a climate of confidence amongst stakeholders.

3. Leveraging Public Procurement

The financial volume represented by public procurement is difficult to assess. It is estimated at €71.5bn annually for the State, public institutions and local communities (depending on what is included; certain estimations mention a figure of €200bn). Just as private stakeholders need to be able to grasp the challenges of AI and become its purchasers, public authorities also need to be able to use it for their own requirements. The mobilization of this capital could thus fulfil a triple objective: satisfy certain of the State's requirements in terms of AI, support the ecosystem through public procurement and assist in the creation of a showcase that would be exportable to Europe and worldwide.

Public procurement remains insufficiently geared towards innovative procurement

Public procurement remains insufficiently geared towards innovative procurement. Reasons for this abound: customers are uninformed about the appropriate procedures for innovative procurement, there is an aversion to the legal risks involved in the operation of current regulations, and there is an aversion to the operational risks involved in the purchase of innovative solutions. Indeed, procurement should meet the needs of public authorities and be under an obligation to achieve results; this obligation should then be passed on to the contractor. With this in mind, innovative procurement represents a risk in terms of the quality, performance and sustainability of the product delivered for which, in the event of a defect, the public purchaser may be held liable.

Finally, the regulations expressly rule out—except in the case of exceptions—the exercise of European preference in public procurement even when the market is completely out of balance vis-à-vis foreign stakeholders. An additional challenge therefore consists in mobilizing public procurement so as to benefit the European artificial intelligence ecosystem, particularly since, on an international scale, certain States do not hesitate to act according to national preference: a perfect example of this is the *Buy American Act* in the United States. So we must not be naive and we must make the best use of the economic weapons we have at our disposal.

Adjusting the Thresholds for Applying the Regulations at European Levels

In France, the financial threshold above which the public authorities are subject to the public procurement order is €25,000 excluding tax; Under this threshold, the customer is only obliged to choose an appropriate offer, make good use of taxpayers' money and forgo systematically committing themselves to the same supplier when there are several other offers available that could meet their requirements. However, the thresholds above which European regulations apply are considerably higher[17]: €144,000 excluding tax for public supply contracts and State services; €221,000 excluding tax for public supply contracts and local and regional authority services and for central public authority public supply contracts operating within the domain of defense; and €443,000 excluding tax for public supply contracts and services from contracting authorities.

In order to free up the exercise of public procurement in the domain of AI, it could be useful to lower the French thresholds for the application of the Public Procurement Order so as to bring them in line with Europe.

Using Public Procurement to Support European Industry

There cannot be healthy competition amongst European and foreign actors if the former are not subject to the same rules a position in terms of access to public procurement. This is especially true at a time when a more and more blatant imbalance is evident in the global AI and—more broadly speaking—the digital industries.

For France and for Europe, this is a major issue of sovereignty: in AI, and more generally in all fields, there is a high risk of becoming dependent upon foreign technologies with no other choice than to use them under conditions established elsewhere. Worse still, to maintain our independence, we could be forced to deprive ourselves of major technological advances. When it comes to AI and all things digital, the State therefore needs to set itself the objective of reinforcing an industrial and technological base for the key sectors which are of strategic importance.

Therefore, at a European level, we need to introduce the possibility for public authorities—within the context of the awarding of contracts—to make allowances for the state of the European industrial and technological base by, for example, giving priority to a European actor when it is clear that there is an imbalance in the

17. According to the 2018 update: regulations 2017/2364, 2017/2365, 2017/2366 and 2017/2367.

competition. Such rule could only be possible within the context of real commitment and European negotiations.

Revitalizing Innovative Public Procurement

The contract engineering capacities available to administrations and their operators are liable to greatly vary. It is therefore essential to leverage the experience acquired by those who have already put these procedures into practice, particularly where the State Procurement Directorate and the Defense Procurement Directorate are concerned. The dissemination of experience gained could be achieved through the provision of access to documentary repositories, through the exchange of good practice and through greater communications concerning concrete results.

Therefore, in coordination with legal affairs directorates, we need to produce these documentary repositories and guides to best practice so as to be able to inform the public purchaser about innovative procedures and limit the perceived risks in investing in innovation. The creation of networks of purchasers extends beyond the context of AI, but these could be of great benefit in terms of the acculturation they would provide.

The priority should be given to the development of two specific processes. Firstly, that of innovative partnerships: during the tendering process, this covers the need for initial research and trial stages right through to the purchasing of the finished product, without having to reopen competition among the stakeholders between these various stages. This point is one of the major problems associated with making an exception of R&D, a point we will return to: when the work has been completed, if it is successful and the public purchaser is willing to proceed to the operational stage, he is obliged to reopen competition even when the results of the trials have been satisfactory and promising. The situation is not helped by the fact that frequently, as a result of this reopening of competition and after it has already undergone trials, there is found to be no market for producing the finished version of the solution, very often for financial reasons.

The second mechanism of interest in terms of innovative procurement is competitive dialogue. This is a suitable solution for complex procurement contracts, in which the public purchaser is not able to define alone and in advance the type of technology that would meet his requirements, or alternatively for which he is not in a position to set up a suitable legal or financial arrangement. This process offers public purchasers the possibility of a much broader dialogue with tenderers, with the aim of improving the quality and the innovative nature of the proposals that are submitted to them; it is not a means to accelerate matters.

Broader communications could come from purchasers who have had a successful experience with innovative procedures, especially when they have only been made use of to a limited extent, as in the case of innovative partnerships. We should, however, bear in mind that the implementation of these processes, which are very costly, requires a high level of commitment from administrations.

Establishing Protection for Public Purchasers in Order to Create Incentives for Contract Engineering

Contrary to preconceived ideas concerning public tendering, current regulations offer public purchasers a great degree of freedom. The problem is that signatories to public tenders tend to display an aversion to risk which limits both the take-up of certain measures and, more generally, innovation in the field of contract engineering. It is not enough to introduce flexibility into these processes; we need to consider the risks associated with signing contracts which involve the personal responsibility of the signatory authorities. Obviously, this may lead to conservative reactions on their part and a reliance on tried and tested procedures, especially when the regulations give them more room for maneuver.

In order to limit these risks and provide incentives for innovative contract engineering, protection for certain purchasers could be put in place. This measure could take the form of a specific identification of innovative purchasers which would formalize the required risk-taking so that lack of success would not be penalized. This would need to be accompanied by a system of devolved responsibility, where the responsibility of the State would take precedence, except where there is shown to be foul play or deliberate abuse. The aim would be to create a favorable setting for contractual experimentation which would include an acceptable element of risk in their structures and effective protection for those taking part in these trials.

Making Exceptions to the Public Procurement Order the Rule

Public authorities have unique room for maneuver within the framework of making exceptions in special cases: for research and development contracts, for contracts which concern the vital interests of the State and for defense and security contracts[18]. In these circumstances, they may choose to remove the standard constraints contained within conventional rules of procedure in order to exercise European and national preference, for example, or alternatively to award contracts by mutual agreement in accordance with appropriate procedures. On paper, these exceptions give a lot of freedom to the public purchaser. In practice, the public purchaser tends to be very cautious since, in these cases of exception, the conditions of use are not very clearly defined.

We need to make these exceptions wherever possible and they need to be accompanied by guides to best practice; this could contribute to safeguarding public purchasers, in particular regarding the exception made for research and development.

The exception made concerning the vital interests of the State itself raises specific questions; in particular, it is not always easy to determine precisely which cases are covered. In the field of health, for example, we can easily consider that creating and controlling a repository of data which relates to citizens' health would come under this exception. In any event, it would appear crucial to define its scope.

18. These are, however, governed by specific decree.

4. A Clear Choice: Focusing on Four Strategic Sectors

In order to strengthen the French and European artificial intelligence ecosystem, we must make the utmost of our economy's comparative advantages and niches of excellence. In other words, we must determine the priority sectors in which our industry can seriously envisage playing a leading role at global level and compete with non-European giants. Limited budgets also mean that we must not be tempted to spread our resources too thinly: public support to innovation must focus on sectors in which there are the greatest opportunities over the short and medium term.

Such choices bear on sectors that have acquired sufficient maturity to launch major transformation operations requiring large-scale investment. Even so, every effort must be made to foster experimentation across all other sectors, helping them to mature at little expense and assess the potentialities that AI has in store.

How have we identified these strategic sectors?

Impact: *it should bring about far-reaching transformations from an economic point of view as well as in terms of general interest;*

Ecosystem: *the ability to create and maintain momentum requires having a group of robust public and private actors to rely on from the start;*

"Initial fuel": *this may take a variety of forms, but whichever it takes, there must be enough of it available and usable over the short term. In this context, financial aspects play a lesser role. It would appear more important, at least initially, to provide one (or more) of the following: data, use cases, business knowhow, resources, flexible framework, market, etc. Data is obviously a key factor and constitutes a major comparative advantage.*

Finance and resources: *the financial aspect remains crucial although not enough on its own, and sectors identified must be able to mobilize public and private funding alike, along with the human resources required;*

Markets and openness: *actors' ability to make best use of their knowhow on public and private markets in France and abroad is also important with regard to scaling up and seeing the emergence of large-scale ecosystems;*

Duality and percolation of fields: *even when effort is focused on specific fields, these latter are also chosen in order to enable "technological percolation" (i.e. a technology developed in one field being rapidly transposable to another).*

Impetus from the State: *finally, the sectors concerned will require major initial involvement on the State's part in order to transform themselves, which is not on the cards for a great majority of industrial sectors.*

After considering the above requirements, our mission recommends that four sectors in particular be selected: health, transport/mobility, environment and defense/security. Each of them represents a major challenge from the point of view of general interest and they are likely to crystallize the continuing interest and involvement of public and private actors alike. The State could well play a key role in the structuring of AI industrial policy in these sectors, by providing the substance required for setting things in motion and structuring the ecosystem, by playing the

role of first customers via public procurement and by creating the conditions required for the emergence of a market able to stand sustainably on its own feet. Why not decide to prioritize other sectors, other niches of French and European excellence—banking or insurance, for example? Because it would appear that their development is less a matter of public initiative as it is of private impetus, largely initiated as is, and, in the opinion of the actors concerned, any State involvement in it would be undesirable. As regards the selected fields, however, strong action on the State's part is essential to creating the required momentum.

For all selected fields, the ecosystem to consider is a broad one, set to include (among others) companies, researchers and sector professionals, along with the ministries and other government bodies concerned. Generally speaking, the current obstacles to progress observed in these fields cannot be blamed on these various actors. On the contrary, they often seem to be well aware of the issues involved and to have a real desire to see AI developing in their sectors. The causes of such hindrances are thus to be sought elsewhere:

- Organizational limitations: administrations are not structured to make use of AI, as, by its very nature, it cuts across all their various missions;
- A historical legacy: appropriation of AI often comes up against a culture and ways of operating that are unfavorable to its development, especially as it bears on processes, purchases, and practices with regard to information systems and exploitation, acquisition and openness of data;
- A change of paradigm: AI invalidates conventional means of expression of need and specification in a context where emerging needs sometimes go hand-in-hand with solutions;
- A silo effect: the lack of forward-looking, cross-cutting thought on future uses leads to prioritizing systems designed in isolation, incompatible with future developments in AI. This lack often goes alongside a fear of losing control of one's data, a fear that keeps such "silo logic" going and greatly hampers circulation of data (including in-house).
- Material absence of platforms adapted to sector constraints and bringing together data of interest to AI, computing resources to exploit it and the software stacks required to develop experimental and operational applications;
- Regulatory and legal frameworks that may seem ill adapted to needs connected with AI development.

Implementing a Sectoral Policy Around Major Challenges

As regards artificial intelligence, fundamental changes are required in traditional forms of industrial policy. The obstacles mentioned above, the industrial landscape's complexity—startups, SMEs, right up to large industrial groups—and the frenetic pace imposed by such technologies make them ill adapted to the conventional tools for supporting innovation. The technological difficulties surrounding AI are very real. There is, however, a tendency to greatly underestimate problems arising from organizational, structural and cultural aspects of its development. Within the same organization, difficulties also arise with regard to the various actors' ability to communicate with one another. Consider, for example, data governance, which

requires full cooperation between lines of business, engineers, researchers and administrators. In this respect, AI significantly challenges organizations' historical legacy.

In the priority sectors selected, such transformation must be based on three focuses.

First of all, end-to-end support to innovation. Development of AI application is achieved through iterative confluence of data, occupations and algorithms. Emergence of an innovative AI technology is not in itself enough to enable its percolation for use in industry or the public sphere: AI technologies are designed to integrate into larger, more complex systems, of which they are only a component. On the border between overall transformations of organizations and transformation of occupations, support to AI should be envisaged at all levels, from upstream phases, by supplying the required material (data, computing capacities, usages and professional expertise) up to dissemination and marketing. This requires the involvement of all stakeholders (industrial concerns, administrations and occupations) from the outset.

Next, mobilization and structuring of ecosystems around major sectoral issues and challenges. It is not a matter of developing AI for its own sake, as an end in itself, but rather of channeling the energy expended into development of applications and usages that contribute to the improvement of our economic performance as well as the common good. In short, of making development of AI relevant.

Finally, organizations must be receptive of innovation, whether in technologies, usage or business models. Development of AI requires rethinking traditional methods of carrying out projects so as to be able to develop, experiment and (where applicable) fail in short, dynamic cycles. Such requirements—specific to digital technology in general and AI in particular—often contrast with conventional methods of project management, which are usually based on far less agile rationales, undoubtedly because goals to achieve are better defined and needs clearer.

From this point of view, the model provided by the USA's Defense Advanced Research Projects Agency (DARPA) is inspiring. Set up in 1958, and attached to the Defense Department, the institution is responsible for a whole range of technological revolutions, including the ARPANET network (ancestor of the Internet), the GUI computer and GPS. DARPA also gave initial impetus to the development of driverless vehicles.

There would be no sense in trying to replicate this model. Financial capacity, methods, culture and mentalities are not the same on the other side of the Atlantic. In addition, DARPA's success has much to do with a historical context of major integration of the military-industrial complex, which has no real equivalent in France or Europe.

Some of the Agency's methods and the spirit in which they are implemented should inspire us nonetheless (see inset). In particular as regards the President of the Republic's wish to set up a European Agency for Disruptive Innovation[19], enabling funding of emerging technologies and sciences, including AI.

19. President Macron expressed such a wish in his speech on Europe of 26 September 2017.

What makes DARPA's programmes so successful

Programme directors: they are acknowledged experts in their fields, well able to identify and promote promising applications and technologies. They are independent: they take decisions autonomously so as to avoid frictions due to hierarchical decision-making chains, and are appointed for a relatively short duration (3 to 5 years) in order to maintain momentum and expertise;

Risk-taking: by their very nature, technically ambitious projects have significant risks attached to them. Risk-taking must become an integral part of project culture and success rates as low as 10% should be nothing to be afraid of.

Dynamic short cycles: developments mainly focus on proof-of-concept projects and working prototypes with major market potential. They are carried out over short periods (5 years maximum), with the ability to start and stop projects immediately over the course of time and as successes and failures dictate.

Programmes with specific objectives: it is of key importance to clearly specify objectives sought for and take care not to be overly prescriptive with regard to technologies likely to solve the problems involved.

Funding: each programme is provided with an ample budget and finances several teams, while keeping their numbers down (between 3 and 5), and periods over which projects must be carried out are kept short.

In order for these methods to work effectively, special attention must be paid to three points: acceptance of risk-taking, ability to implement dynamic short cycles, and funding. Risk-taking poses a problem for cultural reasons: public money must be spent to meet specific needs and results must be guaranteed. A major policy choice must therefore be made, and publicized in order to implement it.

Ability to implement dynamic short cycles is often hindered by contractual constraints, which implies that public procurement procedures should be reformed (see the corresponding recommendations).

And finally, as regards funding, agreement must be reached on financing several teams for one and the same project over a short period of time, which means initial additional cost but finally ensures greater innovation capacity and result quality.

However, a number of cross-cutting problems (to do with security and ethics in particular) must be taken into account right from the start, as they cannot be integrated *a posteriori*. This is a lesson learned from the world of cybersecurity: it is not possible to integrate security aspects as an afterthought without destroying much of what has already been constructed. It is essential to make project leaders and architects aware of the fact to ensure that it is taken proper account of from the start of AI projects.

Determining and highlighting major sectoral challenges

A fundamental change in our industrial policy must be brought about in this regard: structuring of support to innovation around major sectoral issues, ambitious long-

term objectives in industrial strategy, which go beyond AI itself but help provide a suitable environment for its development. Such issues may be wide-ranging, and specific to each sector: early detection of pathologies, P4 medicine[20], elimination of medical deserts, zero-emission urban mobility, and so on.

The interest of this approach is threefold. First of all, it leaves existing ecosystems free to structure themselves in order to propose solutions. As AI on its own does not enable such objectives to be met, it should make major contributions to them while enabling them to catalyze its development. The second advantage of these major issues is that they do not close the door on disruptive innovation, whether in terms of technology, usage or business models. Setting over-specific goals would come down to taking a technological stand that might well become obsolete over the short term, whereas major issues will point us in the right direction over time. And finally, industrial policy must be given a clear direction, enabling broad structuring of ecosystems around mobilizing projects.

What sectoral organization?

For each sector, the major issues concerned might be determined by sectoral committees tasked with publicizing them and facilitating their ecosystems. It is yet to be defined how such committees would be constituted; in certain cases, however, they could well be based on existing structures but also involve representatives of administrations, professions concerned, industrial concerns (startups, SMEs, mid-caps and large groups) and the public research sector. Such diversity would ensure that aims are ambitious enough, of operational interest, and significant in technological, social and industrial terms alike.

What funding?

Issues determined would be integrated into the conventional innovation-support systems overseen by BPI France[21], which could be complemented by special schemes:

- **Fluctuating subsidies** for provision of aid to development of highly innovative products. Fluctuation would enable examination of projects throughout the year and assign aid along the way as required. Inspiration might be drawn from the RAPID scheme (*Régime d'APpui à l'Innovation Duale*/ Regime for Support to Dual-use Innovation) implemented by the Defense Procurement Directorate (DGA) to ensure aid is provided within 3 months for approved applications.
- **Competitions**, models of which have already been implemented[22] and overseen by BPIFrance. They could possibly comprise a series of phases, with aid awarded becoming increasingly significant.

20. Predictive, preventive, personalized and participatory medicine.
21. Such as funding of collaborative R&D projects and funding of R&D.
22. In the context of the *Investment for the Future programmes*, to the tune of €35m and €40m per annum respectively for the world innovation competition and digital innovation competition.

- **Investment in equity** following a free-standing competition or in a final phase of a competition of the type referred to in the preceding point.

The rationale behind such schemes, however, remains the same in all cases: they all aim to enable emergence and acceleration of innovative AI projects that contribute to solving the major sectoral issues determined, and are therefore in no case prescriptive.

It should be borne in mind that such structuring cannot and is not intended to replace private investment. It is, however, a first-rate means of creating a technological showcase with financial aid. As regards private funding, Europe must put itself on a footing with its international competitors and ensure consequent further development of venture capital.

Organizing major challenges in combination with sectoral issues

Major long-term issues will not be enough on their own, however. Parallel thought must be given to emulation of the ecosystem over time, and emergence and implementation of innovative solutions in the meantime. They must therefore be complemented by implementation of schemes for development of operational capacities as early as possible (e.g. prevention of nosocomial infections or real-time detection of cyberattacks).

Support to innovation in the form of challenges currently has little place in the public approach to provision of such support, even though the method has proved effective, in particular in the United States with the abovementioned DARPA model. Such challenges must have clearly defined quantitative and operational goals and, in spite of everything, be ambitious enough to stimulate the ecosystem's innovation capacities, with major financial rewards as a key incentive. It is therefore suggested that innovation challenges be organized in each sector, with a view to funding development of technologically ambitious operational capacities over the short term, which also contribute to advances made on major sectoral issues.

The DARPA Grand Challenge

DARPA held this competition in 2004 and 2005 (as well as in 2007 in an urban context), with the aim of developing fully autonomous ground vehicles capable of:

- completing the selected course in under 10 hours;
- using GPS and possibly other available civilian signals;
- operating completely autonomously without receiving any orders while competing on the course;
- not hitting any other vehicles intentionally during the competition.

At the end of the competition, the 3 top teams were awarded prizes of $2m, $1m and $500,000 respectively.

There could be a variety of ways of organizing such challenges, depending on objectives and on whether public purchasing was on the cards. It is essential, though, that such challenges involve all interested parties, from researchers to industrialists,

for whom they would also provide opportunities to forge ties and set up common projects, so facilitating technology transfers.

Without the prospect of public purchase, such events could be held for purely R&D or experimental purposes, in which case financial rewards could be in the form of subsidies or procurement contracts as regards R&D. In this case, results would not be directly usable by the public authorities.

With the prospect of public purchase, in the event of results being intended for operational reuse by the public authorities, challenges should be directly designed and integrated into a procurement contract allowing for a post-operationalization phase. Such challenges would then constitute an initial phase of assessment and selection of the procurement contract, following which the public authorities would decide whether or not to make the final purchase. This in itself would constitute a further motive and reward for the challenge's winners, with an immediate prospect of final purchase. This model would also solve the problem of successful direct transference of innovation into operational circles, which is not possible with other models without reopening competition.

Whatever the case, implementation of such challenges will require major involvement on the part of administrations and their operators right from the very start, whether for facilitation, support, or purchase in the event of possible public purchase being incorporated into the competition. Organization of such challenges might also draw on any existing innovation structures.

Testing Out Sectoral Platforms

A platform is a service that plays the part of an intermediary in access to information, content, and services or goods published or supplied by third parties. It is a formidably efficient development model that provides so many Chinese and American giants with their strength. The term should not therefore be taken to cover a physical or technological implementation, but rather a **functional logic**: a platform enables ecosystems to structure themselves around functionalities it makes available to them. It must enable the design and deployment of products and services in connection with all its users, publics and private alike, in a logic of creation and distribution of value. It is therefore highly likely that, in the very near future, users (citizens, sector operatives, industrial concerns, etc.) will have access to a full spectrum of applications, from public services to private apps, even including public research experiments. This is the strategy that large platforms have implemented. Take personal assistants (Google Home, Amazon Alexa, etc.) as an example—they make resources available on their "clouds", the peripheral installed in households and the infrastructure that enables it all to operate. With development kits on public sale, third-party companies can deploy new functionalities on constituted markets.

Such logic should be adopted with all due speed, in order to reinvent public/private collaboration. The risk? Let others deal with it! We can see it with every passing day: the digital ecosystem is characterized by an omnipresent "winner takes all" logic and dominant positions seem increasing difficult to challenge. And the fields covered by AI are no exception, which is why it is up to the public authorities to introduce "plateformisation" into these various sectors, if only to avoid value being vacuumed off by a private actor in a paramount position.

Sectoral platforms should enable stakeholders in such ecosystems—industry, the public authorities, academic researchers, citizens and associations—to develop and market new functionalities adapted to the sectors concerned. In particular, they should:

- Gather data relevant to the sector and organize its capture (connected objects and specialized sensors) and collection (existing data);
- Set up secure differentiated accesses to application programming interfaces for ecosystem users (researchers, companies and public authorities) or, in some cases, direct to data;
- Give access to large-scale computing infrastructures including hardware resources and software adapted to AI and sector data;
- Facilitate innovation by possessing a capacity for experimentation within a controlled framework, especially if it puts forward rules constituting exceptions to common law;
- Enable continuous development, testing and deployment of operational and commercial products on one and the same support;
- Create ecosystem and platform logics providing users with direct access to national markets at the very least, by enabling them to deploy their applications in a "continuum" of services between public and private sectors.

Certain conditions must be met in order to ensure the success of such an initiative. First of all, setup of differentiated access is required in order to control and secure use of applications and (sometimes sensitive) data hosted there. Such requirement must enable operation of a proportionality principle between the goal sought by an individual wishing to access these resources and the means required to achieve it.

The second condition is that of openness and transparency. As regards questions of technological and economic sovereignty and questions of efficiency and performance alike, it is essential to prioritize use of open technologies ("open source" and "open hardware") as much as possible, so as not to fall victim to closed-shop mindsets. Public awareness of the platform and the data and resources it contains is a major factor in membership and mobilization of the ecosystem under consideration.

The final condition is that sectoral constraints in development of platforms must be taken into account, such as, for example, consent management in the context of personal data management. Apart from the question of compliance, this should finally provide the various interested parties concerned with a pre-approved toolbox, so dispensing with complementary developments they would otherwise have had to consent to.

Setting Up Innovation Sandboxes

It is essential to simplify the AI innovation pathway, in particular in priority sectors. A common complaint in all these areas is that there are too many regulations and too much time is taken over examination of applications to implement such experiments; the two problems are not unconnected.

Which is why our mission recommends to setup innovation sandboxes with a threefold purpose: temporary lifting of certain regulatory constraints in order to leave the field free for innovation, helping actors to take account of their obligations, and—last but not least—providing means of carrying out experiments in real-life situations. The Law for a Digital Republic made a first step towards such an initiative, in particular by the opportunities it opened up for experiments in telecoms[23].

As regards regulatory obligations, each sector has its own problems. Take airspace drone trials, for example, which are strictly regulated by the Civil Aviation Authority. On the technological side, experimenting with new applications may be submitted to a range of operational constraints, including use of cryptography techniques, database partitioning, interconnection and interoperability constraints, and tightening up of collection systems.

Actors must be provided with the opportunity to experiment under "real conditions"

In order to speed up AI development in priority sectors, actors must be provided with the opportunity to experiment under "real conditions". This is a major factor in the innovation ecosystem's attractiveness and is above all a one-of-a-kind advantage that only the public authorities have the power to provide. It is also an opportunity for the latter to try out new regulatory and technical frameworks, better adapted to AI problematics, under real conditions.

Sandboxes should thus act to facilitate experiments on full-stack basis: from iterative design to deployment of AI technologies in connection with their future users.

In order to ensure the necessary rapidity and simplicity of such an initiative, participation in sandboxes should be upon application on the part of actors in innovation, with examination of a single submission file and a 3-month "silence means consent" deadline. This would acknowledge the importance of such regulatory authorities as the CNIL, while making deadlines and conditions compatible with innovation.

Implementing a Data Policy Adapted to Each Sector

Access to data in priority sectors is of strategic importance, and is a question to be considered alongside the industrial policy and sectoral issues detailed above. We must cast a wider net in the hope of creating a "snowball effect" and increasing the range of possibilities open to innovation promoters. Drafting data policy does not simply mean thinking up ways of accessing or recovering existing data, it also means considering setup of new means for collection of quality data. And for such new means to emerge, technological expertise needs to be maintained and developed in Europe and is indissociable from expertise in AI.

Data governance

Data and platform governance is regularly underestimated, both as regards collection (what needs to be collected and how) and data management over time

23. ARCEP has already implemented this initiative and is still only just getting started.

(structuring, storage, life cycles, needs management, etc.), all of which require setup of a decision chain. This is a critical problem, and specific, identified decision chains therefore need to be implemented.

As pointed out above, however, access to raw data is sometimes not enough. It must be annotated in order to enable its optimal use by AI. This may require major investments and developments, which should nonetheless finally turn into profit through the value created by AI applications that would otherwise not have emerged. Setup of such means of collection should be regarded as a financial and operational priority for AI development, without anticipating what applications might finally be developed.

As regards annotation, inspiration could well be drawn from approaches similar to those implemented for "captchas" (see inset), either by inserting annotated data collection systems into operational systems or by adding means of measurement into tools already in use (activity monitoring tools, for example).

Annotation collection by "captcha"

The principle consists of requiring Internet users to read/identify an image, text or sound to differentiate themselves from a robot before they can submit a form. A part of the data thus generated is used to operate this distinction, while other items are not tagged, providing a subtle way of obtaining annotated data at little expense, both for the system operator, and, transparently, for its user.

There is also an opportunity to develop software designed as an aid to data structuring, so facilitating collaboration between human and machine for production of data usable by AI technologies and making such data preparation work less tedious. In the field of medical information, for example, it would enable pre-structuring of data based on free texts produced by physicians in order to minimize action on the part of medical information specialists.

5. Initiating European Industrial Momentum with Regard to AI

Europe has everything it needs to become a leading player in the global AI race: it is the largest market in terms of volume and possesses major academic and industrial advantages. In order to start on development of a European industrial policy on AI, our mission recommends that, initially, work should be carried out within a Franco-German axis. Italy (the north in particular) should also be seen as a possible serious partner, all the more so because of its advances in the field of robotics. Similarly, despite its specific position vis-à-vis the European Union, Switzerland possesses a wide range of industrial and academic skills that might be made good use of.

As regards the priority sectors, not all of them are suited to direct developments at European level. Concerning health, defense and energy, legislative and regulatory disparities between Member States would make a two-phased approach more appropriate, starting with consolidation of our domestic ecosystems and then going

on to deployment at European level. In this respect, launch of a special mission to study the possibilities of a European AI policy in these sectors and map the various obstacles to harmonization would be welcomed.

Developing European Robotics

Although robotics and AI go hand-in-hand in the collective imagination, the two fields are yet to truly converge. Many robotics applications are not within the purview of AI and vice-versa. There is, however, a whole field of exploration ready and waiting, and in which Europe has everything necessary to play a leading role, whether in terms of industrial robotics, for example, or agricultural robotics. This is particularly true considering that American domination of the field is yet to be established—despite such highly mediatized results as those published by Boston Dynamics.

On the same subject, development may take place on a Franco-German axis complemented by a partnership with Italy, which has a great deal to offer, in particular in the north of the country. In addition, a European flagship project has been submitted[24], in which this Franco-German-Italian triptych is very well represented.

Making Development of AI for Transport One of the Future Agency for Disruptive Innovation's Priorities

Plans to set up an Agency for Disruptive Innovation have been announced by the President of the French Republic, citing the DARPA model. This is a good sign, as it is essential to foster the development of large-scale projects with dedicated management and adequate funds. This would also be a way of mobilizing Europe's collective imagination around great ambitions—which has often been behind the successes of European construction—while employing methods for supporting innovation that have proved to be relevant in other countries.

Support to development of artificial intelligence must obviously be one of such an Agency's top priorities, keeping in mind that disruptive innovation should be designed in immediate contact with the people who work in them on a daily basis and are familiar with data and operational issues concerned.

If there is one sector that is particularly well suited to integration into a European scheme of this kind, it is the transport and mobility sector—one of Europe's longstanding strengths, bringing together all the above mentioned conditions combined with a very sizeable market, largely due to Franco-German constructors' and parts-manufacturers' importance in the automotive sector. The other priority sectors (health, defense and environment) do not lend themselves so easily to direct treatment at European level, although it would be useful to get Germany involved in the initiative in order to eventually convergence possibilities.

24. The document is available at this address: http://www.roboticsflagship.eu/

Innovating in the Components Industry Adapted to AI

The first postulate is as follows: all developments with regard to AI and digital technology rely on the existence of advanced components of various kinds (CPUs, GPUs and other variants for processors; embedded memories). It is such components and the power they provide that have made the latest advances in AI possible. The computer component industry is a very high technology sector of great strategic importance to Europe. Three requirements need to be fulfilled: develop R&D activities and maintain them over time, retain key skills and maintain adequate means of production. It has to be said, however, that very few industrial concerns in Europe are able to meet these requirements, especially as regards production capacities. There would be grave consequences if this European industry were to disappear: it would result not only in a dramatic dependence on hardware producers outside Europe, but also in Europe's inability to understand, design and produce electronic systems.

The fast-growing use of GPUs for AI

One of the reasons for the recent take-off of deep learning was the fact of GPUs (Graphical Processing Units) coming into general use. Their ability to carry out mathematical operations (essentially multiplications of matrices) in massively parallel fashion and the accessibility of their programming to this end were both determining factors. While a CPU contains a dozen independent processors, a GPU contains thousands of them, so dramatically accelerating the speed of calculations made and consequently of learning and data processing by machine learning algorithms, deep learning algorithms in particular.

Over recent years, technologies have emerged around industrial use of such processors. Tensorflow, PyTorch and Theano, for example, have enabled a wide audience not expert in GPU programming to access means of creating, training and deploying new models in extremely short loops.

Although European groups are managing to maintain strong positioning in some sectors, such as sensors, the situation as far as digital technology goes is alarming: they have abandoned production of advanced digital semiconductors, focusing instead on objects and peripheries rather than on the hearts of advanced digital systems. Many great European initiatives designed to support industry do already exist and are necessary in order to keep up skills, but the situation gives little cause for celebration though: Europe today is neither sovereign nor autonomous with regard to the entirety of the component production chain. This is especially the case with what would seem to be the most important of all: advanced processors and memory. If we include manufacture in key factors of independence, Europe possesses neither technologies nor means of production advanced enough (in terms of fine engraving, for example) to compete with Asia. The same goes for memory production. The issues of sovereignty and existence of extraterritorial regulations (such as ITAR) make it necessary for Europe to ask itself a tough question: how much importance should we give to autonomy in an industry in which it would seem hard to catch up lost time?

There is still some cause for hope, though: the past few years have marked the end of Moore's Law, which up until recently guided the components industry's R&D—a fact that requires the entire ecosystem to reinvent itself and innovate off the beaten track. This opportunity might be made the utmost of in the context of AI in order to produce new approaches, not in a technology race in the sense the term is usually understood, but rather to produce new, innovative, energy-efficient architectures.

Among the avenues envisaged, in-memory computing and neuromorphic approaches (see inset) would seem of particular interest. How well a system performs, of course, partly depends on the quality of its components, but not nearly so much as on the system's architecture as a whole (processors, memory and dataflow in the machine).

What is neuromorphic technology?

This technology draws its inspiration from the brain's internal organization and is capable of impressive cognitive tasks with less consumption than a light bulb. We speak of "neuromorphic chips". Neuromorphic systems are extremely energy-efficient in comparison with processors and graphic cards, due to their exploitation of two strategies. First of all, they bring computing and memory as close together as possible, so limiting data exchanges, which are currently the main source of energy consumption in processors. Secondly, they carry out computing less accurately than processors but in a much more energy-efficient way, either by using low-precision digital circuits (with small numbers of bits) or using the intrinsic nonlinearities in electronic components, which are an essential part of modern approaches such as neural networks. It should be borne in mind, however, that neuromorphic technologies do not necessarily solve all learning problematics. Several articles have tried to quantify energy gains obtained via such technologies: IBM's TrueNorth neuromorphic chip, for example, consumes 20 mW/cm2 compared with 100 W/cm2 for conventional computers, for implementation of neural networks. Learning is carried out offline, however. Online learning via these technologies remains a challenge our researchers have yet to resolve, and could therefore be the subject of an innovation challenge.

Source: The Centre for Nanoscience and Technology's contribution to the mission.

As well as provision of general support to the semiconductor industry, it might well be necessary to organize another innovation challenge bearing on construction of a supercomputer, for example, or embedded means of computing adapted to AI and only requiring European technologies. The aim of such a challenge would be to come up with new architectures taking advantage of European technological innovations—in the fields of in-memory computing or neuromorphics, for example. Such a challenge could well further the development of the transport sector at European level, especially in the event of setup of a European Agency for Disruptive Innovation.

Sunway TaihuLight

China has succeeded in producing the first supercomputer to be listed in the 'Top500' and the 'Green500's' top 5 using only Chinese components manufactured in China, even though it does not possess the most advanced production technologies at global level. To do so, it relied on preceding technological generations, which it succeeded in implementing in an innovative architecture.

Accelerating Setup of European AI Infrastructure

There are several types of infrastructure required for development of AI, covering the various phases from research to development and on to marketing of the product itself. In certain AI fields, such as machine learning, life cycles comprise two main phases: the learning phase and the inference phase. The speed and performance of the learning phase are conditioned by the scale of the material resources allocated, in particular as regards dedicated processors (GPUs, for example). Hence, infrastructure size conditions productivity and efficiency of research and development. The second phase, that of inference, has much less need of material resources, and can even be carried out inside embedded peripheries (an AI in a smartphone).

Learning and Inference

This is basically how AI techniques based on learning work: first of all, they go through a learning phase during which an algorithm seeks out all parameters enabling the model to carry out the required task at the best possible performance level. Once this phase is over and the model's parameters are set, an inference phase follows in which the task for which the model has been trained during the learning phase is carried out.

During the learning phase, one must distinguish between several types of workflows. Cases in which a supercomputer dedicated to AI calculation is fully mobilized (typically with resources numbered in thousands of GPUs) are quite rare and only concern a limited field of research. The great majority of applications require far fewer resources (numbered in dozens of GPUs, for example). This type of need that could well be met by an "AI cloud".

Setup of such an infrastructure requires very considerable investment and is the preserve of a specialized branch of activity: infrastructure, data centers and the cloud itself have to be taken into account. It is therefore a matter of pooling such resources as far as possible, at least for the public authorities overseeing the development of key sectors.

Why the cloud?

By making use of interposed networks, the cloud provides possible access to computing resources (networks, servers, storage, applications and services) that may be distant from users, transparently and with minimal intervention on the part of the service provider.

Although our instinctive reaction may be one of wariness, putting one self's data in the hands of a cloud provider does not mean giving up on security. On the contrary, making the cloud choice means turning to a specialized supplier which, by its very nature, will be more competent than the overwhelming majority of organizations, in particular as regards security. Due to the volume concerned, the quality of the service provided is even greater because it avoids the pitfalls usually encountered: assembling one's own small-scale computing infrastructures is financially, ecologically and functionally inefficient. As regards the powers that be, from central government to local authorities, it would be for the best if they put themselves in the hands of suppliers whose core business it is.

Deployment of such infrastructures should rely on European actors whose core business it is, in a context where the giants in the field are mostly American and Chinese. Thought should be given to the possibility of implementation via a public/private partnership seeking to help a European concern to make a showing specifically on the subject of AI. Such concerns can be counted on the fingers of one hand: according to the experts, the only company that would currently seem to have the capacity to hold its own on an international market is OVH.

This being so, the "*datacenter-as-a-service*" concept proposed by a number of economic actors would provide us with infrastructures managed by a specialist in the field and also increase its knowhow on the subject of AI. This would enable public research to combine the flexibility resulting from having a private cloud with guaranteed service quality.

6. Transformation of the State: Leading by Example

Together with businesses, the State must undertake a transformation in order to be capable of integrating AI into public policy management. Transformation is crucial to modernize and improve the effectiveness of public action, and also in terms of the 'State leading by example': it must therefore position itself as the primary user and buyer of AI technology.

Together with businesses, the State must undertake a transformation in order to be capable of integrating AI into public policy management

Appointing an Interministerial Coordinator to Implement the Strategy

Considering the scale of the transformations announced, the need to ensure sustainable coordination and management is vital: a relevant response to this could

be to nominate a High Officer for AI, as the British have done by setting up the Office for AI (OAI) in a bid to implement their recent strategy[25] (see inset).

Placed under the authority of the Minister for the Digital Sector (attached to the Prime Minister), this figure would be responsible for coordinating government policy, notably in terms of internal ministerial transformation and for forming the interface between the public and private sectors. They could call on the assistance of the administrations supervised by their ministry in order to perform this task.

In line with profoundly interministerial logic, they would be responsible for the daily coordination of a network of contacts within the various ministries and administrations in order to accelerate implementation of the transformations. This coordinator could refer to the technical expertise of DINSIC (Direction interministérielle du numérique et du système d'information et de communication de l'État —Interministerial Directorate for Digital Technology and the Government Information and Communication System) in order to assist administrations in their understanding of AI.

The British Office for AI

The British government announced the creation of the Office for AI (OAI) following publication of its strategy last November. Its role is to initiate the transformation of public policies using artificial intelligence, to encourage the appropriation of AI tools in the private sector, and to forge strong links with the economic and academic worlds. Jointly led by the Department for Digital, Culture, Media & Sport and the Department for Business, Energy and Industrial Strategy, the OAI and its director are responsible for leading the operational implementation of the UK's transformation strategy.

Creating a Joint Centre of Excellence for AI at State Level

Not all administrations possess the same level of maturity in terms of reflecting on the usage of AI in their specialist areas and their implementation processes. A major difficulty resides in the capacity to source the right skills for keeping up with the pace of innovation, identifying their applicability, and potentially transforming them into an initial proof of concept.

In this context, public authorities must rely on an organization whose mission is both to recruit profiles adapted to AI transformations and to act as an advisor and a lab for public policy design. This is a temporary arrangement only: over time, these skills should exist and be sustainable within the various administrations, which should be able to recruit specialists in AI from their own sectors.

DINSIC seems best placed to take on this role. With its directorate operating under supervision of the Prime Minister, DINSIC is currently responsible for coordinating the activities of administrations in terms of information systems.

25. Document available at the following address: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf

Creating a joint center of excellence for AI within DINSIC

DINSIC could incorporate a hub of excellence for AI internally and coordinate a network of skills found within administrations and their operators. Composed of thirty or so officers, their tasks could be to steer advisory assignments within administrations, ensure monitoring and mapping of innovations accomplished by the State, and also to create proofs of concepts and assist implementation on a larger scale in the event of success.

Lastly, the hub could be responsible for generally supporting the acculturation of public policies and could instill agile approaches within project management.

Creating a 'Reserve for AI' to support DINSIC

Following the model of Cyber defense reserves, the AI center within DINSIC could be supported by a community of citizens participating in a voluntary context (researchers, entrepreneurs, non-profit actors, activists, etc.). The objective: to establish open relations with society and external experts in order to build on expertise that may not necessarily exist internally.

The 'AI reserves' could be mobilized in the form of a jury or panels in order to clarify both the views of DINSIC and technological choices made by administrations. 'Citizen reserves', of which certain members would be experts on issues concerning 'predictive policing' (criminologists, data scientists, etc.), could be called on to give their opinion on technological solutions envisaged, within a multidisciplinary approach.

Strengthening of DINSIC and its right to notify

DINSIC currently plays an adjudicative role by assessing the performance of digital services. Its director is notably informed of the main projects envisaged by ministers—these projects have a provisional budget of between €5m and €9m. The directorate also has the right to veto projects in which costs exceed €9m. Their opinions are forwarded to the Prime Minister, to the ministers concerned, and to the Minister for the Budget.

Additionally, the right to notify (the Prime Minister and ministers concerned) is granted to the director for projects "which present stakes or risks that justify specific provisions and governance" and following consulting assignments "for any project or system of significant importance for which development or operating conditions appear to pose elevated risks or stakes in terms of scheduling, costs, quality or security".

This role must be strengthened with regards to State-led AI development. Reporting and right to notify thresholds could be lowered in order to provide DINSIC with an effective right of review and supporting role in important projects.

Integrating AI in the State's Digital Strategy

Whilst AI and digitization are not equivalent, it is apparent that the first would not be possible without the second: both must be considered as two successive waves.

As such, sailing on the first wave of digitization without regard for the second would result in losing ground from the get-go.

The AI component must therefore be immediately integrated into the State's digitization strategy, notably within the framework of the 2022 Programme for Public Action (*Action Publique 2022*). Or else, if this does not happen, a thorough review would be required before being able to manage the second wave of major transformation. Immediately capitalizing on this wave of digitization is essential in order to capture the opportunities AI presents to public services.

Consideration of AI has several implications: the design of data repositories which must integrate the possibility of them being used for ulterior purposes such as AI from the outset, purposes unknown during the design phase which potentially benefit initial third party use; the data capitalization policy: data which powers repositories is an asset; the collection and annotation of this data should be automatically questioned and studied even when it may not be of immediate use.

AI at the Ministry of Economy and Finance

AI has significant potential in areas such as user support, or even in the fight against fraud. The Minister of Economy and Finance has therefore launched initial projects in this respect:

1. a 'chatbot' has been developed by the CISIRH (Centre interministériel de services informatiques relatifs aux ressources humaines —Interministerial Centre of Information Technology for Human Resources), providing easy access to regulations concerning human resource management in the civil service for the benefit of managers within the Ministries of Culture and Social Affairs.
2. a 'chatbot' has been put in place by the AIFE (Agence pour l'Informatique Financière de l'État—Agency for French Government Financial Information Systems) for users of the information system 'Chorus', primarily composed of SMEs and microbusinesses;
3. a 'supervised deep mining' algorithm is used by French customs in order to detect fraud within value declarations, as well as an algorithm to analyze natural language designed to detect cases of identity fraud or import trafficking;
4. artificial intelligence modules have been developed within the SIRANO programme to fight against financial trafficking as part of TRACFIN, the unit fighting against money laundering and financing of terrorism.

Implementing dedicated and multiannual budgets for promising applications

Conventional operation of administrations does not always lend itself to testing, nor the emergence, of promising applications: in a tense budgetary context where emphasis is placed on choosing what not to do rather than what to do, it seems necessary to protect resources in order to avoid the issue concerning choice, which forcibly gives precedence to urgency.

This could involve the implementation of dedicated and multiannual streamed budgets, which incorporate the potential for cost-savings in order to encourage examination of promising applications, study of impacts and the launch of pilot projects. This concerns increasing flexibility in order to seize upon transformations linked to AI within an adapted working mode and pace.

This level of dedication makes it possible to move away from short-term needs; the multiannual, streamed element enables the evolving and responsive nature of AI to be broached, in contrast with annual scheduling tools, in as much as opportunities continually present themselves, projects come to fruition, fail and succeed.

Lastly, incorporating the potential for cost-savings enables negative costs to be incentivized, so as to avoid favoring saving one euro next year, against saving 10 or even 100 times this amount over the following years. The vehicle for multiannual programming laws could be studied.

Developing the reliability, safety and security of AI technology

Metrology

Public authorities must act in order to develop and implement standards, tests and measurement methods in a bid to make AI technology more secure, more reliable, useable and interoperable. In contrast to expert systems for which reliability and safety can be developed and tested by design (in theory in any case), systems which implement AI make decisions based on models built using data. In this way, protocols should be developed and incorporate new metrics in order to be applied to data, performance, interoperability, usability, safety and confidentiality.

In this regard, responsibilities of the LNE (Laboratoire National de Métrologie et d'évaluation —French National Laboratory of Metrology and Testing) could be expanded, within the realms of its historical remit, for it to become the competent authority in terms of assessment (for metrology) in the field of AI, and to build test methods required in order to achieve this.

Safety

Whilst AI fosters the emergence of new opportunities, it also fosters the emergence of new threats. A case study on this topic was the subject of recent publications which showed that it was possible to arbitrarily skew results produced by certain models informed by neural networks, which poses a significant safety issue for critical applications.

Whilst AI fosters the emergence of new opportunities, it also fosters the emergence of new threats

The example of the driverless car is significant in this regard: the existence of means used to skew its perception of the surroundings (deliberately causing poor interpretation of a stop sign, for example) could cause severe incidents. Safety is therefore a significant subject, notably for critical systems and systems with a physical component capable of causing damage in the event of attack.

Amongst the problems raised, we will notably discuss the possibility of the following occurring:

- arbitrary skewing of algorithm results due to the manipulation of input data;
- manipulation of data inputted during the learning period carried out by an AI algorithm;
- creation of new attacks based on the weaknesses of current AI techniques.

Safety is of clear concern to experts, but not uniquely. Collective awareness on the issue is required. Generally speaking, and more specifically in terms of AI, collective awareness must be considered from the outset of any process in order to avoid 'patch' culture, and safety should be considered from the design phase for technological products and solutions.

This is one of the reasons why it is useful to call on the support of specialist actors, who are able to propose solutions thanks to their experience and expertise. It is especially critical since recent events continue to report on the occurrence of security breaches, both in terms of software and material products.

The task of monitoring, foresight and study on the subject of safety and security issues posed by AI could be allocated to the ANSSI (Agence Nationale pour la Sécurité des Systèmes d'information —National Cybersecurity Agency), for which it could facilitate a skill network at State-level in the fields of cyber defense, defense and critical systems.

Standardization

One of the specific aspects of AI is the creation of de facto standards, notably of a technological nature: this is the case for deep learning for example, where technology such as TensorFlow (developed by Google) was adopted by an overwhelming market majority as soon as it was released, whether by individuals, startups or academics. Whilst these building blocks may avoid an ecosystem in which the same solutions are continually reinvented, they contribute to enforcing de facto standards.

This situation could prove to be highly detrimental if members of GAFAM (Google, Apple, Facebook, Amazon, and Microsoft), who remain the beneficiaries, decide to recover all of the developments made in AI that they enable.

As such, the greatest risk in terms of AI is not presented by the algorithms themselves, but rather by the technology (and human) "stack" which facilitates their implementation. In this context, standardization is not conceivable without maintaining very tight connections with the ecosystem as a whole: research, industry, innovation.

This approach must consist of reducing the trend for monopolization and logic of confinement. It will notably concern the establishment and application of non-proprietary interoperability standards within a proactive and coordinated approach, as well as local outputs for personal and non-personal data production tools.

Part 2 — Towards Agile and Enabling Research

Part 1 of this report particularly addressed the global data competition that is currently being played out—and one whose first rounds have so far gone the way of the world's tech giants. But there is another contest in evidence when it comes to AI: to do with human resources (HR). On the one hand, breakthroughs in science and technology are very often down to high-level researchers[1]; on the other, properly trained specialists are already in short supply in the global economy (a phenomenon which is only set to get worse in the years ahead, see below) and more highly qualified teaching staff are needed to train such specialists. The luring of our top talent by the AI behemoths through a premier HR policy is therefore taking its toll on both public research and the training of tomorrow's scientists in artificial intelligence (researchers and engineers).

Background and Requirements

French Higher Education and Research (HER) in AI has always played a leading role at international level thanks to the renowned excellence of scientific training in France—a constantly updated wellspring of the world's very best researchers. And what was true back in the 1980s—with the choice of the French programming language Prolog by the Japanese Ministry of Economy (METI), for its Fifth Generation Project—still holds today, with the rise in deep neural (deep learning) networks, several of the main stakeholders of which are French. The most famous of these is Yann Le Cun, who has spent many years working in the United States, currently New York, sharing his time between NYU and Facebook AI Research (FAIR).

The line between public and private research has become more blurred

And yet the AI research landscape has changed dramatically in recent years, and the line between public and private research has become more blurred: all of the foremost stakeholders of AI have opened hi-tech fundamental research centers, located in areas conducive to scientific development, and where there is a wealth of talented students and researchers to hand. Facebook has just announced the scaling up of FAIR's Paris center, and Google the opening of a research center in Paris. These success stories of France's appeal in this domain are praiseworthy indeed. But we should be wary of the drying up of the local AI public HER pool, as these private research centers are big draws for both high-level researchers and talented new graduates alike.

This means that not only has there been an endemic brain drain towards foreign academic institutions for a number of years now, owing to the differences in earnings and working conditions, but there is also a brain drain of researchers towards the major industrial players (GAFAMs and other unicorns). And, because of the necessary association between research and high-level education, the knock-on effects of this gain in pace is now sorely being felt at the training level—not least because industry-wide demand is rising (see inset).

1. In the rest of this chapter, academic "researcher" should be understood in the sense of a "professor, researcher or research engineer", employed by a university, graduate school or research organization.

French capacity in terms of university training and supervision at Master or PhD level has become critical (the master's course options in the field are now having to turn away bright students as their lecture theatres are so packed—Stéphane Mallat's classes at the Collège de France are fully booked out, etc.). The thriving private training programmes are of unequal quality—and they are cut off from research which, in the rapidly changing sector of AI today, is not viable over the long term. And yet there is a proven need on the job market for high-quality AI engineers, at all skill levels (see inset). It is therefore crucial that the French higher education potential in AI be enhanced considerably—in close conjunction with research.

A shortage of engineers trained in artificial intelligence

A McKinsey Global Institute report from 2011 had already warned of a deficit of 190,000 Data Scientists in 2018, as well as 1.5 million managers and analysts capable, quite simply, of understanding the ins and outs and of making decisions in the AI context.

The study published in early 2017 by Burning Glass Technologies, BHEF and IBM, meanwhile, predicts a 28% rise in the number of Data Scientist and Data Analyst jobs worldwide over the next five years, to a total 2,720,000, and that 39% of these jobs require a master's or PhD.

Lastly, in December 2017, according to a study compiled by Tencent Research Institute, there are just 300,000 "AI researchers and practitioners" worldwide at present—when the market demand is for millions of roles (even if there is not much detail on how such figures were reached). Tencent suggests that the bottleneck in this case is education. Incidentally, the study identifies the US, China, Japan, and the UK as the top contending countries in the AI race, with special mention made of Canada and Israel particularly in terms of education. France does not feature.

Another endemic problem plaguing French research (and not just in AI) is its poor performance in terms knowledge transfer to industry, whether to startups or multinationals (European where possible). Whilst the situation has improved in recent years, with the appearance of several structures aimed at fostering such transfer, there is still a brain drain of entrepreneurs, who prefer to venture abroad (primarily outside Europe, and usually to the US) on account of the better conditions they find there—both in terms of available funding possibilities and the business ecosystem and swiftness of decision-making processes.

On a final note, AI is well on the way to infiltrating all research areas, and it is being hampered in the process by French research's lack of interfaces between disciplines. For the mere fact of having gifted researchers in maths or IT on the one hand and, for example, in physics-chemistry or medicine on the other is not enough in itself to ensure robust interdisciplinary research.

This phenomenon is even more acute when it comes to data science: when the digitization of other scientific disciplines until now came up against problems essentially of a technical nature (such as the recording, handling and storage of data—whether or not on a massive scale), the arrival of AI is bringing with it challenges that call for specialists to work very closely together. To make the most

of the AI revolution across the scientific spectrum, it is necessary to set up a fully-fledged interdisciplinary approach between AI specialists and researchers in other disciplines. The latter bring original challenges to the table in the former's case, which enable research to move forward; in turn, solving them leads to disruptive innovations. But this virtuous circle cannot begin until there has been at least some effort on the part of researchers to familiarize themselves with each other's disciplines—beyond merely attending joint seminars.

The key factor setting AI apart from other scientific disciplines is its all-encompassing impact society-wide. This is not just some passing trend or media phenomenon, far from it: its implications are poised to be long-lasting and game-changing worldwide. AI is seeping into all sectors—economic, social, political and cultural alike… And most of the economic heavyweights, whether national or private, are fully aware of this fact and are investing massively in AI. The key question now is nothing less than what kind of society we wish to live in tomorrow. If we do not want to see such choices made for us by others, we need to protect our independence in this regard. And one of France's rare strengths in this field is the excellence of our scientific training, and the talented graduates this has produced. We must do everything possible to safeguard, enhance and turn it into scientific and economic success stories which champion our values.

1. Building a Network of Interdisciplinary Institutions for Artificial Intelligence

The flagship measure advocated here has three interdependent objectives: (re)shaping attractive and prestigious research environments that are capable of significant breakthroughs at international level and are grouped under a single, high-profile and renowned label; dispensing high-level scientific training in AI, for the researchers, engineers and entrepreneurs of tomorrow; enabling smoother interfaces between disciplines and between academic research and industry, expediting the transformation of ideas into proofs of concepts (POC), scientific applications and groundbreaking technology and intellectual property, capable of forging the fabric of startups and SMEs on which the industry of tomorrow will depend.

The three Canadian institutes

Canada, with a population of 36m, is considered one of the four world leaders in AI today. In its 2017–2018 budget, the federal government of Canada is devoting CAD 125m (EUR 80m) to a Pan-Canadian Artificial Intelligence Strategy, to support research and attract and retain talent in Canadian universities (master's students and trainees). The CAD 125m (EUR 80m) earmarked will be shared out between the cities of Montreal (CAD 40m/EUR 25.5m), Toronto (CAD 40m/EUR 25.5m) and Edmonton (CAD 25m/EUR 16m) and provide the main research institutes located in Montreal, Toronto-Waterloo and Edmonton with funding. Fund management is entrusted to the Canadian Institute for Advanced Research (CIFAR), which is also receiving CAD 35m in federal funding over five years from 2017–2018.

Setting Up a Nationwide Multidisciplinary Network of AI Research Institutes

France is a global leader in terms of research in mathematics and artificial intelligence. And yet research expertise is fragmented between universities, graduate schools and major research centers: the National Center for Scientific Research (CNRS), arguably the most involved in fundamental research, the National Institute for computer science and applied mathematics (INRIA), whose work ranges from fundamental research to transfer to industry and society, and the French Alternative Energies and Atomic Energy Commission (CEA), whose initial mission bore on the development of nuclear applications to the military, industrial and scientific sectors and which has since broadened its scope to become a key digital player in general, not least in the AI field. Mention could also be made of the other research institutes (INSERM, INRA and IRD among them) on which AI is also having a direct effect and which have therefore honed an expertise in AI geared towards their requirements.

Is this a consequence of this fragmentation? The fact remains that the weak links of French research are still to be found at the interfaces: between disciplines and between academic research and industry.

The 3IA institutes

In such a context, it is proposed to set up four to six Interdisciplinary Institutes for Artificial Intelligence (3IA institutes) nationwide, organized into a network: the National Network of Interdisciplinary Institutes for Artificial Intelligence (RN3IA). Set up in response to a call for tenders, immersed in a scientific ecosystem abuzz with potential collaborations, directly involved in higher education and closely connected with industry, these 3IA institutes will have to provide the whole of the chain from research right through to innovation with fora where productive collaborations can take place and knowledge associated with AI can be shared—and providing a significant proportion of the motivated stakeholders concerned (researchers, students, entrepreneurs) with direct access to cutting-edge research. In terms of research and innovation, the RN3IA will ensure national coverage of the AI fields by fostering the geographic and thematic diversification of the institutes: efforts will be made to avoid thematic redundancy between institutes, particularly as regards application fields or multidisciplinary research.

The research themes

More specifically, regarding research themes, a balance will be sought between concentrating endeavors and financing on the star subjects of the moment, and uniformly allocating across all present and past themes.

The scientifically and economically dominant themes of the moment (learning, and the different strands of data science, or data analytics, also known as Big Data) need to be delved into further; semi or unsupervised learning, reinforcement learning, representation learning and domain transfer as well as unstructured data learning (textual data, tweets, blogs and other electronic media for example) also need to be on the programme, widening the scope beyond deep learning alone.

Some lower-profile areas of AI (such as knowledge representation, the Semantic Web, distributed AI and game theory) should not be overlooked, for they ensure a diversity that must not be forsaken and which perhaps carries the seeds of the next AI revolution. Diversity is all the more important given that the GAFAMs seem to be focusing all their energy on the highest-profile areas, and it is therefore, perhaps, these other areas that we may be able to turn to our advantage, catching these AI giants with all their immense resources by surprise.

A series of strategic themes, which will be touched on throughout the report, has to do with ethics and the validation and certification of AI technologies, the aim being confidence on the part of all stakeholders in their results: from validation in terms of theoretical proof to explicability, transparency, causality and fairness.

More broadly speaking, theory (and that of deep learning in particular) is lagging behind practice today, and close collaborations with other fields of mathematics and the information & communication sciences and technologies (without really being able to talk about interdisciplinary, then) are to be set up, from game theory to logic and formal proof, from information theory to geometric approaches. In particular, France excels to such an extent in terms of proof of program correctness that a partnership between the two AI and proof communities can only result in major advances being made.

Without immediate theoretical progress, it is vital, to ensure swift dissemination of learning techniques, to acquire the means of choosing the right algorithm and then of configuring it, entirely automatically on the basis of data. We are talking about similar research to research in program synthesis here … which, incidentally, currently involves deep neural network approaches as well as other less conventional approaches.

Along with image and video analysis and vision processing, natural language processing has most likely gained the most from the arrival of deep learning (machine translation, textual entailment and understanding, generation), and language interfaces are being mooted as interfaces of the future, even though there is still some way to go here, too, before an AI program can pass the Turing test (the examples of the racist rants tweeted by the chatbot Tay[2] again throw into sharp relief the importance and difficulty of research on certification of AI techniques). In more general terms, the whole of the human-machine interface spectrum is already benefiting from the recent leaps forward made in AI, which are paving the way to new fields (other than security) such as Lifelong Learning in an open and uncertain world. And address concerns in the robotics sector, for which very close interaction between researchers from both fields seems necessary.

Another crucial area is that of optimization, whether or not in connection with learning. The field of operational research and combinatorial optimization has thus been highlighted by IVADO (Montreal) as one of their three areas of expertise, and one which they therefore distinguish from AI. In any case, the economic repercussions (all of the logistics for starters), and the impacts on AI in general (planning, constraint solving, etc.) are countless. All of the fields making use of

2. "A peine lancée, une intelligence artificielle de Microsoft dérape sur Twitter", LeMonde.fr 24 March 2016, http://lemonde.fr/pixels/article/2016/03/24/a-peine-lancee-une-intelligence-artificielle-de-microsoft-derape-sur-twitter_4889661_4408996.html

modelling obtained, thus far, by applying the basic principles now have a choice of alternative models to work with, which can be built using data. The ideal solution probably involves a combination of the two approaches to get the best of both worlds.

But, and this is certainly worth repeating, research fields that have not been mentioned here should also be retained, for the sake of encouraging originality. Along similar lines, although taking a slightly different tack, through the 3IA institutes it would also be advisable to encourage interdisciplinary research, of which no mention has yet been made.

Interdisciplinary

The multidisciplinary implications of AI have been touched on, calling both for fully-fledged joint research, rather than simply applying AI techniques to other disciplines, and, at the same time, training in AI for students and researchers in other subjects, to enable them to attain a genuinely twin skill set. Depending on the teams available (and willing) locally, each 3IA institute will focus on a small number of areas already being probed in its ecosystem, from HER to entrepreneurship, which it will be able to bring on board.

The target areas may, for example, have to do with the social sciences, economics and law, physics and chemistry, biology and health, ecology and sustainable development, computer-aided engineering, the human-machine interface or culture for example.

One particular area is the social sciences, because of the game-changer that AI is proving for the whole of society, which is naturally raising all sorts of ethical issues that come within its purview (see the section specifically devoted to this subject). All 3IA institute stakeholders will need to be made aware of such issues insofar as they can take different forms depending on the area we are dealing with.

Bringing Together Researchers, Students and Businesses

Researchers

The 3IA institutes will welcome world-renowned French researchers, drawn by the opportunity to return home, by the French culture or by the scientific reputation of their fellow researchers. They will also be tasked with rekindling the possibility of a bright future in France for young researchers who have received a world-class French education.

The perception that public research, though earnings are insufficient, offers up significant scope for freedom, needs to be challenged. First, because the freedom that researchers in the cutting-edge R&D laboratories of the GAFAMs is real enough—even if it varies from employer to employer. Second, because public researchers' freedom is seriously hampered by the need to fund their research programmes. They are spending increasing amounts of time responding to calls for

tenders, with very little chance of success[3], and that's without considering the time they spend reviewing peers' projects without the possibility of pursuing their favorite projects even. The alternative is often to seek out research subjects addressing industrial priorities, when the latter should only represent an ongoing, rather than binding, source of inspiration. This is all on top of the ever-increasing red tape that public researchers have to deal with constantly (hiring, purchasing, assignments, etc.), which also plays no small part in the lack of appeal—and competitiveness—of the French research environment. Not to mention the administrative burden that grows ever heavier the denser the administrative jungle to be navigated in the national HER becomes.

It is not realistic to imagine being able to close the gap in earnings between the public sector and the GAFAMs, but a high enough wage—enabling researchers to live in the Parisian region for example—is paramount for all that[4].

There is a genuine opportunity for improving the situation in public research by tailoring the employment conditions to ensure job stability—which alone will enable long-term research with peace of mind.

The working conditions—ranging from the computing facilities to the admin facilities particularly for foreign nationals—will be discussed below. It must be pointed out that the administrative jungle to be navigated is overly complex for the foreign researchers we wish to welcome and completely off-putting for French researchers abroad, who are acquainted with them through their peers and friends: *ad hoc* conditions shall have to be laid down as an absolute priority.

Researchers' obligations will amount to participating in teaching (one or two modules a year), leading a seminar and, on a voluntary basis, taking part in discussions with affiliated industrial members.

The institutes will be able to host researchers according to several statuses; it will be for the institutes themselves to choose the practical details (and recruitment methods) (principle of independence).

- Fellows will have full-time roles, be temporarily assigned from the civil service or hired directly, depending on the case. In addition to an appropriate salary, their funding shall come from an administrative budget earmarked for a team of doctoral students and postdoctoral researchers working on their research programme. They shall uphold the reputation of the Institute, hold scientific leadership responsibilities (research seminars, guest speakers, etc.) and be involved in the local higher education landscape. On a voluntary basis, they may also coordinate relations with the affiliated industrial members;
- Associate fellows will have a part-time role to play, devoting the rest of their time to their tenured position; they will also benefit from chair-type funding: earnings supplement and budget for a small research team around their project;

3. The success rate is around 12% for the ANR 2016 and 2017 calls for tender, and less than 5% for the FET calls of the European H2020 programme.
4. Today, a researcher just starting out, after 8 years of higher education, can expect to earn around 1.7 times the minimum wage.

- Affiliated researchers hold a permanent position with close thematic ties to the Institute. They will be co-opted by the Fellows to share the responsibilities and benefits, with a minimum obligation of regularly taking part in associated discussions and seminars;
- Researchers will be invited as visiting residents for periods of 3 months to 1 year, possibly spread over several years, making the most of their sabbatical or summer periods, for example, in a similar way to the Blaise Pascal chairs, or international Inria chairs. They may benefit from subsistence or accommodation allowances, and may also invite their students for shorter periods of time, and even hire 1 post-doctoral student for the duration of their stay.

Each of these statuses must be accessible at various levels of seniority: seniors will ensure the scientific coordination and reputation, while juniors will play an active part in supervising doctoral and postdoctoral students.

Education

3IA institute researchers will have to make a significant contribution to the higher education of AI in the region they are working in, bringing about or strengthening leading higher education programmes that are appealing by the presence of high-level researchers within the faculty. Recruited members' level of commitment in education will involve teaching one to two modules a year.

Continuing professional development for affiliated businesses and researchers working in other disciplines—which is as important as traditional bachelor/master education—may be delivered, for example, by the postdoctoral researchers hired by the 3IA institute. The organization of challenges could be a useful way of sharing data, experience gained and of passing on best practices in terms of "outlining a problem" and validating a solution in AI. Inspiration could also be drawn from the experience gained by the laboratory of excellence (labex) AMIES[5] with emphasis on the fact that demand largely outstrips supply. Good ideas abound, and are greeted with enthusiasm by industry and students alike: the barrier is the lack of time experienced by competent and motivated supervisors.

To overcome this, solutions involving bringing master's students on board to help teach Bachelor's students could be explored. The drawing up of challenges by the former for tackling by the latter (as practiced at the Université Paris-Saclay for example) is a striking example of practical and pedagogical innovation, in a field like AI where the importance of challenges is clear for all to see.

Four key points will need to be borne in mind regarding education in AI: they concern the diversity of students on the one hand, and the impacts of AI on society on the other (see the part on ethics in this report). Some people learn by proving, and others by doing. Students geared towards theory need to be taught in a way that does not separate out mathematics and computer science. At the same time, intensive testing modules must be on the programme, backed up by the appropriate

5. The *Agency for Interaction in Mathematics with Business and Society* provides First Exploratory Projects Support (PEPS) for laying the groundwork for budding ideas; an employment forum, etc.

computing means. Thirdly, the dangers of a blind implementation of AI call for the definition in the institutes of interdisciplinary courses (Maths/Computer Science/Social Sciences) that are likely to outline problems to do with the ethics of the types of AI we will be rolling out in the future—and, across all AI programmes, core modules aimed at making students aware of these issues. The network of 3IA institutes will be able to assist with dissemination in this respect. Lastly, students will need to be able to receive education for entrepreneurship, taking direct advantage of the presence of the institute's partner startups—and vice-versa, for example, as part of project-based learning.

Businesses

That businesses need to be able to attract, retain well-trained engineers and have a short circuit of interaction with cutting-edge research, is a demonstrated fact at all levels. The need for expertise, particularly with regard to the choice of technological solutions, is also proven, and the lack of this expertise is clearly damaging to French Tech's solutions.

The 3IA institutes will meet these needs, supply the industrial fabric with account taken of its diversity and represent a springboard for harnessing and transferring research findings jointly with industry. The chosen approach[6] is to get the entire chain working together on a daily basis, ranging from fundamental research to industrial transfer, from researchers to R&D engineers and private entrepreneurs, at both formal and informal events.

Most of the businesses conducting research into AI for their own innovation needs do not have the means to invest in scholars involved in fundamental research (unlike the GAFAMs and French or European multinationals, in their occupational sectors). By getting them to participate in these institutes, they will be able to: sustain an advanced business intelligence stance in the rapidly changing sector of AI; benefit from advice from researchers and the wider ecosystem, including computing means, for projects conducted jointly with the researchers; very swiftly bring to fruition, with a minimum of red tape to accomplish, the most promising projects in POC (Proof of Concept); and perhaps even launch more ambitious projects in partnership with researchers, such as joint laboratories or startups for example.

The direct interaction between public research and innovative businesses, startups, SMEs or "institutional" multinationals is currently hampered by red tape—even when all the technical partners have reached an agreement. Framework agreements will form part of submissions to the call for tenders launched for setting up the 3IA institutes. They may also feature in the administrative facilities provided by the national coordination of 3IA institutes. The aim is to obtain very streamlined decision-making processes and the associated formalities (ranging from Memoranda of Understanding to intellectual property agreements), on a timescale of one week for example.

6. This strand will be carried out in close liaison with the EngageIA initiative of the network of technological research institutes (IRTs), which provides its private partners with the first measures for assessing what they can gain from AI.

Businesses will be able to participate in the applications submitted by the institutions, gradually, depending on the level of maturity and financial strength, without being permanent members for all that ("affiliated" members status). In this way they will be able to access research seminars and properly trained students, as well as the advice of the 3IA institute members, on a voluntary basis by mutual agreement; the 3IA institute administrative support could comprise a framework agreement for this type of collaboration, with the possibility of researchers and industry working together in a reciprocal manner.

In practice, it should be possible to get involved in the 3IA institutes at several levels, corresponding to different degrees of participation in scientific life (this will again be for each 3IA institute to decide on):

- The permanent members would make an annual contribution, in the form of a fixed-rate amount, to the budget of the institute and the host institution. Some of their researchers (in-house research engineers) may be assigned to the institute's premises, on a full—or part-time basis, working together with the researchers—particularly benefiting from the irreplaceable ongoing advice through the "co-location" framework. To avoid the pitfall of this status only being accessible to multinationals, the annual contributions paid by the permanent industrial members of the 3IA institute may be adjusted based on the member's financial strength.
- The "affiliated" members will pay a smaller annual contribution, and their representatives may attend the institute's seminars to talk with the researchers about the occupation-specific problems they are encountering. This status would particularly allow startups who can already see what they can gain from it to interact at regular intervals (a few hours a week) with leading researchers (see the example of the Technion).
- The "occasional" members would, for a fixed-rate price, be able to benefit from a number of hours of consultancy with the institute's researchers, on a voluntary basis.
- Some potential entrepreneurs could benefit from guest status, for a short period of time, with a view to studying the viability of innovative ideas *in situ*. On this point, institutes adopting this model would almost play the role of a startup studio for innovations requiring the involvement of researchers, i.e. calling for the development of new fundamental approaches, and not just the application of existing technologies (also see a more general recommendation below on the creation of startups): the RN3IA could act as a correspondent and coordinator in this regard too.

In all cases, it is assumed that the R&D engineers and entrepreneurs taking part in the scientific life of the institute already have a good grounding in AI and an awareness of what they could gain from it.

The "Affiliates" programme at the Technion

The Technion (Haifa, Israel) is the oldest public university in Israel. Its Industrial Affiliates programmes (IAPs) are aimed at facilitating direct links between academic research and industry. Funded by membership fees, these programmes enable the

industry to access the research programmes and departments bearing on their areas of interest. The affiliated members can also attend working meetings and seminars, receive copies of reports and other publications and have facilitated access to students whom they may invite to come and present their work in internal seminars, with the prospect of perhaps being recruited. They are also privileged partners with the possibility of submitting collaborative project proposals to various national and European calls, and benefit from facilities for setting up joint research centers with the University's departments. These programmes transcend the traditional boundaries between academia and industry, resulting in win-win relationships against a backdrop of scientific excellence.

Support

Underpinning the appeal and effectiveness of the 3IA institutes are three support strands: 1) access to virtually unlimited computing means; 2) administrative procedures that have been streamlined as far as possible; 3) assistance with living conditions, not least for foreign researchers.

The *computing* strand is essential: the private stakeholders involved in AI research are equipped with vastly more advanced computing facilities than public laboratories and access to platform data which are quite simply out of public researchers' reach, on evident grounds of industrial property. The 3IA institutes will offer dedicated computing facilities, with pooling of an open data set (proposal).

The administrative burden must be eased

The *administrative* strand is also essential: any researcher drawn to France, only to be put off by the overly complex red tape and response times, will be a source of lasting bad publicity for our entire system. The administrative burden must be eased. The facilities enabled by not-for-profit associations or foundations, for example, must become the rule within the 3IA institutes—particularly when it comes to recruitment and purchasing. The civil service pay scales are dissuasive, a point we hope to have made clear, in light of two facts: 1) productivity varies by several orders of magnitude between researchers in this field; 2) there are numerous competitors who are able to pay high salaries after a short decision-making process.

Aside from recruitment, in terms of purchases and assignments, the burden of proof must be reversed, by systematically authorizing purchases by purchasing cards, and by conducting *ex post* checks on assignments.

Lastly, offering foreigners assistance with their arrival is essential for hosting not only researchers, but also students, who may be non-French speakers: residence permits and contact with the Prefecture; assistance with housing; helping the other spouse to find a job; help with finding school places for children; cultural help. Similar forms of support must also be offered in the event of mobility in France.

Setting Up a National Coordination Strategy

The 3IA institutes will be expected to nurture close, robust relations between them, both in scientific and organizational terms.

Scientifically, this will involve sharing seminars (lecture theatres with high-performing video conferencing tools), organizing discussions via video conferencing, pooling a maximum of teaching aids, tutorials, challenges, etc, as well as internships and sharing their results. A system for sharing expertise between the various 3IA institutes will also need to be set up for comparative assessments of applications (recruitment or other projects launched by one of the 3IA institutes). It will only be possible to guarantee the flexible and swift progress of such exchanges via a shared information hub. Organization of an annual event for taking stock and sharing experiences in person might also be an option, for the financial backers and the public alike, which will also require a certain level of national coordination, scientifically and logistically.

In organizational terms, all of the aforementioned administrative procedures will have to be finalized (and maintained over the long-term), on the basis of specific administrative and legal expertise as regards intellectual property and framework agreements among other things. Obviously, this should only be done once, with a single point of contact being appointed for the whole of the 3IA institute network. The network in itself must be assured a high international profile, as a single gateway for steering enquiries towards the institute most qualified to respond.

Whatever the chosen mechanism for easing the administrative burden may be, it might not be relevant to duplicate it. Instead, we should set up just one instance of it at national level and then enable all of the network's 3IA institutes to use it. This requires smooth movement of information and financial flows, thereby allowing each institute to remain in control of its own budget whilst delegating part of its management.

A coordination structure which has oversight over all of the administrative expertise from the research sphere to the innovation sphere is therefore required. What would be more qualified to house such a structure than a research institute with the very mission of *placing scientific excellence at the service of technological transfer and society* and that already has all of the necessary expertise as well as the necessary culture to deliver?

Beginning the Process with a Call for Proposals

The legitimacy of the 3IA institutes will only be ensured if their setup is supervised by an independent international jury (a prerequisite, even if it is not always enough). It is important not to repeat the same failing highlighted above, where the financing of programme-based research results in a waste of time and energy for researchers, but to build large-scale instruments, in the medium—and long-term (at least seven years, on a renewable basis)[7].

A call for proposals will therefore need to be launched. A two-stage process (short application, then full application), but a single jury, is recommended to avoid unnecessary work on the part of candidates and the jury alike. Care must also be taken over the geographic and thematic spread of all of the accepted institutes, which could be the main purpose of the first selection stage, therefore resulting in a

7. Note that the process being considered here is similar to the one applied for selecting the "Instituts Convergence". The only such institute in the digital sector, DATAIA, incidentally ticks nearly all the same boxes as a 3IA institute.

shortlist of candidates. A second call for proposals could, where applicable, be envisaged where the quality of long applications from the first round is deemed inadequate overall by the jury.

In all cases, each institute must be allowed considerable scope for independent vision, organization and governance, in order to enable original solutions to be developed that are tailored to distinctive local features and specific expertise. Several general conditions will be defined to ensure a coherent whole; the requirement for extensive freedom may allow for the fact that these conditions are not all met by a given institute at the same time.

- The project must be fronted by one or more existing research or higher education institutions, to avoid overcrowding the French research scene. Given the strong commitments to be made in terms of education, it seems to make sense that at least one educational institution (university or school) forms part of the project leadership, willing to provide means if possible (such as premises for the institute to operate without incurring real estate expenses)—even if the principle of independence already mentioned must also apply here. In return, it could be possible for some of its Fellows to be recruited on a permanent basis after a productive work period (at least 5 years). But this could also be a research institute, which may be more inclined to provide support staff for example.
- The project must include a site occupancy plan, preferably in existing premises made available by one of the project leaders. Provision will nevertheless have to be made for an additional budget, for redesigning or even building new premises—when duly justified.
- The project must be fundamentally interdisciplinary.
- The project must play a significant part in enhancing AI education. In particular, all of the researchers who may be associated with the 3IA institute must commit to teaching at least one class a year. The creation of new streams, especially interdisciplinary streams that lead to joint honors, on a project basis (see inset) must be strongly encouraged, and form part of the jury's selection criteria.
- The project must include an industrial affiliate programme, and the commitment of a certain number of local industrial stakeholders concerning their participation, supported by framework agreements bearing on the sharing of intellectual property.

Funding

Significant basic funding must be provided by the public authority when the 3IA institutes are set up. This must enable the institute to operate at a minimum level, with day-to-day running costs covered: the usual overheads; financing for research support staff (assistants for systems engineers and research engineers); and financing for a few research projects (chairs, interdisciplinary projects, guest speakers, etc.)—and, where applicable, as mentioned above, justified real-estate funding.

Over and above the call for projects approach, private funding will also be called on, but on an equal public-private basis, with the State supplementing any private funding. It should be possible for industrial affiliates to get involved at several levels (see above), so that they can benefit from expertise dedicated to their requirements.

On another note, though fairly uncommon in France, the option of a call for corporate sponsors should also be explored, particularly among successful AI startuppers.

The needs in terms of budget of the measures recommended for setting up the 3IA institutes and the RN3IA (and more generally of all of the recommendations in this chapter) are very modest in comparison with the expenses that would need investing in the other sectors, when the return on investment, over the medium—and long-term admittedly, will be immense if the aim of creating a thriving fabric for entrepreneurs is achieved.

Integrating this network in the European AI research area

Via its national coordination, the RN3IA will be able to become the lead correspondent for our European partners to ensure French research in AI connects with the main European AI centers (DFKI and MPI in Germany, Alan Turing Institute in Great Britain, IDSIA and Écoles Polytechniques Fédérales in Switzerland, CWI in the Netherlands, IRIDIA in Belgium, Sapienza Roma and the other robotics and AI research centers in Italy, etc.), not least amid the emergence of a major European AI network, within which this network of French institutes will naturally be expected to represent the French ecosystem. Initially, precedence will be (and already is being) given to the Franco-German partnership. The form that such a network might take is not yet known, but it could be modelled on the EMBL (European Molecular Biology Laboratory), which has been operating successfully since 1974.

There are also hopes for other European partnerships, with the instruments of the H2020 programme[8], like the current public-private partnerships in robotics and Big Data.

But each 3IA institute will, of course, be given scope to forge cooperations with other partners, whether or not European, based on its specific features and the personal relationships of its researchers. We have partnerships in place with our counterparts in Quebec, for example, which would be worth nurturing with respect to the momentum in AI that is gathering at the moment.

A stronger researcher presence alongside French Tech entrepreneurs must also be encouraged at flagship European or international events (Consumer Electronic Show, Web Summit and Founders Lisbon for example).

2. Computing Means for Research

The 3IA institutes must have computing tools at their disposal that can rival the almost unlimited means of the leading private stakeholders. There are, however, a number of needs, of different types, which cover the different stages of research, development and life of products. Indeed, in some areas such as machine learning, the development cycle entails two key stages: learning and inference. The speed and performance of the learning stage depend on the scale of physical means devoted to them, particularly in terms of dedicated processors (mainly GPUs today).

8. *Horizon 2020*, the current EU Research and Innovation programme.

The productivity and effectiveness of research & development is therefore conditional upon the size of their supporting infrastructure. The second stage is inference, which is not nearly so demanding in terms of physical resources.

Several types of work flow in the learning stage do need distinguishing, however: cases which will for the most part rely on a supercomputer adapted to AI (resources typically entailing thousands of GPUs) are fairly rare and only concern part of the research focusing on AI. The vast majority of applications require a much smaller equipment setup (entailing a few dozen GPUs for example). These two complementary requirements should not be confused but kept distinct, therefore, as they are very different in nature and setup:

- A requirement in terms of supercomputer that is designed and dedicated entirely to AI;
- A requirement in terms of "cloud adapted to AI", the beneficiaries of which will include research.

Developing a Supercomputer for the Requirements of Research

The recommendation is to set up a supercomputer designed specifically for artificial intelligence applications (such architecture differs significantly from conventional HPC supercomputers), solely for use by French research, beginning with the members of the 3IA institutes, described above, and their industrial partners under joint projects. By limiting access to free access for research is the only way to achieve an access mechanism that is simple in both administrative terms and in terms of everyday use. For opening up broader access on a paying basis for some users ends up attaching conditions to use of the tool in practice.

Private businesses specializing in the field will have to be called on to design such infrastructure and specifications will have to be drawn up that are specific to AI. Moreover, technical management of the infrastructure may be delegated to an organization like the Key National Infrastructure for Supercomputing (GENCI). This already taps into a range of necessary skills for this purpose (engineers, administrators and so on). These will nevertheless need extending to the specifics of AI (which, and if we might stress this point, are not the same as those of the traditional HPC), so as to be able to integrate the constantly upgraded high-performance equipment, software stacks updated with the latest algorithmic advances and data storage capacities for guaranteeing secure access (modelled, for example, on Teralab) so that the national private partners feel confident entrusting it with their data for research purposes.

Note that this recommendation is fairly similar to the one outlined by the working group of the alliance Allistène for a HPDA infrastructure dubbed GENIAL (which stands for Key National Infrastructure for Artificial Intelligence), the key points of which we will be able to set out again word for word, not least as regards the importance of guaranteeing the most flexible type of access possible, and the mechanism proposed for achieving this, of an open-access section with *ex-post* checks and another section where access is reserved.

Negotiating a Pass in a Private Cloud for Research

However, this supercomputer will not cater to all research requirements bearing on cloud computing for quick and often complex testing in terms of hardware or software configuration, as might be found on the main Private Cloud stakeholders (AWS, azure, etc.).

For the AI cloud requirements, setting up an access package to a cloud adapted to AI is recommended. This package (including computing time and storage space at least adaptable) could be allocated based on the teams and their needs. The aim here is also to maintain maximum flexibility in terms of access to the infrastructure. Given the scale of such a project, this AI cloud will be expected to develop at European level through a privileged partnership with a specialist European stakeholder in the field. Using it for research would have a twofold advantage, firstly for research, but also for "initiating European industrial momentum with regard to AI", an idea explored in Part 1 of this report.

3. Enhancing the Appeal of Careers in Public Research

Improving the Profile of Professor and Researcher Careers, Particularly in the Early Stages

Unrealistic though it would be to try and compete with the financial offerings of the GAFAMs, the pay gap is currently so large that it is putting off young graduates—even those most attached to public research and the common good. At least doubling starting salaries is a necessity, otherwise we risk seeing the arrival of young graduates ready to invest in higher education and academic research dry up completely.

Doubling starting salaries is a necessity, otherwise we risk seeing the arrival of young graduates ready to invest in higher education and academic research dry up completely

Increasing France's Appeal in the Eyes of Expatriate or Foreign Talent

The status of permanent researcher is one of the last surviving advantages that French research has over all of its competitors, whether we are talking about the major digital industries or public institutions abroad (Europe included), and it is often the only argument to counter the eye-popping salaries offered. But for a senior researcher, entering the French system is a very big step financially speaking (buying back pension contributions over a long period), which could partly be accomplished with *ad hoc* assistance, something which seems nigh on impossible today.

Other measures of ad hoc assistance could also be considered on a case-by-case basis (see the recommendations concerning 3IA institutes that are not intended to be limited to the RN3IA.). Similar assistance should also be offered for national mobility schemes, so as not to put at an indirect disadvantage researchers who already have permanent positions in France, thereby paradoxically creating

incentives to emigrate, even if the aim is to improve conditions upon return. Once again, this measure is not unique to AI, but is now more necessary than ever given the competition from the GAFAMs.

Training More High-level Specialists in AI

On the one hand, there is a blatant shortage of high-level graduates in AI, a problem which is only going to get worse in the future if all the predictions are to be believed (see above)—even if it is also necessary to train intermediate-level scientists too (with two or three years of higher education). The numbers of students taking AI Master's and doctoral studies therefore need to be increased drastically.

On the other, the motivation of professors is also tied in with the possibility of leading a research group—which begins with the possibility of recruiting master's and doctoral students easily and quickly. And yet even though some of the aforementioned recommendations will, if adopted, result in an automatic rise in the number of doctoral students, such efforts should not be limited to these schemes. Master's and thesis grants, channeled towards AI and possibly based on a facility & administrative budget, must be offered in all doctoral schools, awarded following consultation of a specific local committee, for example including researchers affiliated to the nearest 3IA institute. Note that such a quota-based financing mechanism already exists in contexts of funding in addition to doctoral school grants, provided by several entities—be they local or regional.

4. Stepping Up Interaction Between Academia and Industry

A great many schemes have been set up in recent decades to try and resolve the lack of transfer often observed between academia and industry in France. We recommend rounding these off beyond the 3IA institutes, at the individual level of researchers, by facilitating their partial involvement in industry. Such measures—which are in no way specific to AI and could be applied across the academic spectrum—are now essential in the short-term on account of the recent uptick in the brain drain towards private AI stakeholders.

Encouraging Shared Work Between Academia and Industry

Permanent research and higher education staff must be encouraged to share their time between academia and industry—up to 50% for example, by authorizing supplementary pay at the competitive level of the private sector (this would mean abolishing the rule of not doubling pay). Regarding professors, where someone switches to part-time, they would be obliged to find a replacement to teach their classes (from their research group for example). The university board, which now decides whether or not to allow the buying back of teaching hours, should not be able to refuse where funding is ensured, for example by the private employer, and teaching is ensured, for example by experienced postdoctoral researchers.

Taking Account of Stints Spent Working in Industry when Reconstructing Career Paths

In order to encourage to-ing and fro-ing between academia and industry, care must be taken that periods spent in the private sector are not penalizing for professors, whether in terms of career progression (reconstructing one's career path, pension contributions, etc.) or before different recruitment and promotion panels.

Appointing AI Researchers on Boards of Directors

As part of a reform of the State-as-a-shareholder policy, and both to involve researchers more in industry and bring AI culture into boardrooms, researchers specializing in AI could be appointed as State administrators within companies making up the portfolio of the Agence des participations de l'État (special agency of the Government of France managing the State's holdings)—as is practiced on a much more regular basis by our German neighbors.

Solving the Problem of Sharing Intellectual Property

The best collaborative projects between academia and industry inevitably encounter problems regarding the sharing of intellectual property (IP). But is there another solution than to proceed on a case-by-case basis? One possibility would be to seek out agreements to share out the value created and IP rights fairly. As suggested in the plans for a European intergovernmental institute, another option would be to simply abandon IP to industrial partners, as long as they are clearly and unambiguously European. A third way, as practiced in Germany, entails clearly separating out the applicative and fundamental spheres, with the IP shared out accordingly. What strikes as essential, whatever approach is adopted, is that this does not become a stumbling block or factor in holding up common projects.

Encouraging Researchers to Create Startups

A civil servant wishing to create a startup on the basis of his or her research findings must also be given encouragement, for example by providing financial assistance for 2–3 years in the early stages and/or free hosting in an incubator for startups—as well as systematically granted leave of absence (for the same duration). The IP should also be subject to an agreement between the institution which produced the research and the startup (see above). There are several possible models, from holding shares (this is prohibited for public institutes today) to revenue sharing agreements, which still need looking into, and could be proposed to entrepreneurs. One of the key points is to keep up with the very quick pace of innovation in the AI sector and to have correspondents for the legal and financial setups who understand what is at stake in these situations. Examples now exist, such as the venture capital fund, Partech Fund, launched by the Université Paris-Saclay.

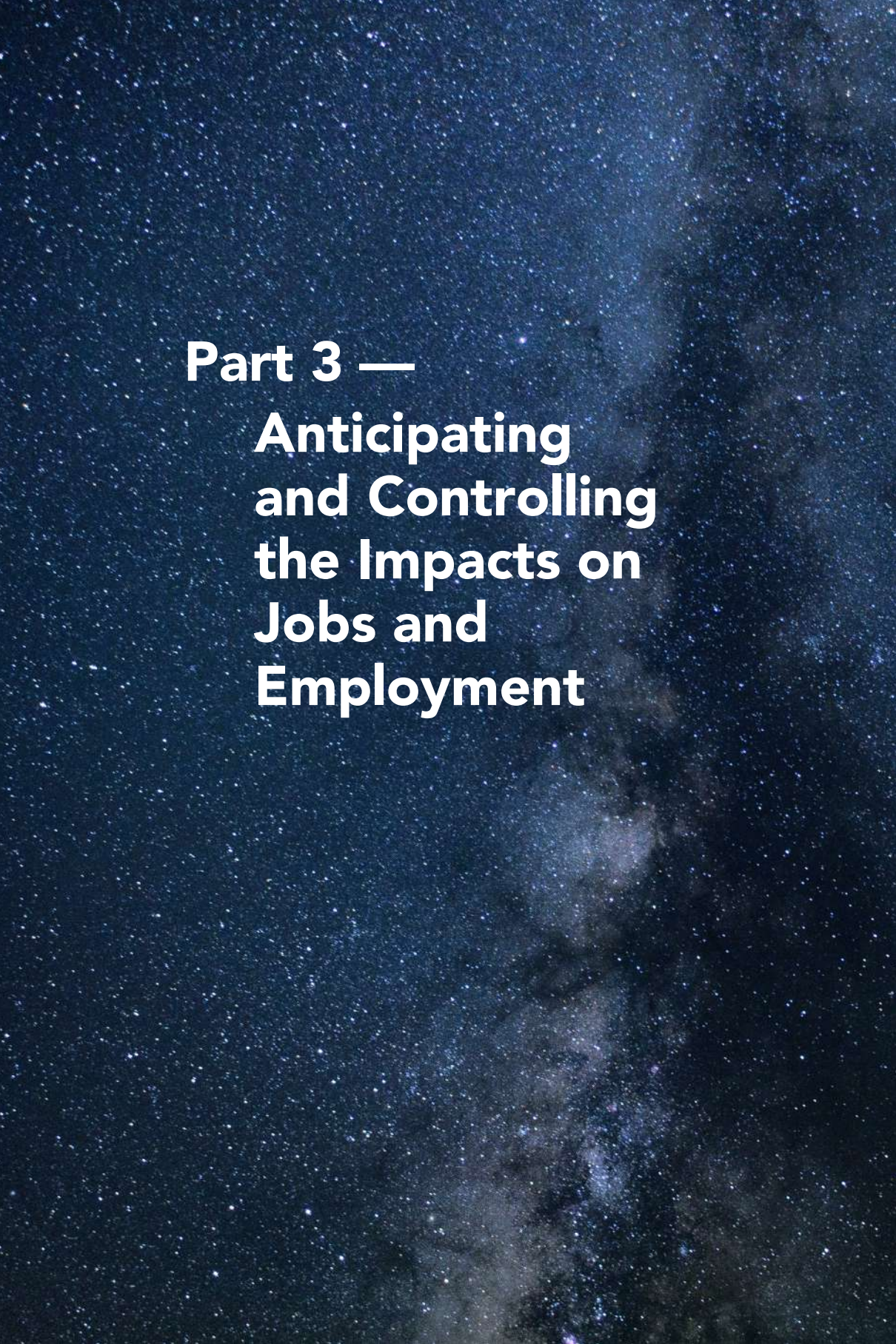
Encouraging the Setup of Industrial Chairs Through Co-Financing

Chairs (for example the initiatives of excellence [IdEx] chairs of excellence) tend to be associated with chairholders, involving funding for a limited period of time (5 years) and including an earnings supplement and facilities & administrative budget for a small team of doctoral students and postdoctoral researchers who are starting up a dedicated research team. A fairly widespread variant is the industrial chair, fully financed by an industrial stakeholder (a multinational, given the level of investment), including pay for the chairholder.

The aim here is to encourage the creation of such chairs; the chair must be set up for at least 5 years, and financing must make it possible to typically recruit one or two doctoral students and, some years, postdoctoral researchers. Public funding may supplement the overheads for the industrial chairs as follows:

- The candidate already holds a position in higher education or research. In this case, the chair must finance an earnings supplement (but it is important, here again, to line up with the market as defined by the GAFAMs), plus a small team.

The candidate does not (yet) hold a position—this is typically the case for an expatriate we are trying to bring back into the French research fold. The public funding of the chair after 3 years could therefore depend on the chairholder obtaining a permanent position.



Part 3 — Anticipating and Controlling the Impacts on Jobs and Employment

Part 3 — Anticipating and Controlling the Impacts on Jobs and Employment

The professional world is not yet sufficiently prepared for the unprecedented changes looming over it. Strictly speaking, the development of artificial intelligence is not yet considered as a fourth Industrial Revolution, but what is becoming increasingly clear is that it has an impact on most occupations and organizational procedures. For its development will enable a great many tasks to be automated. We are therefore entering a major period of technological transition, and this is sparking significant concerns: history tells us that it has not all been plain sailing during previous transitions and that the political readjustment processes have at times been brutal, often to the detriment of populations who were already the most vulnerable.

Whilst it is important to distinguish between automation, artificial intelligence and robotization, it is difficult to know to what extent these three phenomena are each responsible for changing work practices, and they must, therefore, be taken as a complex whole if we want to analyze their effects.

What are the Forecasts?

In light of the sheer scale of the phenomenon, the temptation of getting carried away is great. It is stoking apocalyptic forecasts on the mass destruction of jobs: the first study on the subject to have really hit the headlines, by Frey and Osborne from the University of Oxford, predicted that 47% of total employment in the United States was at risk of vanishing over the next two decades[1]. In France, one consultancy firm, Roland Berger, estimated that 42% of jobs were under threat within a similar timeframe. The most recent study to have been published on the subject by the Employment Advisory Council, attached to the Prime Minister, adopted a different approach and predicted that 10% of jobs were at risk of disappearing, but that 50% of jobs would potentially be automated at more than 50%. The bottom line is that the scale of change we are talking about here is far-reaching and must be matched by collective planning.

The professional world is not yet sufficiently prepared for the unprecedented changes looming over it

And there are several ways we could plan in light of these forecasts. First, it may be tempting to deny that there is a problem, maintaining that new jobs we know nothing about yet will be created in high numbers and that, owing to the interplay of prices and demand, individuals will naturally move into new roles. But we now know that things are not so simple, that the human and social costs are often very high during economic transitions and that the simple pressures of the market are seldom enough to distribute the supply of work in the best way possible. The risks of higher unemployment and inequality may be high. On the other hand, we could choose to adopt a gloom-mongering approach, often for the sake of better defending hidden interests: when it looks like disaster is all but upon us, it is easy to forget certain principles which have until now

1. Future of Employment, C. Frey and M. Osborne, Oxford.

guided our collective practices. To avoid both of these pitfalls, we need to address this challenge head-on, without succumbing to panic, in spite of the major uncertainties weighing down on us. Incidentally, these uncertainties are tied in with the theoretical oppositions that are gripping the world of economic research on these subjects.

Thinking in Terms of Complementarity

Tackling the problem head-on first of all means recognizing that what we are dealing with here is a large-scale upheaval of the job market, the distribution between human tasks and machine tasks and value production modes. And that, first and foremost, a great many tasks are going to be automatable, no matter how quickly this automation takes place. In light of what may now be considered an inevitability, in the medium-term we need to be pushing on with discussions on alternative modes for producing and redistributing value. The priority must be on developing the means for effective complementarity between human tasks and machine tasks.

Which Tasks are Automatable?

Obviously, it is difficult to provide any precise criteria. In 2003, three academics (Autor, Levy and Murnane) defined a broad criterion in a formative article[2]: it is the repetitive nature of a task that makes it susceptible to automation, i.e. if it follows clear rules, as opposed to the ability to solve problems in an autonomous manner. There is nevertheless some debate over this definition as it remains too general and, above all, its scope is very flexible: should driving a vehicle be considered a routine task in this case? A fuller set of criteria therefore needs adopting to consider which tasks will be liable to automation. In its commendable 2017 report on automation, the Employment Advisory Council (COE) defined four main criteria for determining whether a task can be readily automated:

- no flexibility: the work pace is set by a machine speed and the task is regulated by hourly production standards and involves continually repeating the same series of movements and operations
- no capacity for adaptation: there is no need to interrupt an ongoing task to carry out another unscheduled one, and the task entails a strict application of orders or instructions;
- no capacity for solving problems: when an abnormal situation arises, the worker calls in other people to solve the problem;
- no social interaction: contact with the public is limited and the work pace is not set by outside demand.

There is still room for further fine-tuning these criteria, which do not fully determine the propensity of a task to be automated. Other factors obviously come into the equation too, such as the state of technology, as well as the social acceptability of automation or its overall cost—it is not necessarily more advantageous to automate a task, even from a purely economic point of view. What is more, these criteria are

2. Author, Levy, Murnane, *The skill content of technological change*, 2003.

not set in stone: the division between the non-creative machine and the creative human is ever less clear-cut, which makes it difficult to attribute the shared fields. No occupation can assume itself safe from change because of these criteria.

That said, they do throw down the broad outlines of the bottlenecks of automation and give an idea of the skills and abilities that need to be developed. What are these skills and abilities?

Put simply, they fall into four categories:

- cross-cutting cognitive skills (understanding language and numbers, ability to solve problems, etc.);
- creative abilities;
- social and situational skills (teamwork, independence, etc.);
- precision abilities relating to perception and handling, which should not be overlooked, such as manual dexterity for example.

Of course, these skills are not distributed equally among the working population and, although automation may affect all occupations across the board—even highly skilled ones which do not rely on cross-cutting skills—it will have a greater impact on low-skilled workers. For the COE, the most vulnerable jobs which are proportionally more represented relative to their share in total employment are usually manual, low-skilled occupations, particularly in industry, such as unskilled workers in the process industries, unskilled mechanical handling workers, unskilled workers in construction finishings, cleaning staff, unskilled mechanics, cashiers, etc.

Breakdown of the most "exposed" jobs: the main occupations in terms of volume (where the automation index is at least 0.7), source: COE

	Number of jobs exposed	% jobs exposed
Cleaning staff	320,215	21.05%
Skilled workers in the process industries	95,545	6.28%
Unskilled mechanical handling workers	85,965	5.65%
Unskilled workers in the process industries	83,304	5.48%
Domestic helpers and home help	76,198	5.01%
Cooks	70,306	4.62%
Skilled mechanical handling workers	62,047	4.08%
Market gardeners, gardeners, vinegrowers	49,875	3.28%

Drivers	48,786	3.21%
Skilled workers in structural construction	48,455	3.19%
Unskilled workers in structural construction, civil engineering works, concrete and extraction	46,517	3.06%
Employees and supervisors in hospitality and catering	44,362	2.92%
Household employees	43,880	2.89%
Cashiers, employees in miscellaneous services	43,770	2.88%
Skilled workers in construction finishings	37,156	2.44%
Unskilled workers in construction finishings	34,226	2.25%
Skilled mechanics	32,899	2.16%
Crop farmers, livestock farmers, foresters and lumberjacks	31,985	2.10%
Unskilled mechanics	31,732	2.09%
Other	202,628	13.32%

These changes could also have a direct impact on low-skilled and intermediate service occupations, which means that "white-collar" workers are also largely concerned. A certain number of studies corroborate the fact that automation is only going to further polarize the job market. In its article entitled "Schumpeter et les robots : le cas de la France", Patrick Artus, by drawing on recent economic research[3], thus defends the argument of the bipolarization of the job market: where automation would lead to the creation of mainly highly skilled jobs on the one hand, and basic jobs in the domestic services on the other. Opinions are still divided on this argument, however—the COE believes that it is only applicable to a polarization geared towards the highest-skilled jobs. There is also uncertainty over the question of which new jobs will be made possible by artificial intelligence (whether directly or indirectly).

Indeed, the macroeconomic consequences of automation on the distribution of jobs and work will partly be determined by the objectives set collectively and by the

3. It is generally confirmed that robotization, as we have just described, leads to polarization of the job market. (Author-Lévy-Murname (2003); Goos-Manning (2007); Michaels, Natras, Van Reenen (2014); Author-Dorn (2013)). But other publications suggest that robotization does destroy jobs overall (Ford (2015); World Bank (2016); Arntz-Gregory- Zierahn (2016); Acemoglu-Restrero (2017); Graetz-Michaels (2015)).

means that will be put in place to meet these objectives. Thinking in terms of complementarity resonates with this ambition to sketch out what the changes wrought on the workplace by automation might look like. As such, rather than forming specific predictions on occupations that are yet to be invented, it would be more worthwhile to begin with what we already know—most occupations will change dramatically, especially low-skilled occupations—to guide these changes.

De-Automating Humans

Guiding these changes does not therefore simply mean adapting the workforce to the new jobs that will be created, complementing machines. Since, in theory, the vast majority of jobs that will be created will involve working with a machine, whether or not it is a highly skilled job or creative job, what we should instead be doing is setting objectives as regards this complementarity.

General objectives in terms of the structure of the job market to begin with: avoiding excessive polarization of the job market and soaring inequality could be the first such objective. Then objectives in terms of ways of working with machines: what, indeed, does it mean to complement machines? This can take several forms, and they are not all desirable: following orders from artificial intelligence, losing control over processes, delegating decisions to machines are all examples of complementary working which, at individual and collective level alike, are bound to result in suffering at work. It must therefore be made clear that not all forms of complementarity are desirable, and that an enabling form should be developed.

For the automation of tasks and occupations could represent a historic opportunity for de-automating human work

What this basically means is developing complementary human skills to artificial intelligence on a massive scale; the fact that some types of complementarity destroy human abilities must be underscored, and we need to realize that we must work together to set the scene for developing a form of complementarity where human abilities can be enhanced.

For the automation of tasks and occupations could represent a historic opportunity for de-automating human work: it enables us to hone our uniquely human skills (creativity, manual dexterity, abstract thinking, problem-solving). We must turn artificial intelligence to our advantage to develop the abilities of each and every one of us: the opportunity is there for the taking.

Preparing for the Transition

Setting political objectives must go hand-in-hand with scaling up our abilities to predict and understand the phenomena at work. For whilst it may be possible to estimate the major macroeconomic effects that the introduction of AI technologies will trigger, it goes without saying that much remains uncertain, that the lines are moving constantly and that the forecasts constantly need clarifying. If we are to address the problem head-on, we therefore need to begin by fine-tuning our understanding of these phenomena: we have to find ways of more clearly grasping what is happening, of maintaining our abilities to look ahead over time and of

moving away from our current fragmented, short-sighted approach to understanding what might lie before us, to see the bigger picture.

Then, it is necessary to prepare for the transitions, at individual and collective level alike. Managing to identify the main risks and skills that we need to develop to work with machines is only the first step. The second step, and the most complex, is to enable the massive transition of individual skills and abilities. Those of people who are already in jobs of course, but also those of newcomers to the job market.

Current learning paths, whether they involve vocational training or initial education, are simply not equipped to see this transition through smoothly

Two main sectors therefore need adapting. First of all, education: current learning paths, whether they involve vocational training or initial education, are simply not equipped to see this transition through smoothly. Consideration of cross-cutting skills, learning creative skills, new teaching practices—all of this is often still sorely lacking in syllabuses. We are on the brink of a major transformation to the education and training sector, one which will be necessary if we are to embrace the development and spread of artificial intelligence. This transformation will first involve testing out, setting up structures that are akin to "pioneering experts" or outposts, designed to do things differently and to test out.

Then, the public employment and vocational training policy schemes: these schemes are not sufficiently factoring in the need to urgently and specifically target certain jobs and individual profiles and, at the same time, to test out. The point here is not, therefore, to overhaul existing schemes, but to provide for scope for testing out, within the overall system.

Training AI Specialists

This is one of the top priorities: the requirements in terms of people trained in artificial intelligence have not yet been met by existing specialists and these skills gaps look likely to get wider still. The existing very high-level training programmes will not be enough. Interfaces also need to be designed between artificial intelligence courses and other subjects (life sciences, social sciences, business skills, etc.).

1. Anticipating the Impacts on Employment and Testing Out

Setting Up a Public Think Tank for Transforming Work

This is the first requirement: ensure that the ability to anticipate is sustainable and ongoing, and above all tied in with public policy. Today, foresight studies, which are published at more or less regular intervals by public or private institutions with this role, no longer match up with the day-to-day delivery of public labor, employment and training policies. The effects of this mismatch are entirely negative: the publication of studies sparks collective debates which, whilst often enthusiastic, do not really lead anywhere, while concrete public policy, which has to address both the

need to manage the day-to-day and the complexity of the delivery channels, is only subjected to minor adjustments and struggles to properly take on board the findings of the foresight exercises that are carried out. This question is relevant to all State functions across the board: how can the need to anticipate often radical change in the future be reconciled with the need to ensure the smooth running of all public services on a day-to-day basis? But it is more pressing for some sectors, including that of employment and training policies. Indeed, changes can take place extremely quickly and the public policymaking processes are complex and difficult to intervene on. Vocational training alone, for example, represents €32bn a year, with a whole host of funding channels and different stakeholders.

We therefore need to create a space where forward-looking capacities, macroeconomic forecasts and analysis of changing uses can be linked with concrete testing capacities that are tied in with measures aimed at certain categories of workers. A permanent structure could therefore be set up, with a "pioneering" role within public vocational training and employment policies. A close link will have to be upheld with sectoral observatories.

The missions

This structure could have several missions:

A foresight role

A traditional foresight role aimed at producing annual studies on the automation of tasks, the most heavily affected occupations, new jobs, etc. Every year, indicators relating to the automation of occupations (automation criteria, the occupations most directly concerned by potential automation) could be updated. This structure is expected to adopt an interdisciplinary approach in this role. A role more specifically dedicated to leading a thought process and to producing analyses on the means for achieving complementarity between humans and machines and the new skills that have become necessary. This thought process will have to bear on all occupations, not least to enable sufficient consideration of cross-cutting skills. Then, an experimental role. These tests must first and foremost address a primary challenge: supporting the professional transitions of employed persons. Admittedly this objective forms part of all vocational (and even initial) training schemes, but room must nevertheless be made for testing out. This is because some occupations, sectors and local areas are going to feel the effects of the technological changes afoot more than others. Trials will need to be performed, in a preventive mindset, to test out transition schemes in these jobs and local areas. The think tank could therefore make targeted changes to existing national schemes (individual training account/CPF or career review guidance/CEP for example), but also issue calls for proposals and support local and national pilot schemes. These trials could also be relevant in the context of more advanced thinking on the reinvention of models for creating and distributing value at the dawning of this new automation era. Obviously, they would be guided by—and would inform, in return—the structure's theoretical work.

A debate leadership role

A role leading open debates, at national and international level, on the changing workplace in the age of automation. This role must take shape, beyond the theoretical aspect, via the establishment and/or networking of living discussion on the future of work, which take the form of think tanks on the new occupations and uses of tomorrow.

The Arbeitviernull platform in Germany

In connection with the congress entitled "Work 4.0", organized on 22 April 2015 in Berlin, the Minister of Labor and Social Affairs (Andrea Nahles) unveiled a "Green paper on Work 4.0" which outlines the main challenges and questions raised by the digital revolution underway. This document is intended to inform another debate on the future of the workplace with all of the stakeholders concerned (economic, political and social decision-makers, experts and citizens). A website (www.arbeitenviernull.de) provides a platform for a broad and open dialogue with all of the stakeholders.

Constitution of this structure

To operate most effectively, it seems necessary for this structure to conduct discussions on an inter-sectoral basis and to take a tripartite approach (State, trade unions, local authorities), as well as drawing on the foresight capacities of academia or specialist public institutions. The link with technical experts will be paramount if it is to succeed in its missions. This structure may be encouraged to play a part in international debates, via the nomination of foreign experts for example, or via participation in a global policy network.

Experiments and funding

The experimental role strikes as being central to this structure's purpose, hence why calling it a “lab” in French (think tank) appears justified. The testing ground is extensive.

The new learning methods

Trials could bear on the new learning methods and on the way to organize the range of vocational training programmes to best cater to needs that are difficult to address because they are not directly related to business-specific skills: creativity, cross-cutting skills, general cognitive skills, etc. In this respect, in keeping with the creativity plan (see below), this structure shall have to help finance calls for proposals for the attention of the vocational and initial training ecosystem. It could also help bring about proofs of concepts (POC) and demonstrators, tailored to specific professional transitions, in connection with the think tanks on the occupations of tomorrow (see below).

The Coding Boot Camps in Israel

The Israeli Government has set up a "coding boot camps" programme, implemented by the Innovation Authority. The overarching objective of this programme is to encourage private initiatives aimed at the vocational training of seniors, minorities or a career change towards hi-tech occupations—particularly in computer programming. In Israel, national service also plays a key role in fostering encounters, ecosystem rationale and work on defined projects, for the benefit of the nation as a whole.

Lastly, trials aimed at expediting links between initial training and continuing professional development could be launched: a reminder clause could be introduced, which would encourage certain initial training courses, identified as having learning outcomes at risk of being automatable in the near future, to get back in touch with their alumni after a certain number of years to offer them modules to develop their skills.

Focusing Certain Schemes on Jobs that are at Higher Risk of Automation

The experimental approach could serve to bring about different ways of thinking to the mindsets that are currently adopted in vocational training. Current schemes are mainly to be taken up at the initiative of employees, in a mindset of individual accountability. Given the potentially lightning-quick, exponential even, nature of these changes, for the existing general schemes it seems difficult to be able to cover all situations, enabling both consideration of the needs of the whole population and the need to act in a targeted and urgent manner. What is more, individuals are not all equally equipped in the face of their changing jobs to be able to adapt and build alternative career paths.

In this regard, trials could be carried out to develop schemes for specific target groups, whose jobs are considered to be at the highest risk of automation and who will struggle to make the necessary career changes themselves. The aim is therefore to move partly away from the sole view that individuals are accountable for their own professional transitions. Accordingly, a trial could be conducted to transform the career review guidance (CEP) or to come up with a new type of individual training leave.

Examples of trials

The CEP: the resources of the career review guidance (CEP) could be proactively geared towards the individuals working in these occupations, without waiting for them to initiate the process themselves. This free, personalized support scheme, available to anyone wishing to review their professional situation, is a key feature of the vocational reform, and yet it remains vastly underused today. Indeed, to work effectively, mechanisms that make individuals accountable for their own training choices, via the individual training account, are reliant on the latter's abilities to

make the right training choices. In the context of these trials on the occupations under threat of automation, the structures organizing the CEP (Pôle Emploi, mission locale, accredited joint collecting organization for the individual training leave/OPACIF) could be extended to include other stakeholders, in a networking perspective. In this way, a community of expert advisers (volunteers, professionals in the sector, etc.) could be set up along with platforms for discussions and sharing experiences in a more flexible way;

The CIF: There are debates over the merits of the individual training leave (CIF), which enables an employee to train over long periods of time via funding from the OPACIF, insofar as it appears to be able to double up with the individual training account (CPF). It could be upheld but turned into a tool that facilitates major career changes. It would thus become leave for embarking on a career change, with a target list of occupations that are identified to be at-risk;

Supplementation of CPF points: the trials could include supplementing the points in the individual training account (CPF);

EDECs (Commitments to Develop Employment and Skills) could be funded by this structure for identified local areas or businesses.

Job rotation in Denmark

Job rotation involves businesses, employees and jobseekers. After receiving training, the latter temporarily fill the position left vacant by an employee who embarks on long-term training. This choice guarantees that the position is not left vacant, whilst providing the jobseeker with work experience that s/he will then be able to put on his/her CV. The scheme is working well. According to the figures released, this method enables 6 in 10 jobseekers taking part in this scheme to find a new job.

Trials in local areas

Finally, these trials could also finance initiatives in specific local areas, which are trying to invent new models to deal with the automation of occupations for example. The debate on the basic income, championed by different stakeholders with diverging—or altogether conflicting—political visions, or the debate on commons and methods that contribute to wealth creation could thus gain from the setup of concrete trials. Several trials are beginning to emerge in local areas, aimed at rolling out jobs and occupation transition models: the Aquitaine region has pledged to test out a basic income, zero unemployment areas are aiming to hire long-term jobseekers on permanent contracts, using funding from grants, and the intermunicipality Plaine Commune has launched plans for a contributory income in its region… Such a structure could fund and/or support other similar trials and organize the public sharing of feedback.

Setting up local branches of think tanks on the occupations of tomorrow

Think tanks could be set up, at the local level, as local branches of the national structure. They will have to provide open places where individuals are able to think collectively about how their occupations are going to change. Employees, apprentices and students will able to decide to come and train in these think tanks through research, experimentation, trial and error and the sharing of new approaches and of practices developed in France or abroad. These laboratories of sorts, on the occupations of tomorrow, could be places for research-based training and for enhancing the status of apprenticeship tutors—turning them into key players in the vocational training system. They will have to give employment stakeholders the opportunity to speed up the pace at which they make their own professional changes.

These structures could specialize in one or more occupations in particular, depending on the way in which the employment catchment area in their local areas is organized. This recommendation reflects the same thinking as the report on the learning society, overseen by François Taddéi.

The Cap Digital cluster's Edfab

Cap Digital considers it paramount that close attention be paid to the development of human capital, skills and talents if the digital transition of businesses and organizations is to succeed. The "Edfab" initiative fits in with this ambition. It is a new venue devoted to innovations in the training, education and transformation of occupation sector, set up in the Maison des Sciences de l'Homme, Saint-Denis (Paris region). Four main thrusts underpin Edfab's activities: Find out, Meet, Learn and Experiment. EdFab particularly organizes workshops for getting to grips with data and artificial intelligence.

Financing Experiments

Under the Big Investment Plan 2018–2022, €15bn have been ring-fenced for vocational training, primarily for the benefit of low-skilled jobseekers and low-skilled young NEETs (not in employment, education or training). This budget forms the skills investment plan (PIC).

It seems necessary that part of the PIC funds be coordinated by the think tank for transforming employment, in order to nurture a targeted preventive and above all experimental mindset. The trials would be aimed at people still in employment, but who need to undergo some sort of professional transition, which is not currently covered by the PIC.

2. Developing Complementarity Within Organizations and Regulating Working Conditions

Complementarity between humans and machines is expected to be on an upward trend, not least owing to the potential gains in productivity. But this complementarity

may come in a wide range of forms: in some cases, it may end up enabling the development of general cognitive skills and creativity, thereby enhancing individual skills. In others, working in collaboration with a machine may increase the routine nature of tasks and reduce capacity for personal initiative and thinking, at times under the guise of improving working conditions. Although a certain form of automation may evidently make life easier for employees, the longer-term risks nevertheless need highlighting. Major retail logistics warehouses provide a typical example of this ambiguity: for the automation of processes may lead to employees solely following orders from a machine.

This example, which is just one among others, shows that relying solely on the microeconomic choices of businesses in terms of how to implement artificial intelligence technologies within them can give rise to less than optimum situations.

It therefore seems necessary, for an optimum picture to emerge regarding the use of artificial intelligence in conjunction with human intelligence, that an enabling form of complementarity develops within organizations. There must be a broad dialogue on the definition of this form, first and foremost among employees. The aim will particularly be to reconcile the desire to build individuals' room for maneuver and the potentially negative effects of calls for creativity, which can be problematic for many individuals.

Developing a Positive Complementarity Indicator for Businesses

To go about this, complementarity must first be defined, for example by developing an indicator, with all of the stakeholders on board (trade unions, State, researchers, etc.) and by producing information and documentation, for the attention of businesses and social partners. The structure discussed above could particularly take on this role.

Next, the ways in which the collective choice of positive complementarity, enabling complementarity, could be made, must also be defined. There could be several ways of doing this.

Fully Including Digital Transformation on the Social Dialogue Agenda

This firstly requires: An overall approach to digital transformation in social dialogue

Beyond the tool designers themselves, users must also be involved at all reporting levels, and across all roles—particularly the human resource department by making provision, as far upstream as possible, not least via strategic workforce planning (GPEC/GEPP), for co-construction and co-innovation in terms of space and time.

On the one hand, since employees are in the best position for properly gauging the dimensions of their work activity, taking their experience into account makes it possible to design more effective tools. Indeed, by questioning the future conditions for work performance we can identify, from what already exists, what could be improved.

On the other hand, this must enable them to progress from an imposed process to a projective process, guaranteeing a better grasp of the future work situation. In light

of the fact that organizational changes have been singled out as the number one stress factor in the European Agency for Safety and Health and Work survey, employee participation would help to reduce this risk.

In order to take advantage of the contribution that each and every employee could make, it is important to decide on the places, times and common rules together, in advance. This is necessary to ensure that the discussions bear on the actual *subjects* of transformation, and do not just deal, ex-post, with the sometimes unforeseen *consequences* of transformation. Compulsory collective bargaining in companies is one of the frameworks in which such discussions may be held.

Insofar as the strategy is taking on global proportions, the approach based upon social dialogue must also be widened to encompass the myriad dimensions of the digital transformation, whether they concern volume, structure or location of jobs, skills, organization or work situations.

Revising the content of compulsory collective bargaining

The French Labor Code particularly provides for two types of compulsory collective bargaining, one annual on equality and quality of life in the workplace (Art L. 2242-17 of the Labor Code) and the other on strategic workforce planning (now referred to as "*gestion des emplois et des parcours professionnels*" in French—management of jobs and career paths, Art L. 2242-20). This must take place at least every 3 years in companies with 300 employees and over.

The contents of such negotiations could be revised to factor in the introduction of new technology and the digital transformation of companies, in terms of adapting skills and of complementarity between humans and machines.

These discussions could usefully inform all levels of social dialogue (at enterprise, sectors and national level).

Launching a Legislative Reform of Working Conditions as Increasing Automation Beckons

More generally, complementarity raises the question of what framework should govern working conditions in this digital age.

The legislation concerning working conditions has primarily been written with the working methods of the industrial age in mind and, despite one or two areas of progress (not least concerning the right to disconnect), consideration of a series of new risks and situations, related to the development of digital technology in companies, seems to have been more problematic. The result is that, for a large number of businesses and individuals, the framework governing working conditions is too rigid and, in some cases, simply unsuited to new working arrangements (mobility, flexibility, etc.), and does not take the specific new risks on board. The arduous conditions account has thus been set up to bear the strenuous, demanding nature of work in

Complementarity raises the question of what framework should govern working conditions in this digital age

mind. The point is not to call this approach into question, as it is applicable in a wide range of situations.

But in light of the need to provide a framework for new work situations, a specific legislative reform is being urged: exclusive following of a machine's instructions, no possibility of discussing with colleagues without going through a machine interface, etc. Complementarity must take center-stage in this reform of the framework underpinning working conditions.

3. Setting in a Motion an Overhaul of Initial Training and Continuing Professional Development to Make Room for Learning Creative Skills

The rise of artificial intelligence is calling for an overhaul of not just training methods but also training content.

For it is bringing with it new requirements in terms of individual skills, requirements which should not be regarded solely as restrictive since they are also giving rise to new possibilities for freeing up human work from tasks that are overly repetitive or where workers have no say or control. To ensure the best form of complementarity between humans and artificial intelligence, cross-cutting cognitive, soft and creative skills must be developed.

Cross-cutting cognitive skills form the cornerstone of the learning outcomes of the French education system. So although it nevertheless appears important here to underscore the importance, across all curricula, of teaching the foundation subjects to learn how to reason and understand the world in a complex way, there does seem to be widespread acknowledgement of this requirement. We should, for all that, point out that this partly challenges a strict "matching" mindset which has sometimes prevailed in public rhetoric, whereby education courses are strictly tailored to the needs of an employment catchment area—and which, today, may well be training individuals whose jobs will quite possibly be automated just a few years down the line. There is a high likelihood, therefore, that "matching" a training programme too precisely to suit vacant positions would end up leading to a more general "mismatch" of individuals as regards a constantly evolving job market.

That said, solely focusing on general cognitive skills has often resulted in learning of another skill being overlooked: one which is becoming increasingly important today and could even be considered *the* key skill in a constantly changing world. Creativity. This is why it is vital that a reform to the French education system places emphasis on the importance of creative skills.

But this will require a radical change in teaching methods: for creativity should not be viewed as a "personal development" skill solely of use to the "artistic" subjects or associated/open subjects, as is commonly the case today. It is the very way in which the foundation subjects are taught that must be reviewed so that new methods can be geared more towards developing critical thinking and cooperation. There is evidently a wealth of studies and reports available on this point, and thoughts on new teaching methods, right from infant school level, have been expanded on for nearly 30 years now. Conflicting demands are often made of teachers in this regard, in often difficult working conditions: they not only need to follow very precise

syllabuses, with constant statistical assessment required of students' progress and precisely defined monitoring indicators but at the same time they are called to adopt highly innovative teaching methods, which are often very demanding in terms of time, resources and freedom.

This report does not in any way intend to add to their demands, from the outside. On the contrary, teachers' capacities for initiative must take precedence, and we should be enabling them to be acknowledged and to take action when they innovate. For how many teachers, at all levels and even in higher education, which is innovating, are investing considerable effort in improving teaching methods and seeing these efforts go unappreciated and their achievements downplayed? This situation needs to change, so that those who are committed to changing practices receive the recognition they deserve: not to diminish those who, on a daily basis, are doing their jobs to the best of their abilities, but to show that it is possible to progress, to provide examples and, above all, to forge networks.

At the same time, one or more centers for conducting monitoring, foresight activities and pilot schemes could be set up within the Ministry to help bring about and roll out innovative strategies.

Encouraging Creativity and Innovative Teaching Practices

We are recommending launching a series of actions aimed at fostering creativity and innovative practices in the vocational training sector (Ministry of Labor), school education and higher education and research sectors with a view to setting up pooling mechanisms and financing tools that are common to the three sectors. The aim is to develop new experimental teaching methods at various levels (project-based teaching, cross-disciplinary teaching, peer-to-peer teaching, etc.).

Such measures are not intended to create new demands on teachers but, above all, to raise the profile and bring together in a network those who are busy setting up innovative teaching practices.

They will also seek to equip these pioneers and enable them to take action, not unlike what is being done for general interest entrepreneurs, by giving precedence to collective setups.

Priority must be given to revising teacher training ahead of any specific measure. In addition, innovations to the extracurricular system must be promoted and harnessed by the traditional education system: countless initiatives encourage creativity and innovation, among them competitions (French young mathematicians' tournament), educational support initiatives (Curious Mind) and local pilot schemes.

The number one objective is to raise the profile and bring together in a network all those who are breaking new ground in their creativity-enhancing teaching practices.

Setting Up a Platform for Promoting Pioneers of Innovative Teaching Methods

This could be somewhere for:

- accommodating a library of innovative teaching practices, with an explanation for each practice and the rules for putting them in place;

- indexing all of the innovative teaching projects that have been put into practice in schools, higher education & research or the vocational training sector;
- raising the profile of individuals and schools making use of these innovative practices and for encouraging exchange networks to form around such practices;
- launching and communicating on events aimed at rolling out innovative practices more widely.

This platform should be coordinated at the interministerial level (school education, research and the workplace).

Technion in Israel

The Technion—Israel Institute of Technology is a research institute and public university in Haifa, Israel, specializing in the science and technology fields. The classes of some programmes, computer science in particular, mainly comprise sessions focusing on a specific project. The students are placed in scenarios where they must solve a problem and/or create an object or function. The teacher plays a facilitating role. The results of this type of approach are inspiring: the company Waze was founded by Technion students, as well as many Google algorithms.

Expérithèque

The Expérithèque platform, set up by the French Ministry of National Education, keeps a record of teaching innovations nationwide. This type of setup could form the basis of a future platform, which would more largely factor in the aforementioned aims and relies on a dedicated coordination function.

Freeing Up Time and Means for Pioneers of Innovative Teaching Methods

Along the same lines as what is done for general interest entrepreneurs, some individuals or institutions who are leading the way with groundbreaking teaching methods could be awarded special means for documenting and sharing their projects. Two points should be borne in mind, however, when setting about this: the point is not so much to single out individuals as it is to promote collective setups. The main aim of making resources available—or granting leave to teachers from certain teaching duties, for example—must then be to formally put together and disseminate innovative teaching experiences.

The methods for recognizing this type of leave taken from traditional duties, in terms of career, must be developed. Creating a similar system for granting leave to staff who are not teachers should also be considered, to take into account the changes taking effect across all of the Ministry's occupations.

Part 3 — Anticipating and Controlling the Impacts on Jobs and Employment

4. Testing Out New Methods for Funding Vocational Training to Factor in Value Transfers

Vocational training is financed according to payroll. The development of AI is only making the changes affecting value chains more marked and phasing out the correlation between the stakeholders who finance vocational training and those who capture the added value. Accordingly, stakeholders with a very low payroll can be the source of a large proportion of added value of a global value chain, on which they can end up having a significant impact. But for the time being, they are not contributing to funding the professional transition of individuals employed by other stakeholders in the value chain.

To take an example, software accounts for 40% of a car's added value today, a figure which could rise to 70% in ten years' time. If the company that develops the software is not the same one that manufactures the rest of the car, this is an example of a phenomenon of the added value being tapped into by one of the links in the global value chain. This capture evidently comes hand-in-hand with high economic pressure on the chain's stakeholders, who see their economic value captured. These stakeholders are often traditional industrial groups, with a large workforce, which are responsible for ensuring the professional transition of their employees, in terms of guidance and financing. If we look at the automotive sector for example, we could imagine that a startup, with a workforce of 50, sells an autonomous driving software program to a longstanding French company that has tens of thousands of employees on its books, whose professional transition must be ensured. Or, more directly that a company could capture the added value by providing solutions aimed at automating the jobs of other stakeholders.

This phenomenon is not unique to AI, but more broadly has to do with the way in which the digital sector is evolving: so how can we ensure that the stakeholders tapping into the added value contribute to financing the development of individual skills—which they have sometimes had a hand in making obsolete?

The idea of a tax mechanism has been raised in the public debate, for similar reasons. But it is not without problems: first of all it seems to be at odds with support for the development of AI, since it evidently acts as a disincentive. Second, the tax base of such a mechanism is difficult to establish: should the robot, software or algorithms be taxed? Is it realistically possible to define without immediately organizing mass avoidance phenomena? Finally, this idea does not necessarily take on board the value chain and capture of added value, which nevertheless seems to be at the heart of the problem: taxing the company that is automating jobs, without realizing that all of the productivity gains associated with this automation are captured by another stakeholder who is developing artificial intelligence, seems to miss the very point at stake here.

It is therefore recommended to set up social dialogue around the sharing of added value at the level of the entire value chain. This type of bargaining does not correspond to the usual frameworks in which social dialogue takes place, most of the time at a national level and above all according to a vertical organization, by sector. One or two trials could be organized by the International Labor Organization, or sectoral social dialogue committees, on products and value chains that are

particularly symptomatic of value capture phenomena. The purpose of these negotiations will be to envisage how financing for vocational training may be separated from the payroll, at international level.

This would particularly allow for account to be taken, at another end of the chain, of the requirements bearing on career development and training of workers of crowdsourcing marketplaces, such as Amazon Mechanical Turk. For via such marketplaces, these workers are playing an instrumental role in the development of AI by preparing datasets or participating even more directly in the training of artificial intelligence algorithms.

5. Training AI Talent at All Levels

Tripling the Number of AI Graduates

To meet the needs of the workplace, a clear target needs to be set: triple the number of AI graduates within 3 years. To achieve this, and beyond the steps already being taken by the Ministry of Higher Education, Research and Innovation on the subject, the following is necessary:

Tailoring the existing training provision to AI

Existing mathematics and computer science courses, which teach the basic building blocks of artificial intelligence, should naturally be geared towards artificial intelligence learning outcomes. This is the case for engineering schools which already deliver specific programmes on this subject. Communication on existing AI courses should therefore be improved before any new ones are created, to help guide students in their choices more effectively.

A clear target needs to be set: triple the number of AI graduates within 3 years

More specifically, a series of incentives could be put in place for the attention of engineering schools. The organization of an AI challenge for supervised personal initiative achievements by the schools' engineers could form one of the measures.

Specific steps aimed at attracting more women to mathematics and computer science courses also merit being taken. It is recommended to set an ambitious target of 40% women in the digital streams (see the section on diversity and AI).

Finally, over and above the question of increasing the number of courses available, it is recommended that aspects relating to data ethics, privacy and data protection form an integral part of artificial intelligence courses.

AI teaching in Canada

Several Canadian universities offer courses and teaching grounded in the most recent developments of AI and deep learning (such as computer science and operational research). These include: in the Province of Quebec: Université de Montréal, Université McGill, Polytechnique Montréal, HEC Montréal, École de technologie supérieure (Montreal), Université Laval (Quebec); in the Province of

Ontario: University of Toronto and Rotman School of Management, the Faculty of Engineering at the University of Waterloo, the University of Guelph.

IVADO and the Montreal Institute for Learning Algorithms (MILA) also provide training programmes on AI techniques for students and professionals alike. In the Quebec-Ontario cooperation agreement on AI (September 2017), both provinces have made a priority of supporting "the development of curriculum to reflect the evolving AI ecosystem" and of studying "the labor needs of industry and postsecondary education to foster the growth of AI, and the role that immigration and mobility can play to respond to these needs".

Creating new courses and new training programmes in AI

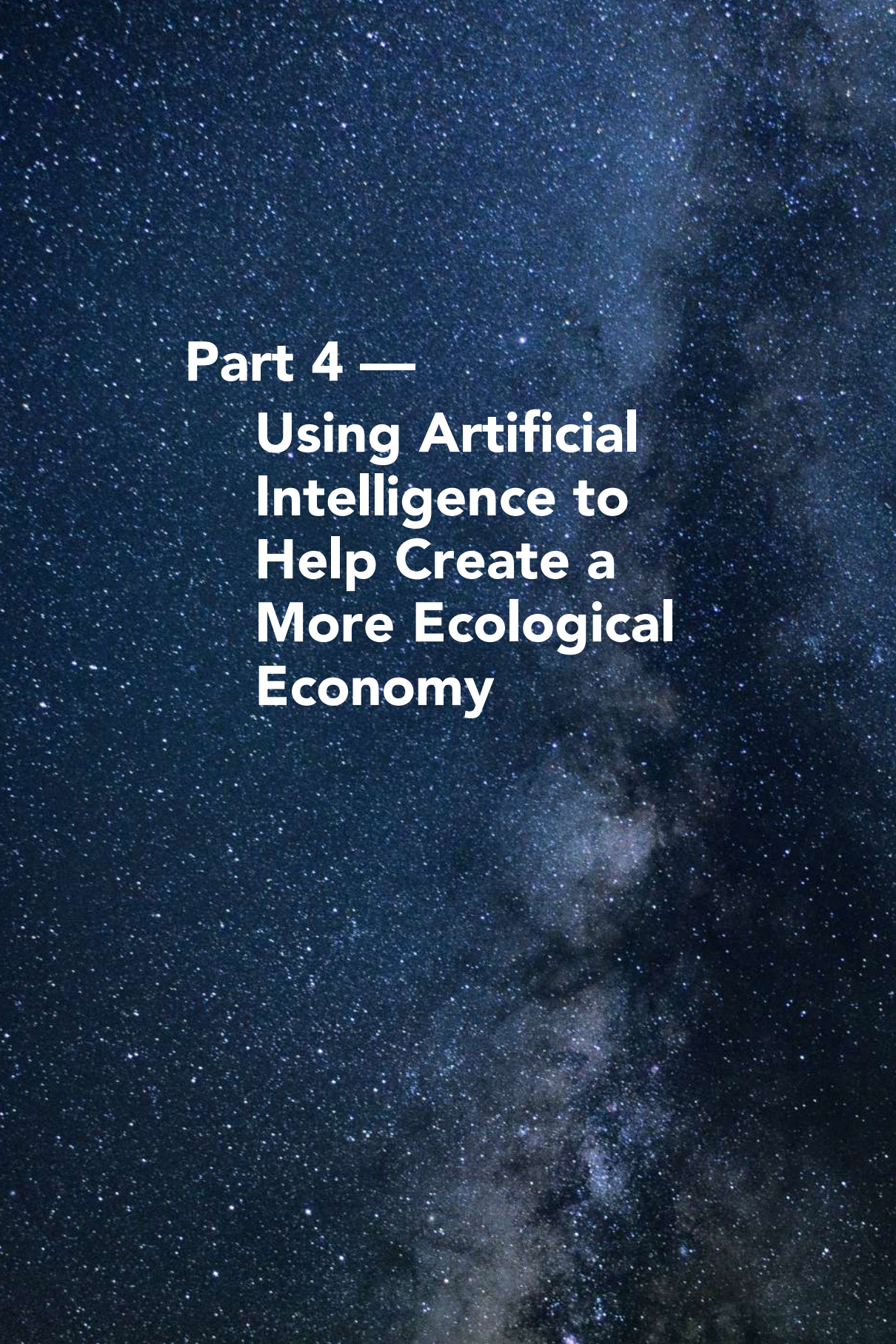
Although increasing the number of high-level training programmes, at master's or Doctorate level, is important, before that it is crucial to revise the way AI is taught. Training hybrid specialists, who are proficient in other skills in addition to skills in AI, is therefore necessary. Cross-linked and joint honors programmes could be set up, particularly with the university subjects where demand is greatest (medicine, physics, chemistry, sociology, psychology, law, etc.), at Bachelor's, Master's and Doctorate level.

Beyond cross-linked programmes, training and research in the social sciences must also be encouraged to tackle these subjects more directly, by adapting their course ranges and steering research towards this goal.

It is also necessary to foster the setup of general modules for students enrolled in another course or working professionals. Given the requirements in this area and the inescapable need to combine skills, open and accessible training programmes must absolutely be encouraged. These modules might comprise short programmes, delivered by universities and engineers undergoing continuing professional development. A call for proposals for setting up specific MOOCs on the subject could also be launched.

On a final note, beyond high-level master's or Doctorate courses, there is a growing need for qualified AI professionals who could be graduating with the equivalent of a Bachelor's degree (so three years' higher education or from a vocational stream). The aim is to train students in the much more technical occupations of AI, in which, although in-depth knowledge of AI is not a prerequisite, it would be deemed a direct asset by companies. Some examples include the industrialization of AI techniques, data visualization and analysis production or the integration and adaptation of AI components. This need is not remotely covered by the current training provision today, and is set to rise exponentially as AI infiltrates the workplace.

If we are to be able to triple the number of AI graduates in three years, political commitment will have to be matched with additional funding, so enabling training institutions to set up the necessary courses and be able to cater to higher numbers of students.



Part 4 — Using Artificial Intelligence to Help Create a More Ecological Economy

Part 4 — Using Artificial Intelligence to Help Create a More Ecological Economy

More than ever before, the revolution triggered by the development of digital technologies and their widespread adoption tends to obscure its impact on the environment[1]. Nevertheless, there is an urgent need to take this on board. Two years ago, the American Association of Semi-Conductor Manufacturers predicted that by 2040, the global demand for data storage capacity, which grows at the pace of the progress of AI, will exceed the available world production of silicon[2].

By 2040 the energy required for computation will equally have exceeded world energy production

Furthermore, by 2040 the energy required for computation will equally have exceeded world energy production; the progress of the blockchain may also cause our energy requirements to rocket. It is vital to educate as many people as possible about these issues and to act promptly to avoid shortages. At a time when global warming is a scientific certainty, it is no longer possible to pursue technological and societal developments if those are completely detached from the need to preserve our environment.

The ecological reality of the digital revolution in industry

Digital energy consumption increases by 8.5% per year and its contribution to world electricity consumption (which is growing by 2% per year) could reach 20% (moderate scenario) or even 50% (pessimistic scenario) by 2030, and therefore be multiplied 10-fold in 20 years' time. Given the global energy mix, the digital contribution to greenhouse gas emissions (GHG) will thus increase from 2.5% in 2015 to 5% in 2020 (2.5 GT).

The production of digital hardware uses large quantities of rare precious metals which are only partly recyclable, and the available reserves are limited (15 years in the case of Indium, for example, the consumption of which has multiplied 7-fold in 10 years); this could result in a technological impasse if this increase in demand does not slow down, especially given that some of these metals are also used in the production of equipment for renewable energy (wind and solar power). Aside from the peak in energy and oil consumption, there is also a growing concern about the peak in the use of these metals, which is also contributing to the peak in energy and oil consumption; since they are continually in shorter supply, more and more energy is required for their extraction. On top of this, both the extraction of these metals and the end-of-life processing of the equipment used (when the facilities for this are inadequate) are a source of soil pollution (this concerns over half the weight of electrical and electronic equipment in France and more than that worldwide).

1. See the Greenpeace report *Clicking Clean*: Who is winning the race to build a green internet?
2. See the 2015 American Semi-Conductor Industry's report: *rebooting the IT revolution, a call to action*

The AI boom is likely to reinforce these trends through the storage and exchange of a growing volume of data, an increase in computer power, the pressure to keep renewing equipment in order to improve performance, etc. The roll-out of new computer architecture which makes more efficient use of cloud computing could act as a brake on some of these trends (the volume of centralized data), but accelerate others (the renewal of equipment).

Source: The Shift project's contribution to the consultation organized by the mission

Although AI is a potential threat to the environment, it is also a potential solution. Indeed, there are many opportunities to use AI in the field of ecology: AI can help us understand the dynamics and the evolution of whole ecosystems by focusing on their biological complexity; it will allow us to manage our resources more efficiently (particularly in terms of energy), preserve our environment and encourage biodiversity. Developments in AI could result in the emergence of new ways to maintain and protect the natural environment, both on land and at sea; from autonomous robots that can remove invasive species of starfish to intelligent fences that can divert fauna so as to preserve them—there are a great number of possibilities for the development of new, more adaptable and respectful ways to interact with nature.

Nevertheless, this leads us to a well-known paradox concerning optimization: gains in energy saving and new possibilities in terms of consumption need to be offset against the fact that AI may result in various rebound effects (see inset). Therefore, AI may prevent us from rethinking our patterns of growth and consumption and change how we measure output, but at the same time result in consumption being at least as great, if not greater, than it was before.

What exactly is a rebound effect?

A rebound effect is a phenomenon whereby the expected savings in energy and resources due to the use of new technology may be partly or completely outweighed by society's response to it. It could take many forms: an increase in consumption of one single item of expenditure, or another one. The response of the consumer is the decisive factor. Investing in sanitary construction work in a building may result in a clearer conscience that may counterbalance the reluctance to book a vacation trip to a remote island. There is also an economic aspect to this: for example, the savings made on a heating bill may be reinvested in another product or activity which then adds to energy consumption.

A truly ambitious vision for AI should therefore go beyond mere rhetoric concerning the efficient use of resources; it needs to incorporate a paradigm shift toward a more energy-efficient collective growth which requires an understanding of the dynamics of the ecosystems for which this will be a key tool. We should take the opportunity to think of new uses for AI in terms of sharing and collaboration that will allow us to come up with more frugal models for technology and economics. AI allows us to model the dynamics and the future of biological ecosystems and thus may contribute to a genuine ecological transition. France and Europe could spearhead this

intelligent ecological transition. France possesses quite a few assets that would enable it to become the champion of sustainable and ecological artificial intelligence: academic ecosystems, important research into nanotechnology and an incredible wealth of data from the energy, farming, marine, water supply and transport sectors for example.

1. Making this Issue Part of the International Agenda

On an international level, France has the capability to take on this leadership. Initially, it could put forward the idea of carrying out a study of the impact of AI on the attainment of the UN sustainable development goals (SDGs), in order to find out to what extent AI is a hindrance in some cases and a help in others. France could be at the forefront with this type of research, which has not yet been initiated by any other country. It could be coordinated with the impetus provided by the Climate Agreement and the Global Pact for the Environment. France could propose setting up a major event along the lines of the COP 21, to showcase exemplary and high-impact initiatives. It could also be more closely involved in the convergence of the two transitions, ecological and digital, within international forums, particularly the G7, where discussions concerning AI were initiated and where France is shortly to take over the presidency.

2. Promoting the Convergence of the Ecological Transition and Developments in AI

Establishing a Meeting-Point for the Ecological Transition and AI

For stakeholders from the two transitions—the digital and the ecological—can come face to face, AI research needs to be confronted with the disciplines which are aiming to understand the complexity of our energy and material resources and optimize their use. This exchange could take place within the framework of the AI research network recommended in Part 2, or on premises where there is already an investment in environmental research. This branch of research could focus on various subjects:

1. The study of new methods of storage which would be more economical but fundamentally disruptive, following the example of DNA storage.

DNA storage

The potential for synthetic DNA storage is being studied by researchers worldwide. In July 2016, the research center at Microsoft converted 200 megabytes of data into DNA. The year later, researchers at Microsoft claimed that the company would have a DNA storage system up and running by the end of this decade. In France, the start-up DNA Script is also working on this topic. The ecological benefits of this method of storage are substantial.

The density of the stored information is much higher: researchers from the University of Columbia and the New York Genome Centre (NYGC) have

demonstrated that it is possible to achieve a theoretical density of 215 petabytes per gramme, by using DNA storage. By way of comparison, Samsung has currently succeeded in storing 512 GB in a 1-gramme chip; the ratio is therefore more than 10^6.

Consequently, these developments are paving the way for significant reductions in the use of heavy metals and reductions in greenhouse gas emissions via the possibility of moving away from the huge data center model and from on-chip storage in general.

Although costs remain high and future advances—in terms of accelerating the storage and reading processes—are still very uncertain, the fact remains that progress is continuously being made in the field.

2. Projects at the crossroads of life sciences and ecology, which follow the example of the *Tara Oceans* project. In fact, measures concerning biodiversity and its monitoring, and ecosystem-based services (for example, water quality, impact studies, invasive and toxic species, fish stock management, key species, etc.) will be put in place in the immediate future using sequencing and high-speed automatic imaging; AI will be used to process and model this data. AI that has been developed for understanding ecosystems could probably be recycled for use in an economic context.

The Tara Oceans Project

The Tara Oceans project involves collecting and freeing-up massive amounts of data concerning the oceans for the purposes of understanding and modelling a planetary biome. Launched in 2008, the Tara Oceans expeditions are travelling the oceans in order to measure the global ecosystem for the very first time and embrace it in all its complexity—from its viruses to its animals and from genes to the structures of its organisms—basing this on physical and chemical parameters. The programme, funded through innovative public/private methods, has produced massive amounts of ecosystem-based datasets (notably >90 Terabytes of DNA data and >30 Terabytes of images of planktonic organisms) which are freely available online from European databases. Now, they form the most complete description of a global ecosystem. The layers of standardized data, from genomes to satellite images—not to mention the images of organisms and the physical and chemical measurements—form a unique case study for the application and development of AI protocols capable of measuring and modelling biodiversity and its interactions with biogeochemical cycles and the climate.

3. Climate and weather research: calls for projects could be aimed at the domains of weather and climate prediction and risk prevention. In fact, AI could make major advances possible in a field where French expertise already has a very good reputation. AI may thus be particularly useful within the context of the prediction of hazards and impacts, particularly where the conventional data used in weather forecasts meet new types of data (typically vehicles and weather stations connected to the Internet). This research will, in addition, prove extremely useful in the development of more intelligent agriculture and

transport, and in reinforcing the competitiveness of our importers (forecasting harvests, etc.).

Establishing a Platform for Measuring the Environmental Impact of Intelligent Digital Solutions

At the moment, there is no clear policy in favor of an ecological evaluation of digital solutions. Experimental devices of this sort already exist[3]. Ademe (the French Agency for Environment and Energy Management) could be responsible for their analysis and development, so as to create a national base that would make this type of evaluation more widespread. This base could, in particular, use incentives to get standardized technical specifications about the products concerned put into open data.

Once this base has been created, a simple tool needs to be invented that would allow any citizen to inform themselves about these issues. A website could be created for the purpose of comparing the ecological impact of the various products and services, software and hardware involved in the digital value chain. This site would need to rely on a database that allowed the evaluation of the environmental impact of *all* the aspects of the ongoing digital dematerialization process, for individuals (the impact of personalized recommendations, chatbots, image recognition techniques, etc) as well as for businesses, so as to allow them to evaluate their digital providers.

This portal could promote simple procedures for minimizing the ecological impact of digitalization. For inspiration, *GreenIT.fr* and the action group Conception Numérique Responsable (Responsible Digital Design), in particular, are developing a green approach to design and ecological tools which have already reduced the impact of one digital service by a third, whilst improving the experience of its users. It could be accompanied by incentives to put the technical characteristics of products into open data. This comparative approach could make it possible to promote smaller operators and new business models, rather than the most frequently used solutions and services.

Apart from this platform and this site, a broader initiative could be introduced that would encourage businesses to use an ecological approach. Ademe could have a role to play in the design and circulation of an evaluation reference system and in the ecological design of intelligent digital innovations.

3. Designing AI that Uses Less Energy

Recent progress in AI has largely been due to the increased use of GPUs, graphics processors, to carry out general-purpose and massively parallel computing. However, like the overwhelming majority of chips, these use silicon in their electrical circuits and transistors.

3. See the Conception Numérique Responsable (Responsible Digital Design) action group's tool *ecoindex.fr*, for example

What exactly is a GPU (graphic processing unit)?

Moore's Law bases the increase in our capacity for computing on an increase in the number of microprocessor transistors. GPUs are developing another approach: an increase in the number of processors working simultaneously in collaboration. GPU technology is very efficient and allows the application of ever more complex and efficient AI techniques that make use of increasing amounts of data. It has therefore contributed to the development of neural networks, the capacity of which largely depends on the available computing capacity and the volume of data they are able to process. The GPU has thus become crucial in the innovation race since the increase in computing capacity has become directly proportional to the capacity for experimentation; however, very little research is being done on the environmental impact of using GPUs for generic computation. Although computation via GPUs may be more economical than computation via CPUs (central processing units), their energy consumption nonetheless remains very high.

Today, GPUs are seen as keys to developing AI; this is why it is also essential to design disruptive innovation in these domains, both for reasons of sovereignty and for environmental reasons (*see the developments in Part 1*). In fact, innovations that can be developed in the field in the short term have the potential to consume less energy. It is therefore urgent to bring the semiconductor industry and its French and European researchers together over these research and experimentation issues, as we have shown in Part 1 *Innovating in the components industry adapted to AI*.

Taking Action in the Greening of Data Centre Value Chains

Through the development of the Digital Single Market, it is essential that we support the European cloud industry in its ecological transition, in order to keep greening the AI value chain. Certain stakeholders are already exemplary in terms of their energy efficiency and it is important to circulate these good practices throughout the sector. In addition, AI can make a valuable contribution to this greening process. In 2016, Deepmind optimized energy consumption in its data and cooling system centers thanks to machine learning, thus increasing energy efficiency in its data centers by 15%.

It is essential that we support the European cloud industry in its ecological transition

Supporting Ecological Moves on the Part of European Cloud Providers

Creating a label and encouraging the use of ecological cloud providers by local authorities and the State

This support could take the form of creating a label which could be managed by the Ministry for the Ecological and Inclusive Transition in conjunction with the Minister of State for the Digital Sector and it could be linked to the implementation of tax

incentives. This move should, however, be coordinated with the present and future efforts of the European Commission on this subject. [4]

It is important to speed up the process of ecological transition in public services by encouraging them to use ecological cloud providers, and by relying on a label created in accordance with Article 10 of decree no. 2016-360 of 25 March 2016 relating to public procurement. The use of European cloud providers needs to be supported since larger computing and storage spaces provide better energy efficiency than multiple smaller data centers[5].

Encouraging the recycling of heat produced by data centers

We need to start planning for the recycling of heat produced by data centers, the new value chains that correspond to this and the investment needed to implement them. The State could assist local authorities in addressing these issues. Although at the present time, the amount of energy obtained is still limited, this is expected to increase.

Several initiatives already exist in this field; they need to be encouraged and supported.

Some examples of recycling the heat produced

The Natixis bank's data center in Marne-La-Vallée supplies water at 55 °C to heating systems in an area undergoing urban development and to the local Val d'Europe water sports center.

The Stimergie company has developed a system which allows the recovery of 1 MWh of heat per server per year, which represents 60% of the heat generated, i.e. the servers' energy consumption is reduced by more than half. The company has signed several contracts to install this system elsewhere in France, including at a block of 40 flats in Nantes and at the swimming pool at La Butte aux Cailles in the 13th arrondissement (metropolitan district) in Paris.

Tapping into open hardware and open software

The *Open Compute* project has shown that the open hardware approach can make significant energy savings possible. Facebook announced that it has saved $2bn in infrastructure costs in 3 years thanks to this project, and that it has gained 38% in energy efficiency and saved up to 24% in operating costs.

4. Following the public consultation on this subject, which was led by the Commission in October 2017, initiatives should be implemented shortly.
5. See the Direction générale des entreprises (General Directorate for Enterprise) Guide to cloud computing and data centers

The Open Compute project

The Natixis bank's data center in Marne-La-Vallée supplies water at 55 °C to heating systems in an area undergoing urban development and to the local Val d'Europe water sports center.

The Open Compute project was launched by Facebook in 2011, in association with Intel, Rackspace, Goldman Sachs and Andy Bechtolsheim; HP, Dell, Cisco, Apple and Microsoft have now joined the movement. The project and its foundation aim to design, use and promote the distribution of the most effective and adaptable computing and storage solutions for infrastructures. Contributions to the project must meet these 4 criteria: performance, scalability, access and impact.

Several European initiatives are currently underway that are designed to support the European cloud industry; these would need to be coordinated with open hardware and open software initiatives in order to reinforce the European market's confidence in cloud. The European cloud initiative could prove useful here, and such an approach would allow an increase in the influence of European stakeholders by increasing their share in the market and their presence in discussions concerning standardization.

4. Releasing Ecological Data

AI is opening up radically new perspectives in terms of understanding and preserving the environment. Whether it is employed in the identification and preservation of biodiversity, the remedying of damage that has already been caused, the modelling of the impact of our actions, the most efficient use of resources, the harnessing of sources of renewable energy or else as a tool for use by shared services, AI can contribute to a reduction in general consumption and can boost all our initiatives towards respecting and restoring regional and global ecosystems. From reforestation using drones to the mapping of living species using new possibilities furnished by image recognition, AI can supply a growing number of increasingly powerful tools to enable us to fully engage in the process of ecological transition.

AI is opening up radically new perspectives in terms of understanding and preserving the environment

How can we take full advantage of this to support France's reputation in the field of ecology? Which initiatives should be prioritized? Two initiatives appear to take precedence: the creation of sets of data which would cross-reference various sources, including genetic, and be available to as many people as possible—researchers, innovation leaders, State-owned start-ups, etc.—and supporting specific objectives. We are putting forward two more: reducing our carbon footprint through greater transport efficiency, and French agriculture's transition towards a more intelligent and less polluting form of agriculture. These two objectives could thus be embodied in specific sector-based policies, along the lines of those described in Part 1. These challenging sector-based proposals are raised in the

sections that focus on the different industries (agriculture and transport). It is obvious that there is enormous potential for using AI in the energy sector; however, in view of the economic characteristics of the sector, a specific initiative on the part of the public authorities would not necessarily appear to be a priority.

OpenSolarMap has proved that great value can be inherent in the most traditional of databases[6]. We need to leverage all the public data that concerns our territories, our homes, our energy consumption, etc. to assist us in developing AI solutions for the ecological transition process.

French legislation is beginning to take on board the importance of data and the need to share it, particularly where it relates to the concept of data of general interest[7]. The latest report from the Ministry for the Ecological and Inclusive Transition[8], which registers the public databases in its domain, demonstrates the large volume of data which is now available. The move undertaken by this same Ministry regarding mapping and data usage should be encouraged, fast-tracked, even extended and linked to other Ministries' data policies (Agriculture, Housing, Health, etc).

This move ought to go hand in hand with informing public stakeholders in ecology about the potential offered by AI. In particular, this would identify the Ministry's current missions and programmes which could then be accomplished more easily with help from AI, following the example of environmental policing and the updating of certain bases using new image-recognition techniques. To prepare for this move, an 'Intelligent Review' of environmental policies could be introduced at national and local levels: climate plans, Agenda 21 programmes, waste prevention schemes, transport plans, the SNTEDD (the stratégie nationale de transition écologique vers un développement durable, the National Strategy for Ecological Transition towards Sustainable Development), etc. in the light of AI solutions which could maximize their impact. The review's objective would be to ensure that these strategies were taking advantage of the potential offered by AI. In addition, AI specialists need to be involved in defining any new public environmental strategies.

Releasing Public Data

In order to develop AI solutions for the ecological transition process, it is crucial to make public data (meteorological, agricultural, transport, energy, biodiversity, climate, waste, land registry, energy performance analysis, etc) available to everyone—European researchers and businesses alike—ideally before 2019. This release of data could encourage innovations which pave the way for rapid action:

6. The *OpenSolarMap* project relies first of all on data from land registries, before involving data from satellites and other contributions.
7. Notably via the Law for a Digital Republic. However, in terms of the environment, European directives concerning environmental and geographical information are an invitation to even greater access to information that concerns the environment, and recent French legislation concerning the energy transition process and green growth or alternatively the legislation concerning biodiversity, including measures which make the sharing of data obligatory. Aware of the importance of the subject, the Ministry for the Ecological and Inclusive Transition was also provided with a Supervisor-General for Data in 2016.
8. See the CGEDD Data Mapping Report from the Ministry of the Environment, Energy and the Seas.

shared housing renovations, developments in renewable energy, energy efficiency, facilitating shorter supply chains, the recycling of household and industrial waste, planning permission, demolition permits, etc. But it could also be utilized in research into more structured innovations: weather prediction without differential equations, improvement in predictive traffic systems, pollution and flood warnings, etc.

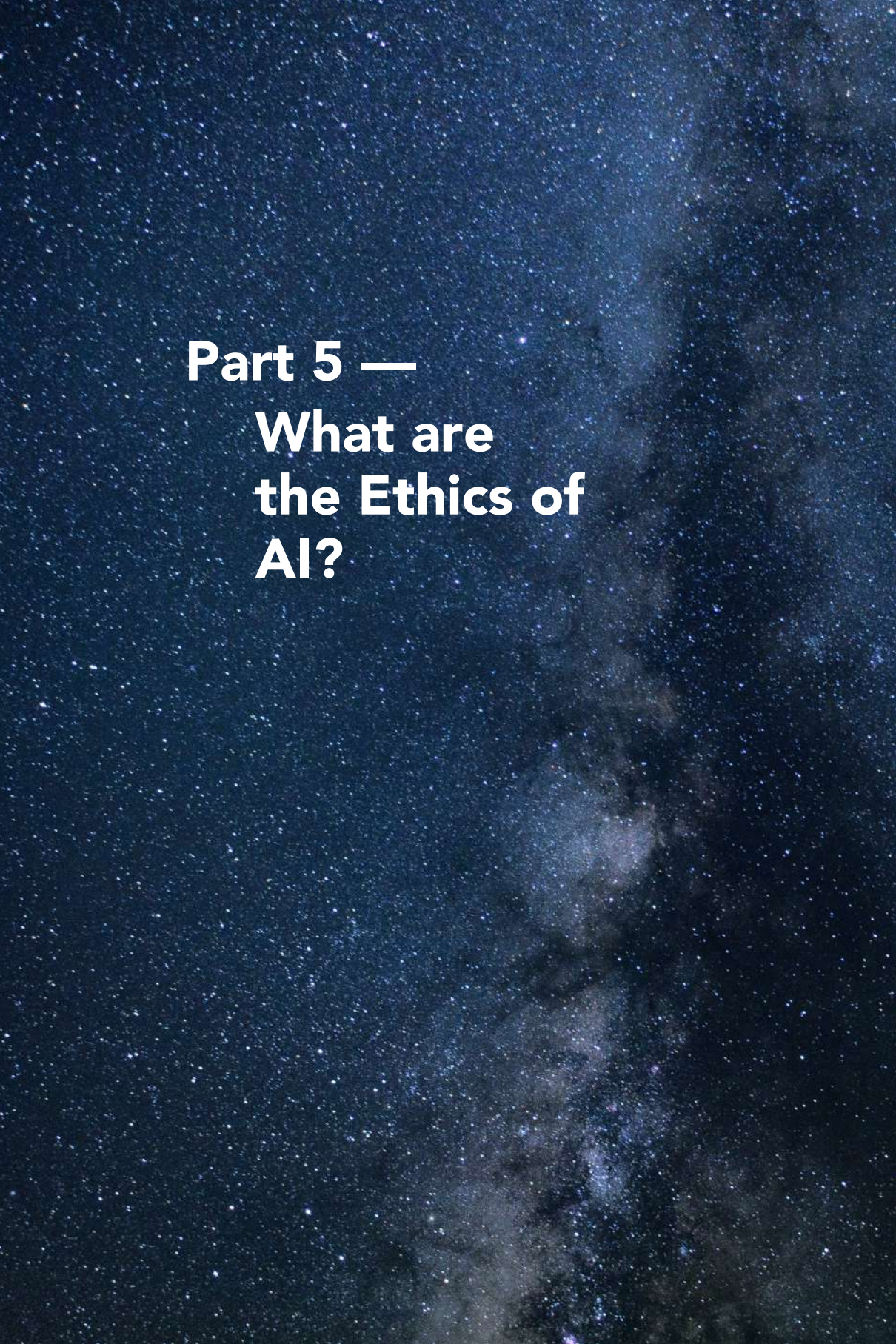
The release of this data should allow the emergence of European leaders in this field; for example, service platforms for construction and housing renovation. Vigilance is therefore required regarding the release of data linked to these issues (energy efficiency, construction data, etc) so as to avoid these sectors becoming overtaken by foreign stakeholders. Access to this data should be promoted within the framework of precise sector-based challenges.

Within the agricultural sector, the release of public data could thus give the digital transformation of agriculture a new framework, by putting it into a wider context than that of just food and health traceability. By combining data linked to the CAP, accounting data, data connected with processing, distribution, the food supply, nutrition and many other types of data connected with industries related to agriculture, public authorities can assist in the implementation of a new French agricultural economy capable of fostering the emergence of leaders in foodtech.

Releasing Private Data

Giving free access to certain kinds of private data needs to be done with care, within the context of the implementation of sector-based challenges (*see Part 1 and its developments in terms of data policy*).

Giving free access to this data should also form part of a policy of incentives aimed at the larger of the French groups in the sectors involved in the ecological transition process (water supply, waste disposal services, etc).



Part 5 — What are the Ethics of AI?

Artificial intelligence now affects every aspect of our social lives. Without always being aware of it, we interact on a daily basis with intelligent systems which optimize our journeys, create our favorite playlists and protect our inboxes from spam: they are our invisible workforce. At least, this is the role we have assigned to them: improving our lives, one task at a time.

Recent progress in AI's several fields (driverless cars, image recognition and virtual assistants) and its growing influence on our lives have placed it at the center of public debate. In recent years, many people have raised questions about AI's actual capacity to work in the interests of our well-being and about the steps that need to be taken to ensure that this remains the case.

This debate has principally taken the form of a broad discussion about the ethical issues involved in developing artificial intelligence technology and, more generally, in the use of algorithms. In different parts of the world, experts, regulators, academics, entrepreneurs and citizens are discussing and sharing information about the undesirable effects—current or potential—caused by their use and about ways to reduce them.

Faced with the need to take respect for our values and social standards on board when addressing the potential offered by this technology, these discussions have logically drawn on the vocabulary of ethics. They occupy the available space between what has been made possible by AI and what is permitted by law, in order to discuss what is appropriate. However, ethics is clearly a branch of philosophy which devotes itself exclusively to the study of this space by attempting to distinguish good from evil, the ideals to which we aspire and the paths which take us away from them.

Aside from these purely speculative considerations concerning AI's 'existential threats' to humanity, debates tend to crystallise around the 'everyday' algorithms

Furthermore, aside from these purely speculative considerations concerning AI's 'existential threats' to humanity, debates tend to crystallize around the 'everyday' algorithms which organize our news feeds, help us decide what to buy and determine our training routines. In 2017, Kate Crawford, Cathy O'Neil and many others reminded us that we are not all equal before these algorithms and that their partiality has a real impact on our lives. Every day, invisibly, they influence our access to information, to culture, to employment or alternatively to credit.

Consequently, if we hope to see new AI technology emerge that fits in with our values and social standards, we need to act now by mobilizing the scientific community, the public authorities, industry, the entrepreneurs and the organisations of civil society. Our mission has humbly attempted to suggest a few ways in which we can start building an ethical framework for the development of AI and to keep this discussion going in our society. These are based on five principles:

In the first place, there needs to be greater transparency and auditability concerning autonomous systems. On the one hand we can achieve that by developing our

capacities to observe, understand and audit their performance and, on the other, through massive investment in research into their accountability.

Next, the protection of our rights and freedoms needs to be adapted to accommodate the potential for abuse involved in the use of machine learning systems. Yet it appears that current legislation, which focuses on the protection of the individual, is not consistent with the logic introduced by these systems—i.e. the analysis of a considerable quantity of information for the purpose of identifying hidden trends and behavior—and their effect on groups of individuals. To bridge this gap, we need to create collective rights concerning data.

Meanwhile, we need to ensure that organisations which deploy and utilize these systems remain legally responsible for any damages caused. Although the terms of this legislation concerning responsibility are still to be defined, the French Data Protection Act of 1978 and the GDPR (2018) have already established its principles.

However, legislation cannot solve everything, partly because it takes much more time to generate law and norms than it does to generate code. It is therefore vital that the 'architects' of our digital society—the researchers, engineers and developers who are designing and commercializing this technology—do their own fair share in this mission by acting responsibly. This means that they should be fully aware of the potentially negative effects of their technology on society and that they should make positive efforts to limit these.

In addition, given the important nature of the ethical questions that confront future developments in AI, it would be prudent to create a genuinely diverse and inclusive social forum for discussion, to enable us to democratically determine which forms of AI are appropriate for our society.

Finally, it becomes more crucial to politicize the issues linked to technology in general and AI in particular, in view of the important part it plays in our lives. To this end, the proposed *Chambre du futur* (Chamber of the Future), announced by the President of the Republic in the context of the reform of the ESEC, the French Economic, Social and Environmental Council, needs to play a major role in the strictly political debate on artificial intelligence and its consequences.

1. Opening the 'Black Box'

A large proportion of the ethical considerations raised by AI have to do with the obscure nature of this technology. In spite of its high performance in many domains, from translation to finance as well as the motor industry, it often proves extremely difficult to explain the decisions it makes in a way that the average person can understand. This is the notorious 'black box problem': it is possible to observe incoming data (input) and outgoing data (output) in algorithmic systems, but their internal operations are not very well understood (see inset). Nowadays, our ignorance is principally due to changes in the paradigm that is introduced by machine learning, in particular deep learning. In traditional computer programming, building an intelligent system consisted of writing out a deductive model by hand, i.e. the general rules from which conclusions are inferred in the processing of individual cases. Such models are by definition explainable, inasmuch as the rules which determine their decision-making are established in advance by a programmer, and it is possible to tell in each individual case which of the rules have been activated

so as to arrive at a conclusion (for example, if your income is less than so much per month, you will be refused a loan).

Explaining the Decisions Made by Machine Learning Systems

The most efficient machine learning technique today, deep neural networks (Deep Learning), does not rely on rules established in advance. In the case of image recognition, for example: if we wanted to develop an algorithm which automatically categorized photos of cats and dogs, the data being processed would consist of images in the form of an array of pixels and it is virtually impossible to write out a programme by hand that is sufficiently powerful to classify all the images accurately from the data, pixel by pixel.

The accountability of systems based on machine learning constitutes a real scientific challenge

At this stage, the accountability of systems based on machine learning thus constitutes a real scientific challenge, which is creating tension between our need for explanations and our interests in efficiency. But although certain models of machine learning are more easily explainable than others (systems based on rules, simple decision trees and Bayesian networks), nowadays their performance does not generally match up to that of deep learning algorithms.

What we do not understand about deep learning

Neural networks and deep learning techniques are routinely condemned by their users for seeming just like black boxes. This argument can be equally applied to a large number of other machine learning techniques, whether we are talking about Support Vector Machines or random forests (the operational versions of decision trees). The reason is not so much inherent in the nature of the model used but resides more in a failure to produce an intelligible description of the results produced in each case and, in particular, to highlight the most important features of the case in question that have led to these results.

This failure is largely due to the dimensions of the spaces in which the data are evolving, which is particularly crucial in the case of deep learning. For example, for image recognition, a deep network inputs images described by thousands of pixels (4K) and typically memorizes hundreds of thousands, even millions, of parameters (network weights), which it then uses to classify unknown images. It is therefore almost impossible to follow the path of the classification algorithm, which involves these millions of parameters, to its final decision. Although in terms of one image, this accountability seems of relatively low importance, it is a lot more crucial in the granting of a loan, for example.

In the long term, the accountability of this technology is one of the conditions of its social acceptability. Regarding certain issues, it is even a question of principle: as a society, we cannot allow certain important decisions to be taken without explanation. In fact, without being able to explain decisions taken by autonomous systems, it is difficult to justify them: it would seem inconceivable to accept what cannot be

justified in areas as crucial to the life of an individual as access to credit, employment, accommodation, justice and health.

Equity, Bias and Discrimination

The obscure nature of this technology is all the more worrying as it may conceal the origins of reported bias, so that we are unable to tell, for example, whether it originates from the algorithm itself or the data used to train it… or both. For instance, some researchers have established the algorithms used by Google in its targeted advertising are more likely to offer less well-paying jobs to women, that YouTube's moderating algorithms are sometimes slow to react when a harmful content is reported and thus allow its viral spread, or alternatively that algorithms that predict criminal behavior recommend a higher level of surveillance in poorer Afro-American quarters. Indeed, all these algorithms only reproduce the prejudice that already exists in the data they are supplied with. But these observations give rise to legitimate fears, and if we are slow to act we run the risk of seeing a widespread distrust of AI on the part of the general public, which in the long run is liable to curb its development and all the benefits it could bring.

The law prohibits any form of discrimination based on exhaustive lists of criteria in the spheres of employment, housing, education and access to goods and services. In these instances, what constitutes discrimination is deemed to be clauses, criteria or practices which seem to be harmless, but which are liable to leave certain individuals at a disadvantage compared to others, except where there is objective justification for these clauses, criteria or practices in the form of a legitimate aim and where the means to achieve this aim are appropriate and necessary.

The use of deep learning algorithms, which feed off data for the purposes of personalization and assistance with decision-making, has given rise to the fear that social inequalities are being embedded in decision algorithms. In fact, much of the recent controversy surrounding this issue concerns discrimination towards certain minorities or based on gender (particularly black people, women and people living in deprived areas). American experience has also brought us several similar examples of the effects of discrimination in the field of crime prevention.

Because systems that incorporate AI technology are invading our daily lives, we legitimately expect them to act in accordance with our laws and social standards

Because systems that incorporate AI technology are invading our daily lives, we legitimately expect them to act in accordance with our laws and social standards. It is therefore essential that legislation and ethics control the performance of AI systems. Since we are currently unable to guarantee *a priori* the performance of a machine learning system (the formal certification of machine learning is still currently a subject of research), compliance with this requirement necessitates the development of procedures, tools and methods which will allow us to audit these systems in order to evaluate their conformity to our legal and ethical frameworks. This is also vital in case of litigation between different parties who are objecting to decisions taken by AI systems.

To date, these skills—even after the event—are almost non-existent for various reasons. In the first place, deep learning techniques are still too obscure (see above) and their audit protocols are still in their infancy. Then, businesses that have invested substantial sums of money in the construction of their algorithmic systems and would like to reap their rewards are necessarily reluctant to see their intellectual property divulged to third parties. The possibility of accountability for automated decisions is in this sense limited by a certain number of legal obstacles, such as the protection of intellectual property and trade secrets, the protection of personal data, the secrecy necessarily surrounding a certain number of State activities and activities concerned with security and public order. As a result, there is a widespread need to introduce a buffer between the realms of secrecy and of legitimate information.

Developing the Auditing of AI

Providing official auditing for algorithms

The appointment of a body of experts with the requisite skills would appear to be essential to the documentary auditing of algorithms and databases and for checking them using any means deemed necessary. This recommendation is in line with recent developments in the field of competition law and data protection, where the action pursued by the authorities is gradually moving from an a priori control of companies to a logic of audit *a posteriori*. Such obligations will, where necessary, be laid down by sector-specific regulatory bodies or for specific domains.

This recommendation is a response to the specific need for certified audits with probative force when it comes to contentious legal proceedings. To confirm one party's suspicions or claims, external observations of the performance and effects of algorithms alone are not sufficient to constitute admissible facts in a great number of cases. Whether this occurs during a judicial inquiry or one being carried out by an IAA (an independent administrative authority), it may be necessary to carry out documentary checks. It is not always necessary, useful or even possible to draw conclusions from an examination of a source code. The auditors may be satisfied with simply checking the fairness and equity of a programme (doing only what is required of them), by submitting a variety of false input data, for example, or by creating a large quantity of system user profiles according to precise guidelines, etc. For example, in order to check the gender equity of a recruitment website, a very large number of CVs belonging to men and women who are following the same career paths need to be submitted; in addition, these need to be representative of all those seeking work who are targeted by the site. The output reveals which applications for interview were granted and the average salaries proposed, etc. The system's provider could be forced to open an API which is designed to test their programme on huge numbers of artificial users (which would also possibly be generated by AI programmes).

As regards court referrals, two distinct levels of requirement have been identified: a primary function that could be called upon for legal purposes within the context of investigations carried out by independent administrative authorities, and a secondary function that would follow a referral by the Defender of Rights.

Developing public appraisal of AI

The potential to evaluate and audit AI should not be confined to government agencies; it should also be provided by civil society. This is a mission which a great many associations have already decided upon. Public authorities have a duty to lend this potential their support and to this end, we need to anticipate the financial problems facing civil defense agencies and journalists in their continuing role as the watchdogs of our digital era. As a guide, *Propublica*, the benchmark investigative media outlet for digital liberty which is financed by the Soros Foundation to the tune of $20m, has at its disposal five highly qualified full-time experts, developers at technology firms and/or post-doctorate students at the best universities, development support teams and a wide range of academic support. It would be difficult to locate similar resources elsewhere amongst French associations or in journalism, especially in the field of machine learning.

Consequently, at the very least we need to oil the wheels of communication between the authorities, research and civil society by maintaining the roles of ombudsmen who are committed to supporting initiatives which aim to mobilize AI in efforts to understand discrimination.

One of the main problems in terms of public auditing is getting access to data, which is frequently held by private stakeholders. There are currently voluntary initiatives on behalf of stakeholders such as Google which consist of making data available for the purpose of studying gender issues or to help us understand the phenomena of the non-use of rights, for example.

In parallel with this voluntary approach to making data available, specific assistance could be put in place for organisations which are not equipped to access it (for example, in terms of their ability to secure data, etc). To this end, funding to assist with the accommodation and boosting of projects (scientific, engineering and legal support, etc.) could be considered, under the auspices of an organization whose independence is guaranteed. In this connection, the efforts of the organisations *Team Up Turn*, *Propublica* and the *Electronic Frontier Foundation* in the United States could serve as examples.

In addition to assistance in terms of access to data, support for testing procedures and reverse-engineering could be introduced. These auditing procedures should not be the preserve of public auditors. To support public auditing, incentives could be offered to the public for making data available for research purposes and to associations who are actively defending social rights and freedoms, in order to help create different profiles and pathways of users, etc. Using the asset of citizens' portability (see Part 1 of this report) could be one of the best ways to achieve this.

Supporting Research into Accountability

In the digital sphere, the most significant scientific progress often results from close collaboration between public authorities, research laboratories and manufacturers; AI is no exception.

Gaining inspiration from the DARPA programme '*Explainable AI*'

In August 2016, the American agency responsible for defense research projects, DARPA, launched a call for proposals supporting research into understanding the nature of AI. This programme was identified as a major priority by the defense industry and aims to finance the development of AI systems which can be explained in terms of their structure. To this end, it supports three lines of research: how to produce models that are more easily understood, how to produce more intelligible user interfaces and understanding the cognitive mechanisms at work in the production of satisfactory explanations.

Although the total amount of funding available has not been made public, initial reports concerning the projects that were successful (13 in total) suggest that it amounts to several tens of millions of euros. Oregon State University alone received €5.2m over three years to fund the work of 8 researchers who are looking into machine learning.

Drawing on the DARPA programme *Explainable AI*, there appears to be an urgent need to support research into understanding the nature of AI by investing in the same three lines of research: how to produce models that are more easily understood, how to produce more intelligible user interfaces and understanding the cognitive mechanisms at work in the production of satisfactory explanations. Each of these areas involves a whole variety of skills—computer science and mathematics, of course, but also design, neurosciences and psychology—and highlights the need for interdisciplinary collaboration: understanding how things work is not just the preserve of developers but involves the whole of the scientific community (see also the recommendations in Part 2 of this report).

2. Considering Ethics from the Design Stage

Incorporating Ethics into the Training of Engineers and Researchers Studying AI

Machine learning techniques have come to play a major role in many fields, such as those of industry, business, public services, medicine or alternatively education. Consequently, the researchers, engineers and entrepreneurs who contribute to the design, development and marketing of AI systems have come to play a decisive role in the digital society of the future. It is crucial that they act responsibly and take the socio-economic impact of their activities into consideration. To guarantee this, we must raise their awareness—right from the start of their training—of the issues involved in the development of digital technology. Currently, this aspect of their education is almost completely lacking in engineering school syllabuses and university IT courses, even though there is a constant increase in the volume and complexity of the ethical questions with which these future graduates will be confronted as they keep pace with the very rapid advances in AI.

It is important to clarify the vision and scope of this aspect of their education so as to alleviate a number of concerns and identify expectations. Firstly, ethics is not reducible to specific morals or doctrines to be imposed on students so as to make them into 'good people'. Neither does it consist of giving lessons in conformity

which exclusively involve respect for all the legislation and regulations contained in company policies; we already expect computer experts to respect the law. The aim of teaching ethics is rather to pass on to the future architects of a digital society the conceptual tools they will need to be able to identify and confront the moral issues they will encounter—within the context of their professional activities—in a responsible fashion. In addition, bearing in mind the practical implications of questions raised concerning the protection of privacy, discrimination and intellectual property, they need to receive practical instruction so as to be equipped to make the connection between normative theories (professional ethics) and their application to particular circumstances. This requirement seems all the more necessary given that a significant proportion of the issues raised are not immediately apprehensible under the law. What can we do about the fact that recommendation algorithms are keeping users living in the security of comfortable filter bubbles, isolated from the realities of living in an ever more complex world? Should programmers work towards pluralism? From another angle, should the selection process for finding the best candidate to fill a post be reduced to merely looking at qualifications awarded by educational institutions and universities? In cases where standards are non-existent, are not mentioned or are insufficient, the developer has an increased moral responsibility. Far from finding immediate solutions, teaching ethics could nonetheless trigger a virtuous cycle: training specialists to be more responsible could lead to the development of more responsible technology.

What should these courses contain? In order to be able to train specialists to be more responsible, the teaching of ethics—and the social sciences in general—should be included in all engineering and computer science course syllabuses. Ultimately, the aim would be to produce graduates with the necessary technical expertise to be able to develop efficient systems and the skills in social sciences necessary for understanding the impact of their developments on society and on its citizens. On the basis of these criteria, various course models could be designed. A major/minor system could be put in place in higher education establishments, allowing students to choose a core subject, computer science for example (major), and a second subject such as Law (minor).

What about lawyers? We cannot leave the responsibility of ensuring that AI systems operate within the law to researchers and engineers alone. It is vital that legal professionals take on their fair share of this task. A precondition of this would be a genuine awareness of this issue within the legal profession and an alignment of the various courses available. Here again, the example of the major/minor system given above could be applied and the options changed to a major in Law and a minor in computer science.

Introducing a Discrimination Impact Assessment

In a certain number of cases, current European legislation requires operators who process personal data to first carry out an impact assessment to find out the potential impact of their activities on the rights and interests of those concerned: this is the privacy impact assessment or PIA. In this way, data carriers are responsible for self-assessing the impact of their activities, taking the appropriate corrective action and, in the event of an inspection, being able to demonstrate that all necessary measures have been taken to give them complete control over the process. This departure

from a system of prior authorization is a major paradigm shift towards agility, allowing manufacturers the scope for innovation. In this case, it would be advisable to capitalize on this approach, which incorporates the right to support for innovation by making a real commitment to equal opportunities in innovation in a digital era.

The guidelines adopted by the WP29, the Article 29 Data Protection Working Party, require a PIA to be carried out when data processing reveals a risk of discrimination or exclusion. This cornerstone in the social acceptability of AI is a matter for separate analysis. The PIA needs to be accompanied by a similar measure which can be applied in cases of discrimination, a discrimination impact assessment or DIA, to force creators of AI to consider the social consequences of the algorithms they produce.

An approach similar to the one that led to the design of the free software made available by the French Data Protection Authority (CNIL)—to assist those with less experience in carrying out their PIA auto-evaluation—could preside over the DIA measure. France could promote a joint investment project—through the EU's intervention or on the basis of voluntary partnerships with certain member states—to provide the necessary protocols and rights-free software. A line in investments could, in particular, be devoted to the engineering of this project (legal and operational support and facilitating the interface between the various competent authorities) so as to be able to implement the solutions identified by research.

3. Considering Collective Rights to Data

Developments in AI have revealed a certain number of 'blind spots' in current (and future, with the advent of the GDPR) legislation regarding the protection of individuals. They stem from the fact that the French Data Protection Act, like the GDPR, deals solely with personal data. However, although the scope for protection offered by this legislation is potentially very broad, artificial intelligence does not merely harness personal data. Far from it: many of the issues raised by the use of algorithms now constitute a 'blind spot' of the law.

Many of the issues raised by the use of algorithms now constitute a 'blind spot' of the law

Legislation relating to data protection only regulates artificial intelligence algorithms inasmuch as they are based on personal data and/or their results apply directly to individuals. This holds good in a large proportion of cases: personal offers, recommended contents, etc. but, in practice, many purposes escape this legislation, despite the fact that these may have a significant impact on groups of individuals, and therefore on single individuals. For example, it has demonstrated that the statistical aggregates that prompt sending a greater number of police patrols or Amazon couriers to certain areas may have discriminatory consequences for certain sections of the population, due to a mechanism which reproduces social phenomena.

From the point of view of developments in artificial intelligence, we could even simply ask ourselves whether the concept of personal data still has any real meaning. The pioneering work of Helen Nissenbaum teaches us, for example, that data is a contextual object which may provide information about several individuals or issues simultaneously. Especially since, within the context of deep learning, data is used on

a massive scale to produce correlations which could affect whole groups of individuals. Everyone has the right (with certain notable exceptions) to be informed in general terms about the fate of data which relates to them (purposes, subsequent uses, etc), even to object. But we do not have the option, neither *de jure* nor *de facto*, of prescribing or proscribing specific uses to which our data is put—except in the act of deciding whether or not to use certain services. At the moment, this power remains in the hands of regulators and legislators who, for example, restrict the grounds on which access to a range of services, to an insurance product, to housing, to work, etc may be refused. An individual may therefore be protected in a granular fashion against the collection of information that identifies him or her, but this protection does not cover the reticular configuration (on the network) that all information acquires.

Making class action effective

Several years ago, the European Union asked France to establish a more inclusive and workable system for collective action. Several of the measures that were adopted in recent years have widened and improved group access to litigation; in particular, the French Act for the Modernization of Justice in the Twenty-First Century introduced the 'personal data' class action which allows associations of consumer protection to act when infringements to existing legislation occur.

This class action is however extremely limited because it only serves to end a particular infringement and does not lead to the awarding of compensation for damages. Class action may be a lengthy and costly process: as things stand, it seems unlikely that associations could take on this type of process. Further, one can imagine the feelings of frustration experienced by injured users who would like to see action taken in court on their behalf and finally obtain… the end of an infringement but no compensation, despite the recognition of their status as victims. We are therefore proposing that compensation for injury sustained be included in this collective action.

4. How Do We Stay in Control?

Boosted by the progress in artificial intelligence, the big data revolution is contributing to the process of making the world more transparent, more quantifiable and infinitely more measurable. This revolution has been made possible by the conjunction of four factors: a massive reduction in the cost of processing information, the arrival of web 2.0 with its user-generated content, the exponential growth of data generated by humans and by machines, and the spectacular progress made in the use of algorithms. This sudden abundance of information has been particularly well received by public and private organisations that are susceptible to risk management. Its greater predictability allows them to be more efficient at anticipating the occurrence of injurious incidents, and therefore to take pre-emptive action in order to prevent these occurrences or at least to limit their adverse effects. In the banking sector, the probability of borrowers defaulting on loans can be more easily predicted, for example, and therefore the optimum amount of credit can be granted according to the level of risk that they pose. At least, this is the promise of big data, all the more since the advent of AI.

In addition to the banking and insurance sector, many other institutions—the courts, the police, the army, immigration—are beginning to make use of predictive analysis systems for a variety of purposes. In France, these scenarios remain largely hypothetical and the development of these initiatives is only at the experimental stage. However, certain foreign governments have already gone one step further; this is the case in Australia. In 2013, the Australian Customs and Border Protection Service installed a system for analyzing the terrorist threat posed by foreign passengers bound for Australia. This system, designed by IBM, cross-checks the data contained in passenger records against data held by the Australian Intelligence Services and social data available online, in order to establish risk profiles.

Following their example, law enforcement agencies could, in the future, rely on algorithms to manage the deployment of their patrol units and armies could use LAWS (Lethal Autonomous Weapons Systems) in operational theatres abroad. Changes of this nature, be it in the fields of health, banking, insurance or more particularly in the context of sovereignty, raise important ethical questions.

Predictive Policing

Police departments, initially in the United States and currently in Europe, are exploring the possibilities of using predictive algorithms within the context of their activities. These methods, commonly known as predictive policing, relate to the application of techniques for the prediction and analysis of big data for the purposes of crime prevention. In reality, they refer to two distinct applications: the first consists of analyzing geographical data in order to identify crime 'hotspots' where offences and crimes are liable to take place so as to increase surveillance in these zones and thus maintain a deterrent force. The second application relates more to the analysis of social data and individual behavior, for the purposes of identifying victims or potential criminals and being able to act promptly. These two applications are already being deployed in several American cities; French and European police services and *gendarmeries* are looking into the possibility of adding them to the tools they use in crime prevention.

The earliest research available on their impact in the United States would recommend proceeding with caution. Predictive policing and legal solutions are not only subject to important technical limits but may equally prove to be infringing fundamental liberties (privacy and the right to a fair trial).

On a purely practical level, we need to bear in mind that, sophisticated as they are, these systems remain fallible; they are capable of making errors, with potentially disastrous consequences for the lives of the individuals they wrongly assess.

The Propublica enquiry

In May 2016, journalists from Propublica (an American investigative newspaper) revealed that the COMPAS algorithm used in the estimation of the risk of recidivism by the American legal system and developed by the Northpointe company =, was racist and inefficient. An analysis of the scores attributed to prisoners revealed that this algorithm systematically overestimated black American prisoners' risk of recidivism at twice that of white Americans. In addition, the latter were often

represented as presenting a low risk, which was inconsistent with their actual rate of recidivism.

This means that this algorithm resulted in the continued detention of black prisoners who would probably not have re-offended (false positives), whilst it allowed white potential re-offenders to go free (false negatives).

Amongst other things, the Propublica enquiry reminds us that we are not all equal when it comes to these systems. Since the COMPAS algorithm was trained with data from police and judicial databases, it is liable to be biased and to reproduce the prejudices currently found in society. The absence of a critical distance in its use could lead to the entrenchment of discrimination in the law and the systematic dissemination of prejudice.

We should also consider the impact of these solutions on those who may be required to implement them—in this case, judges and police officers. Indeed, the increased use of these technical solutions will lead to an increased pressure to standardize the decisions made by institutions: it is far easier for a judge to follow the recommendations of an algorithm which presents a prisoner as a danger to society than to look at the details of the prisoner's record himself and ultimately decide to free him. It is easier for a police officer to follow a patrol route dictated by an algorithm than to object to it. In both cases, they would be obliged to defend their 'discretionary' decisions and in these circumstances, it would be preferable if their approaches or decisions were in line with standard procedure. However, the outcome of this move is very uncertain and there are concerns that it would raise increasing challenges to their individual responsibility. On the other hand, these systems would not be vulnerable to the strain of decision-making which sometimes results in judges freeing fewer prisoners at the end of the day than during the morning…

Another danger linked to the proliferation of systems for predictive analysis is the increased threat of mass surveillance. For predictions to be as accurate as possible and to optimize decision-making, these systems need to have access to as much information as possible, at the expense of individual privacy. More fundamentally, these systems are liable to reduce individual autonomy by encouraging judges to detain prisoners who have already served their sentences or by organizing the systematic surveillance of populations in deprived areas.

Regulating the use of predictive algorithms

To prevent these situations arising, citizens should first of all be informed about their rights: in these two instances, the right to an effective remedy and the right to explanations concerning the processing of data on which surveillance is based. From this point of view, we need to remind ourselves that in 1978, the French Data Protection Act laid down the principle according to which 'no court or other decision involving legal consequences for an individual can be taken solely on the basis of the automated processing of personal data intended to define the profile of the person concerned or to assess certain aspects of his personality', adding that 'an individual has the right to know and to challenge this information and the logic underlying the automated processing when these results are denied him'. These

conditions, extended by the Act of 6 August 2004, demonstrate that the legislator had anticipated—at a very early stage—the pitfalls inherent in such systems. The substance of these conditions has been reiterated in Article 22 of the GDPR.

Secondly, it is vital to ensure that, at any point in the discussion, responsibility can be attributed to a human being via a predetermined procedure. Various scenarios would be worth studying, ranging from those involving individual responsibility (from the individual who makes the decision to the creator of the algorithm or the technology in question) to those involving devolved responsibility.

Finally, developments in this technology should lead us to consider the role automation should play in decisions made by human beings. Are there any areas where human judgement, fallible though it may be, should not be replaced by machines? If so, we should consider taking steps to protect these immediately.

Lethal Automatic Weapons Systems

One of the greatest concerns regarding developments in AI is the subject of lethal autonomous weapons systems (LAW). This is not a new discussion: indeed, France initiated it in 2013 within the UN Convention on Certain Conventional Weapons (CCW) which led to the creation of a group of government experts whose first session was held at the end of 2017.

Aside from the problems of reaching an international agreement on a military issue that is as sensitive as it is strategic, the discussion has been complicated by the question of defining the boundaries of LAW, especially since up until now this weapons system has not actually been implemented…at least not officially. This is where the first obstacle presents itself: just as with AI, it is difficult to draw a clear line between what is and what is not autonomous and we are in fact obliged to envisage a continuum between the two, with different degrees of autonomy. We must steer a course between, on the one hand, a definition of LAW which is too inclusive, which represents an obstacle for proponents of the need for regulations and threatens to undermine existing capabilities or the development of sophisticated capabilities; and on the other, a definition of LAW which is too exclusive and which would not cover any of the relevant systems.

Research into performance is, however, a necessity when we are confronted by increasingly capable competitors and increasingly complex tactical situations involving increasingly sophisticated systems. From a French point of view it is, however, possible to be a driving force behind proposed regulations or the development of good practices without having to forego advanced capabilities *ex ante* or fall behind other States in this important strategic domain.

From automation to autonomy

Developments in weapons systems are in many respects comparable to those made in the motor sector where vehicles are progressively moving towards autonomy. The first successes involved functions which were complicated to use and still required the driver to perform specific actions: changing gear, operating the headlights, indicators, cruise control, etc. The automation of these functions makes them less complicated to use without reducing the role played by the driver, who no longer

needs to be concerned with the mechanical details of driving the vehicle. Similarly, steering correction and avoidance systems allow vehicles to protect their drivers and third parties by reacting automatically and much more rapidly in situations that have previously been identified.

There is no strict definition of an autonomous vehicle; it can be identified on several defined levels[1]. To avoid the pitfalls of an inevitably imperfect definition, it makes sense to establish a scale of autonomy: from landmines to remotely operated and automatic anti-missile defense systems, etc. This would make it easier to target areas of technology needing attention by excluding from consideration those which are not concerned by the developments in AI (landmines in particular, which are frequently cited as an exception) or those for which automation is only relevant for performance requirements. According to such a scale, remotely operated and controlled systems, anti-missile defense systems, torpedoes, navigation and guidance systems and surveillance and detection systems could not be regarded as LAW.

Drawing up this scale would be primarily for educational purposes and make for less heated discussions. In fact, we are still waiting for the technological breakthrough we need to achieve what would amount to an equivalent of Level 5 autonomy in the case of autonomous vehicles, i.e. a capacity similar to that of a human being to adapt to any situation that may arise and react accordingly (for many experts, this is still a long way off and very unlikely). Yet it is this level of autonomy which appears to be of major concern to the general population.

The French perspective

France accepts that mankind is ultimately responsible in the use of lethal force. The major developments which involve AI techniques relate to assisting those taking and implementing decisions rather than replacing them. In this respect, this means relieving human operators of time-consuming and relatively unimportant duties so that they can concentrate on more important tasks. This may also contribute to improving response time in situations where speed of execution is decisive.

We need to remind ourselves that all weapons systems—whether developed, acquired or adopted—are subject to international and humanitarian laws: they undergo tests for legality and conformity to international law in accordance with Article 36 of Protocol I Additional to the Geneva Conventions. France has put forward a proposal[2] to improve transparency and confidence on this point where it relates specifically to LAW.

France was a driving force in initiating international dialogue in 2013 and must continue to play a major role in defining the regulations and guides to good practice that need to be established at an international level. In particular, our country could explore the options worth pursuing in terms of technological solutions for use in determining the level of interaction between humans and machines required, for example in the development, deployment and use of emerging technology; in the

1. From 0 – 5: from an entirely manually-operated vehicle to complete autonomy with capacities similar to those of a human being
2. See the declaration of 15 November 2017 made by the French permanent representative to the Conference on Disarmament.

periodic revisions of systems (for example, in the case of self-learning technology); or alternatively in the addition of a means for self-destruction and abandonment of missions.

Although international dialogue is intended to be pursued within the CCW, it is equally important that these issues should be the subject of an international ethical debate which brings together not only technical experts, but also civil society and NGOs.

Establishing an observatory for the non-proliferation of autonomous weapons

Unlike other types of dual technology where the situation is the reverse, in the field of AI the civil element brings the military element in its wake; this leads to problems concerning its appropriation and adaptation. Thus the issue of proliferation has to be addressed in a context where the technological building blocks required to build weapons are no longer supplied by the military, and where anyone with a grounding in AI can divert its purpose into building weapons for the arms trade.

France has existing regulations which allow it to maintain control over military equipment. According to information published by the French General Directorate for International Relations and Strategy of the Ministry of Defense:

The French system for monitoring military equipment is based on a general prohibitory principle, according to which the whole of the sector is subject to State control; the power behind this is the CIEEMG, the Inter-ministerial Committee for the Study of Military Equipment Exports. The CIEEMG brings together representatives from various ministries including those in charge of defense, foreign affairs and international development, and the economy and finance, who have the right to vote on rulings. It reports to the prime minister and is chaired by the SGDSN, the General Secretariat for Defense and National Security. It assesses all aspects of export initiatives, taking into particular consideration the impact of an individual export on regional peace and security, but also the internal situation of the country of final destination and the practices of the latter in terms of the respect of human rights, the risk of its misuse for the benefit of unauthorized end users, the need to protect the security of our troops and those of our allies or alternatively to control the transfer of the most sensitive technology.

Concerning AI, the issue of proliferation needs to be addressed in a context where the technological building blocks required for the building of weapons are no longer supplied by the military but are developed by private stakeholders for purely civil applications. It should therefore be noted that anyone with a grounding in AI could divert its purpose into building weapons for the arms trade: at the moment, when a detection is made by an algorithm and this triggers a response from a computer, the additional complexity of turning this into a physical response serves no purpose.

In this context, an observatory could be put in place—along the lines of the observatory for the non-proliferation of nuclear, biological and chemical weapons—which would have an ongoing prospective and monitoring role concerning lethal autonomous weapons and the threats they pose.

5. Specific Governance of Ethics in Artificial Intelligence

As development in AI grows, so too does interest in ethical issues and it's now a topic on everyone's lips, from researchers and unions, to associations and businesses (both large and small). Numerous private actors are either already involved or becoming involved in voluntary initiatives concerning in-depth consideration or development of ethics charters.

Two years ago, the French Law for a Digital Republic entrusted the CNIL with the task of reviewing ethical issues and societal questions raised as a result of the development of digital technology. The CNIL chose to respond swiftly by initiating a decentralized cycle of public debates, workshops and meetings which provided the backbone for an outstanding report[3] published last December. Parallel to this, private sector AI giants are seeking to position themselves with respect to every aspect of the global debate, increasing, over the past several months, the creation of ethical think tanks centered on the technology they implement.

The role of ethics in the debate on AI has become so significant that it seems necessary to instate a national advisory committee on ethics for digital technology and artificial intelligence, within an institutional framework. Such a body could be modelled on the CCNE (Comité consultatif national d'éthique - National Consultative Ethics Committee), created in 1983 for health and life sciences. As separate bodies, both institutions could nevertheless study and provide joint opinion on issues to emerge at the crossroads of their fields of expertise with respect to transhumanism, biohacking or the processing of AI data on health, for example.

The ethics committee for digital technology and AI would be responsible for coordinating public debate in an accessible and constructed way within a legal framework. The committee would need to express reasoning on short-team perspectives such as industrial and economic issues, ensuring effective interaction with sectorial committees, whilst also being able to step outside of this mindset in order to take account of long-term perspectives. Forming such a body would not only ensure a high level of expertise, but also independence in terms of special interests.

Independently developed opinions of the committee could provide clarity on technological choices made by researchers, and industrial and economic actors. We could draw inspiration from Germany in this regard, who recently established an ethics commission responsible for ruling specifically on driverless cars. The commission published its first report last August[4], in which recommendations made can serve as benchmarks for resolving ethical dilemmas, and therefore as guidelines for the programming of driverless vehicles. This new committee must also be able to advise the State on its own technological choices: whether at national level (such as choices made by the State concerning the use of AI for surveillance, etc.) or at international level (France's position on autonomous weapons).

3. How can humans keep the upper hand? Report on the ethical matters raised by AI algorithms, published 15 December 2017

4. Report available at the following address: https://www.huntonprivacyblog.com/wp-content/uploads/sites/18/2017/06/084-dobrindt-bericht-der-ethik-kommission.pdf

As is the case for the current CCNE, this commission could be officially called upon by the President of the Republic, members of the government, Presidents of Parliamentary Assemblies, higher education institutes or public establishments; or may decide to act on its own initiative on subjects corresponding to its area of expertise. Nonetheless, this concerns pushing current boundaries and envisaging effective social outreach.

Placing emphasis on social outreach

Alongside the possibility of institutional consultations, public consultations could also be considered. Nevertheless, the technicalities must be defined. The commission could also include members from civil society and public representatives able to participate in the examination of topics, as much as in debates and setting the agenda.

The current CCNE has initiated an interesting approach to social outreach. The Bioethics Law, voted on July 7, 2011 effectively tasked the CCNE with the coordination of general assemblies prior to reforms envisaged for ethical and societal issues. The aim is to encourage citizens to participate in ethical reflection by facilitating their understanding of issues in respect to scientific progress: 'general assembly citizen committees' are therefore composed of a representative sample of French citizens tasked with giving their opinion on topics raised. This approach could be replicated.

Sustaining ethical debate in society

The hive of activity currently surrounding the question of ethics must be encouraged and strengthened. This is why, beyond its initial responsibilities, the committee should be tasked with coordinating and sustaining ethical debate in society by organizing events, holding public consultations both online and off line, making tools and assistance available for the coordination of autonomous debates, carrying out surveys and opinion polls on the various issues, etc.

Lastly, the committee could capitalize on the wide range of initiatives to have emerged in recent months such as union charters, corporate charters, non-profit charters and research work, where various philosophical and scientific approaches, various spheres of legitimacy and expertise on the topic cross over and intersect. The feedback provided by these initiatives is invaluable for achieving more generalized reflection. The commission could be responsible for logging feedback, mapping specific concerns that drive it, but also for enhancing or even certifying it in order to provide elements that other actors in search of guidance on best practice may find useful.

An international debate

A number of international actors, both public and private, have initiated a debate on the ethics of AI. For instance, the French researcher Yann LeCun has been behind an ethical partnership between very large players on this theme; Deepmind has an ethics department; the United Kingdom has already announced the establishment of a national ethics committee. At the European level, certain already imagine a

network of national ethical committees, modeled on the "G29 network" (network of Data Protection Authorities). At the global level, Quebec has just proposed the creation of a international agency that could be housed in Montreal, such as the International Anti-Doping Agency. At the same time, Unesco has launched an international reflection. All these efforts must be encouraged.

Part 6 — For Inclusive and Diverse Artificial Intelligence

Part 6 — For Inclusive and Diverse Artificial Intelligence

Artificial intelligence cannot become another driving force for exclusion: this is a democratic requirement within a context where it is set to become one of the keys to the future. Artificial intelligence provides a vast amount of opportunities for generating value and developing both on a societal and individual level. These opportunities must be of benefit to everyone, and first and foremost to women.

Indeed, almost half of the world's population are women but they represent a mere 33% of those within the digital sector (and only 12% if we dismiss cross-sectional and assistant roles). In order to face the challenges posed by AI, it is important to call on the plurality of expertise. Collective action must be as inclusive as possible. Everyone should have equal access to opportunities to participate in research, development and value creation in AI. As such, the key challenge consists of breaking down barriers and distributing innovation skills.

Almost half of the world's population are women but they represent a mere 33% of those within the digital sector

On a broader level, faced with the scalability of technology and functions linked to AI, our society is bound by a duty of reflexivity and collective vigilance. This is particularly relevant in terms of vulnerable groups in society and those who are already excluded from the digital world, to whom AI may represent even greater risks. In the same vein that AI developments may promise a better society; one that is fairer and more effective, they could also cause the hyper-concentration of profit value for a small, digital elite.

Educating larger numbers of people on the principles of AI and algorithms as a precondition for an inclusive policy is vital. The introduction of a subject dedicated to digital humanities, recently announced by the Minister for National Education, could provide assistance in this regard.

An inclusive policy for AI must therefore incorporate a dual objective. First, to ensure that the development of AI technology does not cause an increase in social and economic inequality. Second to call on AI in order to reduce this. Rather than jeopardizing our individual trajectories and solidarity systems, AI must first and foremost help us to promote our fundamental rights, improve social cohesion, and strengthen solidarity.

1. Gender Balance and Diversity: Striving for Equality

The feminization in scientific and technical sectors is slow but still progressing. At the opposite, the digital sector has not followed suite: gender balance is far from being achieved (*see inset*). Beyond issues concerning competition and performance, gender balance and diversity are societal issues. As digitization becomes omnipresent in our lives, soon to be followed by artificial intelligence, this lack of diversity can result in algorithms that produce cognitive biases in programme design, data analysis and results interpretation, which often go unnoticed (see section dedicated to ethics).

There are numerous examples of this: for example, a number of programmes link words such as 'programming' with 'man' and 'household tasks' with 'woman'. When 'CEO' was entered into a search engine in the United States in 2015, only the 96th photo was of a woman—and even then, the photo was of 'Barbie CEO' dressed in a miniskirt[1]! When the word 'CEO' is typed into a search engine today, the vast majority of photos continue to depict men, despite the fact that almost a third of American CEOs are women.

As such, one of the major challenges posed by AI consists of achieving better societal representation. The prerequisite remains to educate people in equality from a very early age, which must implicate parents, private businesses, the media, associations, and, naturally, all actors within education. From childhood right through to computer science and engineering faculties, educational establishments must foster a culture of equality between both sexes via teaching, educative activities, training and educational material, and ensure that information relating to careers and training pathways is free from any form of gender stereotyping.

Women in the digital sector and engineering careers in France

In 2016, less than 10% of students studying computer science were women, whilst between 1972 and 1985, the percentage of women attending these establishments was higher than that of all other types of engineering establishments.

Within the economy, the percentage of women in the digital sector is woefully low: 33% of workers in the digital sector are women, and if we dismiss cross-sectional and assistant roles, this figure falls below the 12% mark. Additionally, only 11% of workers in cybersecurity are women.

Beyond this comprehensive indicator, it is important to focus more specifically on the roles of female engineers in companies, on their hierarchical positions and their salaries. These are two of the key factors in the appeal of these professions and the attrition rate, the rate at which female employees leave companies after starting at them. Less than 10% of engineers belonging to executive committees or boards are women.

The average pay gap between men and women engineers stands at 30%, exceeding 34% for those aged over 45 years old. The pay gap is four times higher in software firms or engineering companies. Findings within innovative entrepreneurship are equally concerning: only 9% of French startups are founded by women. Last but not least, on average, women raise less than twice the amount of funds than men…

Sources for figures: publicity campaign for "Women and the Digital Industry" by the Centre Francilien pour l'Égalité Femmes—Hommes (Regional Observatory for gender equality), The Hubertine Auclert Centre, study by Syntec and OPIIEC, The economic and social performance of digital startups in France, 2015, study by Mutationnelles, 2014 and startup study Ernst and Young, 2015.

1. Daily Mail: "The first woman to appear in a Google search for 'CEO'? BARBIE... and, of course, she's wearing a miniskirt" http://www.dailymail.co.uk/femail/article-3043673/The-woman-appear-Google-search-CEO-BARBIE-course-s-wearing-miniskirt.html

A second prerequisite is digital education in schools. Currently, training specifically focused on information technology is included in primary and secondary curriculums, but it remains non-compulsory and often insufficient. Beyond being a discipline, digital technology is an educational method which includes course content, methods used, the construction of knowledge and sciences, and even relationships between actors and system coordinators. It concerns establishing effective digital education founded on decompartmentalization and transversality.

Understanding the under-representation of women in science, technology, engineering and mathematics (STEM) teaching: gender bias in the classroom

In 2016, Member States of the United Nations ruled on the role of UNESCO, to encourage women and girls to exercise leadership in science, technology, engineering and mathematics (STEM).

A report by UNESCO published in 2017 responds directly to this request by decoding the factors that hinder or facilitate participation, success and retention of girls and women in STEM education, and notably, what the education sector can do to promote their engagement and interest in these subjects. Generally speaking, the report found that girls seem to lose interest in these subjects in line with their age, particularly between the beginning and end of their adolescent years. This decline in interest affects participation levels in the study of science at secondary level.

Regarding the mastering of software tools, a study carried out in 2013 found that self-confidence levels were lower amongst 12-year-old girls, even within areas in which they outperform boys. The report also cited a study carried out in Vietnam which confirmed that girls approach computer technology with the mindset that programming is difficult. That being said, as they overcome this notion, their skills in programming are improving and they often outperform boys.

Source: UNESCO, Cracking the code: girls' and women's education in science, technology, engineering and mathematics (STEM)

In order to ensure effective gender equality and digital technology education, teaching and pastoral staff must receive training in these areas if they are to encourage significant numbers of young girls to head for the digital sector. School heads could be held responsible for the successful implementation of educative policies promoting equality and digital technology.

Foreign initiatives to teach girls how to code

In India, the social initiative @IndianGirlsCode provides free coding and robotics programmes for young, underprivileged girls. This initiative encourages girls to become innovators in the fields of computer science and technology, and helps them to learn to code and innovate by creating applications designed to resolve daily problems.

In the United States, the non-profit organization Girls Who Code aims to educate, empower and equip teenage girls with skills and resources to pursue opportunities in technology and engineering. Training is delivered through free after-school clubs or intensive summer schools. Over 10,000 girls have participated in the programme, of which many are now studying computer science at the top American universities.

Incisive Action: Ensuring 40% of Students on Digital Courses are Female

The decline of women in computing professions is an alarming" phenomenon and one which continues to escalate despite efforts made in the business, education, and non-profit sector to encourage diversification in the career choices of girls. This observation is a unanimous one: we are experiencing a crisis in the lack of females choosing to study on top courses in digital technology. If we fail to act, a large part of society will miss out on this new economy.

Now is the time to take definitive action to reverse this trend. If educating people in equality and digital technology is a prerequisite and essential condition, gender balance could be achieved by implementing incentives for achieving a female enrolment rate of 40% for digital subjects (preparatory classes and courses in *Grandes écoles*—prestigious institutions outside of the public university system—and universities) by 2025.

A positive incentive policy could be put in place in order to achieve this objective by 2025. In this way, if academic establishments were to swiftly achieve a female enrolment rate of 40%, they could be rewarded with an accreditation or grant.

Promoting courses fully committed to gender equality in the digital field: Grande École du Numérique (GEN), in Paris

The Grande École du Numérique aims to promote gender equality in the digital sector and ensure women have access to opportunities on offer within the field. Accredited courses are therefore tasked with ensuring at least 30% of their student intake are female.

As an example of this, Web@cadémie launched its programme '*Ambition Féminine*' (Feminine Ambition), with a majority female intake. The Grande École du Numérique also promotes courses that enable mothers to enroll thanks to family-friendly timetables. The development of female mentorship, such as initiatives by Social Builder, is promoted with a view to obtaining the accreditation.

Source: GEN, "Favoriser la mixité dans le secteur du numérique" (Promoting gender balance in the digital sector), February 2017

The shift in the number of girls studying computer science and engineering is far from unrealistic if supported by a framework that enables the inclusion of women in digital professions and change in cultures and practices to be envisaged. In this respect, it is interesting here to reflect on examples of foreign higher education establishments that have succeeded in significantly improving the number of female

students and researchers on digital courses. It has been observed that mentorship programmes for girls in computer science gives them a major advantage, improving their attendance and confidence levels in technical and scientific studies and careers.

Programmes for women in a number of foreign computer science schools

The Norwegian University of Science and Technology (NTNU) and Carnegie Mellon University (CMU) launched a large-scale operation designed to reintroduce gender balance in their courses after observing a decline in the number of females studying computer science as of 1995.

Researchers Chantal Morley and Isabelle Collet analyzed such programmes. In 1996, NTNU launched the initiative 'Women in Computer Initiative' (WCI), under the joint leadership of the newly elected Vice Chancellor, the Vice Dean of the Faculty of Physics, Maths and Computer Science, and the Director of the Department of Computer Science. A year later, the percentage of female students had risen from 6% to 38%, and the WCI committee became a permanent fixture. At Carnegie Mellon University (CMU), an action-research programme was launched by a duo composed of the Vice Dean—a professor in computer science—and an expert in gender and education science. The number of students grew progressively and reached 39% in 2000 (compared to 7% in 1995), with a comparable drop-out rate for both genders. The CMU programme is still active today, with a mentorship programme connecting women across the faculty's various departments. MORLEY Chantal et COLLET Isabelle, "Femmes et métiers de l'informatique : un monde pour elles aussi", Cahiers du genre, 2017, no. P. 183-202).

Meanwhile, the Stanford Artificial Intelligence Laboratory has developed a free, two-week summer school designed to train college students in AI. The summer school was organized by volunteer teachers and graduates and was judged to have had a positive impact.

In the United Kingdom, the Athena SWAN Charter launched in 2005 aims to boost the representation of women in science, technology, engineering, medicine and maths. Organizations can apply to be awarded with the prize in recognition of their commitment to equality, diversity, and progress in this regard. The programme has had a positive impact on gender balance in participating institutions.

Initiatives must be backed up with a training and awareness policy for educators on this issue to help them identify biases and encourage them to better guide young women towards these subjects. School heads should be held responsible for the successful implementation of educative policies promoting equality and digital technology.

National Initiative to Promote Gender Balance in Technology

All initiatives advocating diversity in digital businesses could be supported by a national initiative to promote gender balance and diversity in technology. Co-developed with all actors in the sector, it would have a clear and ambitious aim, such

as increasing the number of women in the digital sector by 30% within the next 2 years. The plan could be launched via a national event and should call on the support of existing associations, by giving them more coordination and networking resources. There are a number of very active associations enabling women to fully assume their role within the digital sector (see inset). Unions are also active in this area.

There are a number of active bodies at both national and regional levels seeking to contribute to the impetus of a national initiative promoting gender balance in digital technology, such as the HCE (Haut Conseil à l'Égalité entre les femmes et les hommes - High Council for Gender Equality), the CSEP (Conseil Supérieur de l'Égalité Professionnelle entre les femmes et les hommes - Higher Council for professional equality between men and women), and also the Hubertine Auclert Centre (center for gender equality based in the Paris region).

These networks would be composed of male and female ambassadors who would notably be expected to speak in schools, colleges and higher education establishments and sponsor new arrivals (both male and female) within their organisations.

Examples of associations set up to help women fully assume their role within the digital sector

In France, the associations Girls in web, Duchess France and Women in Tech work to bring about change in this field and to ensure women have access to opportunities within the digital sector. As such, Girls in web organizes monthly events such as master classes and round table discussions, acts as a network, and forges partnerships always with the aim of making women more visible in the digital world and increasing their share in the economy.

Source: from examples in the information report no. 3318 by the French National Assembly, 'Women and Digital technology: overcoming obstacles, seizing opportunities', Delegation on Women's Rights, by Chair Catherine Coutelle, December 2015

A major national campaign could be launched simultaneously in order to raise public awareness of human needs within the digital and AI sectors in a bid to emphasize digital professions and their accessibility to women. More precisely, an information campaign aimed at changing masculine culture within the digital world and fighting against exclusion and self-exclusion mechanisms would be launched. This campaign could be focused on highlighting decision-making biases, integration bias and self-censorship tendencies amongst women.

What are the consequences of hyper-masculine culture on the careers of women?

A number of studies have addressed the high attrition rate (the departure rate) of women in the STEM sectors. In this respect, in the United States, only 25% of women continued to work in the sector ten years after graduating from STEM courses

(source: *Women in STEM: Realizing the Potential, STEMconnector white paper*, March 2014).

After working in the sector for ten years, 41% of women left compared to only 17% of men (source: *Women in IT: The Facts, National Center for Women and Information Technology*, 2010). Understanding the causes behind these departures is difficult. Researchers have shown that the attrition rate peaks mid-career, at 35 years of age, following one or two promotions (source: see *The Athena Factor: Reversing the Brain Drain in Science, Engineering, and Technology*, 2008 and *Women in IT: The Facts, National Center for Women and Information Technology*, 2010).

Beyond this period, women often feel that they are unable to progress further (source Women in Tech survey carried out in the United States on a sample of 1000 women in the STEM sectors, November 2014). They are reported to feel isolated in predominantly male teams, without a mentor, sponsor or project. They leave as a result of reportedly feeling as if they have been treated unfairly, paid less, and are less likely to progress compared to their male counterparts.

In 2017, tech giants were particularly affected by scandals related to sexism. Google was at the center of a controversy after a male engineer justified the lack of female representation in tech due to "biological causes" in a memo sent internally. At Uber, engineer Susan Fowler drew attention to inadequate consideration of sexual harassment cases within human resources, as well as of other forms of discrimination and humiliation occurring daily.

Equally in 2017, Ellen Pao published her book 'Reset' in which she recounts her experience of male chauvinism in Silicon Valley, with discussion on the subject showing no signs of slowing down. In 2018, Emily Chang published her book 'Brotopia: Breaking Up the Boys' Club of Silicon Valley', in which she tells of Silicon Valley's 'e-parties'; in other words, highly sexualized parties to which women were invited based purely on their physical attributes.

Meanwhile, a number of studies have shown the positive impact that role models could have on encouraging girls to study scientific and technical subjects. With this in mind, campaigns could highlight examples of successful female role models. This has notably been proposed by the European Commission in its recently unveiled action plan for digital education: to mobilize stakeholders to provide girls with inspiring female role models to which they can identify. These female role models could be women of our time who are achieving distinct success in the digital and AI sector, or historical figures (see inset).

The role of women in computer science development

The lack of female representation in the digital sector since the 1980s can be partially explained by sociohistorical factors, and particularly by the fact that stereotypes of computer science experts and 'geeks' are increasingly echoed in social consciousness. This consciousness and the absence of female role models have a significant influence on both girls and boys. However, computer science has not always been a male-dominated field: women were the pioneers. The first computer programme was developed by Ada Lovelace in 1843. The first PhD in

computer science was earned in the United States by Mary Keller. The first language processor to pave the way for programming languages was created by Grace Hopper. Women were responsible for building the first fully electronic computer in 1946, and the moon landing was managed by teams led by Margaret Hamilton.

Implementing a National Database on Gender Inequality in the Workplace

A number of studies focus on identifying the factors which cause women to leave the digital sector. From work coordination, internalized collective bias in interactions, invisibility phenomena to difficulties receiving promotions, the possible causes are numerous. However, before beginning to tackle any one of them, it is crucial to obtain more accurate data on male-female discrimination at play in the workplace.

Whilst computing and AI domains are unappealing to women for reasons that can be difficult to formally establish, they can also be unappealing due to reasons that can be indicated in a more objective way. It would therefore be useful to build a database which enables them to be identified, both in this sector and in others.

Quantifying workplace inequalities

Contribution to this database could become mandatory, in the same way that CSR reporting requirements apply to large companies. This database would therefore enable year-to-year progress to be measured and provide a course of action for public policies. The objective of the database is not to condemn particular companies, but rather to drive forward collective reflection and public analysis. Data made public would therefore be anonymized. Businesses to have contributed to the database could promote this fact and highlight their ambitions for diversity (see inset). The database should also be fed by public authorities in order to foster its exemplarity in the field. The database should also be demanding when it comes to selecting indicators:

- gender balance rate in executive committees and board of directors;
- gender balance rate in teams;
- gender balance rate in appointments, promotions and recruitment;
- gender balance rate in terms of grade and job type;
- pay gap between different jobs, at different grades;
- entrepreneurial support and the establishment of women's networks.

indicators produced by French union Syntec Numérique on the feminization

Founded in July 2011, the programme 'Femmes du Numérique' (Women of Digital Technology) launched by Syntec Numérique aims to promote gender equality within the digital ecosystem and highlight the benefits of a career in the sector to young girls. To this effect, digital technology has established benchmark indicators in order to monitor efforts accomplished within the digital sector in view to promoting gender equality in the workplace. This survey could therefore provide elements within the framework of the national database proposed.

This database would enable measurement of the impact of initiatives undertaken to address equality and diversity. Production of these indicators and governance of the database could be led by the CESP (Conseil Supérieur de l'Égalité Professionnelle - Higher Council for professional equality between men and women) within the framework of its task to assess and monitor policies on equality in the workplace.

Promoting Transparent Recruitment and Promotion Processes

It is crucial to back-up the data policy with a policy for transparent recruitment and promotion processes. Whilst large companies often have such processes in place, this is less common in small businesses and startups. An awareness-raising initiative on transparency in human resources could be implemented in conjunction with FrenchTech. These initiatives could be backed up by the provision of advice to young businesses lacking skills in HR.

Carrying out a survey enabling effective measurement of gender discrimination in the digital sector

Two female entrepreneurs recently carried out an experiment in which they teamed up with a male business partner, and in doing so observed a significant shift in consideration amongst financers and investors[2]. This interesting and alarming result deserves to be taken seriously.

Assessments could be carried out potentially via A/B testing which would measure the difference between funding obtained with and without a male business partner. All other factors being equal (qualifications, value proposition identity, etc.), the survey could facilitate improved measurement of the existence (or absence) of financial and recruitment biases and indeed any other decisive factors for career progression and the perception of women—by either women or men—in their relationship with ambition, competition and money.

Promoting AI research in support of identifying discrimination

Solutions capable of identifying workplace discrimination already exist. By analyzing surveys, employee statements, data on salaries and promotion, etc., the company Palantine Analytics was able to detect a number of biases working against women. It would be interesting to promote research which would enable similar solutions to be developed.

Setting Aside Funds to Address Diversity

Funding for projects working on the development of inclusive and non-discriminatory AI could be established at BPI France or FrenchTech, and equally for digital businesses working on projects with high social and environmental impact

2. Le Monde, "Comment deux entrepreneuses s'inventent un collègue masculin pour convaincre les investisseurs" (How two women invented a male colleague to win over investors), 8 September 2017.

and that are committed to diversity. Businesses that receive funding must be particularly active in promoting diversity within their teams, and must commit to holding talks in schools.

Study and research bursaries for women as well as individuals from minority or socially disadvantaged backgrounds could first and foremost be awarded by the private sector. For example, the L'Oréal Foundation rewards women scientists and highlights their work via their programme 'For Women in Science'.

2. Developing Digital Mediation and Social Innovation so that AI Benefits Everyone

It is our collective responsibility to ensure that nobody is marginalised

Given the extent of transformations on the horizon as a result of AI, it is our collective responsibility to ensure that nobody is marginalized. In order to ensure that everybody is able to benefit from advancements in AI, we must develop procedures concerning access to rights and significantly strengthen our capabilities for mediation. Likewise, opportunities for innovation using AI must permeate all fields of activity, including social policy and care systems.

Enabling Access to Fundamental Rights and Public Services

Over the past several years, the number of reports warning administrations of the risks posed by the reduction of access to public services and fundamental rights as a result of digitization have multiplied. These trends are especially serious since they affect a substantial part of the population, particularly those who are in precarious situations and/or alienated from digital technology (see inset), and also since they are likely to intensify over the coming years.

12 million French citizens experience difficulties carrying out common administrative procedures

According to an investigation by the French Defender of Rights published in February 2017, approximately 12 million French people experience difficulties performing common administrative procedures, such as declaring taxes online and downloading or completing online forms. These difficulties result in 12% of users abandoning procedures, which notably concern the legal system (36%), the *Trésor public* (Inland Revenue) (14%) and social security bodies (13%).

In order for the transformation of administrative procedures using AI to improve access to rights rather than encourage polarization, the strategy used by administrations must be twofold. On the one hand, it must be focused on improving public knowledge of administrative rules and their application in personal circumstances, to the extent of automating certain recurring procedures. On the

other hand, it must enable the implementation of new skills in mediation and cross-sectional care for those who need it, in connection with active support networks.

Creating an automated support system for the management of administrative procedures

The lack of information (difficulty accessing information, contacting someone, or the simple unawareness of possible means of recourse), friction points in online procedures (repeated requests for supporting documents, large number of information portals and points of contact, etc.), the non-processing of requests (whether due to omission, error, delay or lack of resources) and the feeling that procedures are pointless largely explain abandonment or the non-take-up of rights.

To resolve this, public authorities must rethink the design of administrative procedures with a view to helping citizens feel more able to approach public services for support. For this purpose, public policies could incorporate artificial intelligence in order to absorb complex administrative procedures, as well as personalize and simplify user experiences for public services (access to high-quality and contextualized cross-sectional information concerning a number of administrations, delegation of certain recurring administrative tasks, etc.).

An open challenge to develop an artificial intelligence-based platform used to manage and perform administrative procedures could be launched.

- The platform could notably help users express their needs and requalify them in administrative terms with the help of natural language processing techniques; draft preliminary assessments based on comparative analysis of similar situations; provide users with personalized and contextualized information to help them perform administrative procedures, or even manage the completion of certain administrative tasks; and redirect users to additional online support or expert help.
- As a first step, the platform could focus on the most simple and recurring administrative procedures; those that are most sought out by users (number of cases) and those that have the highest rates of abandonment or non-recourse of rights (Inland Revenue for tax collection - 12%, the CPAM [caisse primaire d'assurance maladie—public health insurance body] - 8%, the CAF [Caisse d'allocation familiale—Social security office] - 4%, followed by the prefecture [4%] for the allocation of residence permits for foreign nationals, management of commercial and private vehicles and driving licences, identification and registration of associations, and the coordination of demonstrations);
- The platform's architecture and operation should be imperatively focused on user experience: using UX Design experts, coordinating user journey testing, incorporating support from actors in mediation, etc. The aim is to take better account of friction points encountered by users and to resolve them by incorporating more intuitive design of both the interface and interactions.

Drafting of a mediation grid required for digitized public services

Coordinating the possibility for human intervention is essential in order to ensure user confidence in partially or totally automated administrative procedures. Long-term, if tasks are carried out automatically online and administrative decisions are made, this would require user acceptance in terms of administrative action. As such, it seems essential to provide users with the following information, at the very least:

- the right to know who they are speaking with, whether this may be a civil servant or a virtual assistant (identification principle);
- the right to request assistance from a human in the event of an error or problem when using the service (access to human assistance principle)

Varying levels of human assistance could be requested based on a reference grid, according to the nature of the automated service provided (*see studies carried out by Fing on mediation concerning sociotechnical systems*). This grid would enable administrations with access to automated support or decision-making tools to measure the quality of assistance and mediation they provide to users. It would then enable subsequent assessment of the level of mediation required in order for the service to run smoothly, and enable the following to be defined:

- percentage of the initial budget to be set aside for mediation required for the service;
- the percentage of economies of scale carried out as a result of automatic funding of digital meditation initiatives, both within and outside of administrations;
- percentage of staff (and/or time worked by staff) dedicated to activities concerning mediation and human assistance.

The grid could also serve as a source of inspiration for private actors who wish to develop best practices in terms of mediation for algorithmic systems.

Heading towards an accountability scale for mediation (*Nos Systèmes, Fing*)

Fing suggest creating a mediation/accountability assessment scale to enable the number of accessible mediation modes to be noted and to assess the level of assistance and response provided by the system. For example, 'Level 0: no means of assistance, and potentially even difficult to find information', 'Level 1: assistance can only be accessed outside of the system (via email, freephone number, panic button), 'Level 2: auto-assistant systems', 'Level 3: personalized responses', 'Level 4: coordinated and distributed mediation modes, required to respond, etc.), 'Level 5: remote assistance, even outside of the system'.

Mass training of civil servants in digital mediation

The adoption of digitization, and especially automation, within public services will be accompanied by a growing need for on-hand human dialogue and mediation,

particularly in regard to the most vulnerable members of society. The role of administrative staff to assist the public will become all the more crucial.

Development of artificial intelligence within the civil service will only be advantageous if working conditions for civil servants are improved for the benefit of users. Optimization of administrative procedures must echo the aim to empower civil servants (support with looking up information on exceptions to procedures, or that do not directly relate to their field of expertise, automated data entry and transmission, etc.) and to refocus their responsibilities on providing human assistance to those in need, as well as on the development of better institutional coordination between actors involved in care (administrations, assistants, associations, etc.). The reception, guidance and assistance provided to users requires a coordinated approach by actors on the ground, contrary to public, kiosk or device-based approaches which currently prevail.

As such, it is necessary to train civil servants in public assistance on a massive scale, whilst also strengthening links with existing actors in digital mediation and social policy professionals (in digital co-working spaces, associations, foundations, etc.).

Using EPNs (espace public numérique - digital public spaces) to raise awareness of and report on discriminatory biases within automated access to basic services (housing, employment, healthcare, etc.).

A number of studies have shown that machine automated tasks do not necessarily smooth out subjective biases of human procedures. Public authorities must therefore equip themselves with the skills required to better understand, identify and fight against forms of algorithmic discrimination, particularly when it affects access to basic services such as housing and energy, healthcare, employment and training, and credit. These skills could also be of a technical nature (*see proposal on audit procedures for algorithms in the section on ethics*) or institutional.

Public authorities must develop new channels to communicate with citizens in order to facilitate reporting of experiences on the ground and carry out testing in real conditions. To do this, they must call on support from the digital mediation network and associations for the protection of rights.

In conjunction with anti-discrimination and human rights associations, digital public spaces (EPN) could:

- offer awareness-raising conferences on the risks posed by algorithmic discrimination;
- organize citizen panels in order to test and identify possible biases;
- launch research-action groups to better understand the appearance of certain forms of online exclusion or marginalization.

The quality and representativeness of datasets on population groups are correlated to their social groups

In their publication 'Big Data's Disparate Impact' (2015), Solon Barocas and Andrew Selbst (Princeton University) show that the quality and representativeness of datasets on population groups are correlated to their social groups. They

differentiate between several 'at risk' class labels: the 'uncounted' (individuals with little engagement in the formal economy and therefore in its data generating activities), the 'unaccounted' (individuals with limited access to the internet, or with little fluency in the technology required to be active online), and the 'discounted' (individuals who, as a result of their economic situation, are less interesting as targets of observation).

Supporting AI-Based Social Innovation

At present, AI innovation skills are highly concentrated within a small number of businesses. With the exception of healthcare, social fields receive minority shares in private investment. This current coordination of the AI innovation ecosystem has consequences on the speed in which progress is made in social fields.

In order to redistribute innovation skills, public authorities could launch programmes specifically designed to assist AI innovation in the social sphere and equip social actors with tools that would enable them to benefit from AI-related developments.

Providing an AI skills and resource hub for social policy actors (administration, social policy professionals, associations, etc.)

Public authorities could support the distribution of AI innovation skills to actors within the social field who play a key role in social support, notably associations, mediation actors, social enterprises, etc. To do this, they could:

- facilitate networking between businesses and associations concerning AI projects;
- provide resources (data, computing skills, etc.);
- create a hub of excellence (experts in AI, data science, etc) to enable associations to offer and develop social support prototypes using AI.

Data for Good

Data for Good is a community of over 300 engineers and data scientists who voluntarily put their skills to use to resolve large-scale social issues. An acceleration programme is launched three times a year in order to develop a dozen volunteer-led projects. During this 10-week programme, volunteers are paired with mentors and attend workshops to improve their skills. The following are examples of projects undertaken: predictive analysis for food assistance in partnership with the French Red Cross and mentorship matching in support of equal opportunities in partnership with the association, Frateli.

Documenting the effects of robots in social support

Combined progress in AI and robotics has also led to the emergence of new forms of health and social care: robots, or 'automated assistants'. The development of empathy skills within these machines, or in other words, their ability to express a particular emotion to adapt to their human counterparts at any given time, may

prove beneficial in order to attune to and reassure users. However, this has raised significant issues concerning user perception of the technology, as well as the extent to which it is deemed socially acceptable. Before the technology is considered for use, it is important to:

- support research in social and cognitive sciences on the emotional attachments which may develop between users and machines, and their potential consequences (risk of dependence, exploitation of emotional vulnerability, confusion with human empathy, etc.);
- regulate the development of bodies of data on emotions obtained in real-life contexts and their potential use for commercial or surveillance purposes;
- initiate social debate on the position and role of automated machines used to support dependent or disabled people.

Promoting the development of assistive technologies to facilitate digital accessibility

Whilst accessibility requirements for digital interfaces aimed at disabled people are usually applied as an afterthought, certain AI-based solutions enable them to be integrated during the interface design phase. AI could also provide assistance to interface developers and designers thanks to automated digital accessibility testing tools (compliance with the RGAA [Référentiel Général d'Accessibilité - General Administrations Accessibility Guidelines]) or interface design tools (such as 'Thegrid.io' and 'textocode'). Public authorities could therefore support the development of such tools and/or promote their usage as part of the accessibility policy.

Long-term, ex-ante consideration of accessibility rules could become less significant in view of advancements in assistive technology. Certain AI-based assistive technology can improve the living conditions of disabled people; for example, Facebook has developed an object recognition tool for the visually impaired which can be adapted to user preferences and interests. Google's 'DeepMind' tool uses automated lip-reading technology to enable hearing-impaired people to better understand and reproduce conversations. Salesforce recently launched an algorithm which uses machine-learning to summarize and produce texts and for people suffering from attention disorders. Public authorities could simulate the development of AI applications focused on dependence and disability, and in this sense:

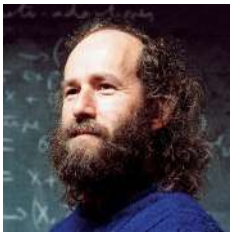
- support investment efforts focused on AI projects within the sphere of dependence and disability, such as the venture capital fund dedicated to artificial intelligence launched by Microsoft in 2017;
- encourage the development of partnerships on AI technology combining businesses, associations, care networks, and research establishments.

The mission

Cédric Villani
Mathematician and Member of the National Assembly

Cédric Villani is a French mathematician and a former student of the École normale supérieure. He received a doctorate in mathematics and he is the winner of the Fields Medal in 2010 and of the Doob price in 2014. He is now professor at the University of Lyon. He has been the director of Institut Henri Poincaré in Paris from 2009 to 2017. He has held various visiting positions at several foreign universities. He is Member of the National Assembly for the Fifth Constituency of the Essonne and he is vice-president of the OPECST (parliamentary office for scientific and technological options assessment). He is member of the Academy of Sciences and has published several books, including *Alive Theorem*, which has been translated in 12 languages.

@VillaniCedric



Marc Schoenauer
Principal Senior Researcher with INRIA

Marc Schoenauer is Principal Senior Researcher with INRIA since 2001. He graduated at École normale supérieure. For 20 years, he has been full time researcher with CNRS (the French National Research Center), working at CMAP (the Applied Maths Laboratory) at École Polytechnique. He then joined INRIA, and later founded the TAO team (*Thème Apprentissage et Optimization*, i.e., Machine Learning and Optimization Theme) at INRIA Saclay in September 2003 together with Michèle Sebag. He has co-authored more than hundred articles and has supervised 35 doctorate dissertations. He has been president of the AFIA (the French Association for Artificial Intelligence) from 2002 to 2004.

@evomarc

Yann Bonnet
General secretary to the French Digital Council.

An engineer by training, Yann Bonnet began his career as a consultant. He joined the French Digital Council in 2013 as General Rapporteur, before becoming Secretary General in 2015.
He was in charge of steering the national consultation on digital transformations, which was launched by the Prime Minister in 2014. This initiative eventually led to the Law for a Digital Republic. Yann Bonnet was also in charge of multiple reports, including taxation in the digital age, the digital dimension of the TTIP negotiations and online platforms fairness.

@yann_bonnet

Charly Berthet
Head of legal and institutional matters at the French Digital Council

Charly Berthet is a French lawyer working at the French Digital Council as head of legal and institutional affairs. He has worked specifically on regulation matters, on data protection and civil liberties. He has been a consultant for the Ministry of Foreign Affairs where he helped elaborate the digital international strategy. He graduated at University Paris II and University Paris Dauphine.

@charlyberthet

Anne-Charlotte Cornut
Rapporteur of the French Digital Council

Anne-Charlotte Cornut graduated from Sciences Po and HEC and is rapporteur of the French Digital Council since april 2016. She worked on the digital transformation of the SMEs and of the higher education and research. She formerly was adviser to the CEO of 1000mercis/numberly, a data marketing company.

François Levin
Head of economic and social affairs at the French Digital Council

François Levin has graduated in philosophy from École normale supérieure of Lyon and in public administration at University Paris I. He joined the French Digital Council in 2015 and is now head of economic and social affairs. He has specifically worked on the digital transformation of work and formation as well as of culture and copyright law.

Bertrand Rondepierre
Engineer in the Corps de l'armement working for the Direction Générale de l'Armement (French defense procurement agency)

Bertrand Rondepierre graduated from École Polytechnique, holds an engineering degree from Telecom ParisTech and is an alumnus of the Master's degree Mathematics, Vision, Learning at ENS Paris-Saclay. He works as a system architect for the DGA, where he runs projects in digital and artificial intelligence fields.

@BertrandRdp

Stella Biabiany-Rosier
Executive Assistant of the French Digital Council

Stella Biabiany-Rosier has spent her career as a Assistant Manager in consulting and law firms, then in ministerial offices. Since July 2017, she has been assisting the General Secretary of the French Digital Council.

For a meaningful artificial intelligence

Towards a French and European strategy

aiforhumanity.fr

March 2018