SIAPARTNERS

RESEARCH

Artificial Intelligence trends for 2022 – 2023.

January 2022


@SiaPartners

Summary

Words from CEO Matthieu Courtecuisse.

COVID has pushed the progression of AI across all sectors. The health crisis and the automation trend significantly challenged business operating models. We are also witnessing an acceleration of investments in these areas from our clients.

From individuals to states and throughout companies, artificial intelligence is now the primary fuel for technological disruptions at all scales.

Far from the fantasies accompanying its first steps, AI contributes to progress throughout all sectors. Without AI and digitalization, lock-down would have been even more damaging to our lives and economies. For AI to continue its fast development, we must encourage education in science and mathematics, a regulatory framework that promotes competition and innovation, and finally aim for carbon neutrality in digital technology within 20 years. This political, social and economic commitment will be decisive for the competitiveness of companies and states. This is the sine qua non condition to catalyze the development of AI and the proliferation of its use cases.

For the past few years, the trend in all sectors has been to integrate toolkits powered by AI, RPA, or blockchain to develop new use cases: forecasting needs, optimizing to increase productivity, improving data quality, helping with decision making, improving customer knowledge, etc. For us, this takes the form of Augmented Consulting: business experts who intervene using technical tools to accelerate their intervention and bring more value to their customers.

The main difficulty for our clients today in seizing the opportunities offered by AI is to identify the right use cases and industrialize them. AI projects are often confronted with the reality of business issues such as large volumes of data, IT legacy, existing business processes, etc. If we exclude the tech giants, which are natively involved in these subjects and which are designed to integrate AI into all of their activities, the big challenge for companies in 2022-2023 will be the ability to industrialize the production of AI use cases at a large scale, by providing a favorable human and technological ecosystem.

This publication by Sia Partners aims to provide an overview of the major dynamics around AI and what the year 2022-2023 has in store for us in this area. This exercise was based on our knowledge and expertise in Data & IA, on the business expertise of our sectoral directors. They support our clients daily in their Data & IA transformation and in numerous consultations with experts in Business & Tech.

1. The democratization of AI.

In 2021, the technology sector had the largest number of IPOs (initial public offerings) with 611 - 26% of IPOs - for a total amount of $147.5 billion. GAFAMs, which are continuing to conquer the world of the future by increasing their hold on data, have seen their stock market valuation double between January 2019 and July 2020. In companies, data teams, data factories and CDOs are on the rise.

An accelerated development through several phases

The multiplicity of data collection channels and the technical breakthroughs in the ability to manipulate large volumes of data have been the main drivers of the Big Data boom, leading to a 20-fold increase in the volume of digital data created worldwide over the past decade. This significant volume has been combined with lower storage and processing costs, making Big Data an attractive and accessible market.
The last few years have also been marked by the emergence of more powerful algorithms. For example, image processing by neural networks has reached such a maturity that it now challenges the most advanced human diagnosis of lung cancer. NLP techniques, on the other hand, enable all sorts of treatments and analyses on all types of textual data.

At the dawn of the big bang of data, the need for data tools and the numerous use cases have given free rein to the imaginations of statisticians to design forecasting or modeling algorithms that, within a limited scope, have demonstrated their value within companies. Two crucial challenges remain: to move forward to the industrialization stage and to ensure that, in addition to statisticians, business units can also adopt and design data driven use cases.

Challenging journey of scaling up POCs

Moving from the prototyping phase to the industrialization phase, remains the reason for many discarded prototypes.

The first prerequisite for industrialization is the availability of an infrastructure that allows the implementation of centralized data models and the creation of DataLakes. These facilitate the encapsulation and harmonization of POCs into a single robust platform. The second prerequisite is the compatibility of AI platforms with existing IS systems to enable the deployment of AI-powered tools in real-world scenarios. Last but not least, one major prerequisite is the translation of algorithmic models into business interfaces adapted to the users' processes. The industrialization of AI models calls for a revision of the algorithm (data volume, processes capacity etc.) in order to robustly expand the scope of the POC to a larger scope of application. In addition, as the complexity of the model increases, we are witnessing the emergence of low-code solutions for non-specialists, enabling the manipulation of databases in a matter of a few clicks.

Data, at the essence of corporate strategies

Last year, 69% of French data-driven companies maintained or even increased their investment in data. This position was reinforced by the pandemic, with more than 80% of business leaders claiming to have gained decisive advantages thanks to data during this period. Furthermore, the startup community is increasingly focusing on data. In 2021, BPI France listed 502 French startups dedicated to AI in all areas (industry, transportation, human resources, etc.).

In addition to start-ups, most companies, large groups, SMEs and ETIs, are embracing data by creating data teams dedicated to data recovery and management. This talent pooling often goes hand in hand with the designation of a Chief Data Officer (CDO). This position, which only existed in 12% of companies in 2012, is now present in 65% of them.

Despite the swelling ranks, the scope and the hierarchical structure of CDOs still varies from one company to another. A sign that a winning combination has not yet been found ?
Furthermore, according to the same survey only half of CDOs (49.5%) have led supervisory authority over data in their organizations, and only a third rate the CDO role as «successful and established".

Data-driven organization : a challenge of acculturation

While it is still difficult for many to measure the business value of chief data officers, their role as coordinators and

(1) https://www.decideo.fr/80-des-entreprises-europeennes-data-driven-affirment-beneficier-d-un-avantage-decisif-alors-meme-qu-elles-continuent-a_a12372.html

catalysts is essential, particularly in terms of advancing the adoption of new technologies and facilitating change management. Gartner mentions the difficulties of effectively implementing intelligent technologies in everyday life, referring to a plateau of AI productivity (Gartner AI hype cycle). Industrialization and data acculturation will be two major aspects of the AI maturation process in the upcoming years.

2. Data-driven organization: a challenge of acculturation.

The application of AI capabilities follows different dynamics depending on the sector. These dynamics can be catalyzed by economic circumstances such as the sanitary crisis for the health sector, but more fundamentally by the quality of the technological ecosystem that drives each sector. While the advent of AI has made it possible to boost business models of various sectors such as energy or insurance, it has been transforming other sectors, such as health, in a more disruptive way. These transformations are already giving rise to major use cases in these sectors. The sectors presented in this article intend to provide a non-exhaustive illustration of sectoral trends in AI by 2022-2023.

Energy & Utilities: a matter of data capture

Since 2010, the massive deployment of smart metering devices, on the electricity, gas and water networks, has made it possible to address new use cases where AI allows large-scale achievements: detection of leaks in water, detection of electricity consumer fraud, predictive maintenance or even enhanced understanding of consumer behavior.

«Smart meters electricity consumption data is a major source of information for AI algorithms in the Energy & Utilities sector»

Based on machine learning, managers of distribution grids can accurately predict flows passing through all the links in their networks and thus anticipate corrective actions to be implemented: intelligent network monitoring via constraint management, activation of necessary flexibilities, better integration of renewable energies, detection of anomalies and finally, appropriate maintenance of the most sensitive equipment.

On the supplier side, the use of this data allows a better knowledge of consumers: portfolio view optimization, dynamic profiling of segments, appropriate definition of pricing policies or marketing campaigns improvement.

“More than 203 million smart meters will be installed worldwide in 2024, or as many points of data collection”.

Beyond the proven short-term return on investment of these use cases (in particular in fraud recovery for example), these algorithms also make it possible to plan for the medium term and are key tools used to plan the design of tomorrow’s networks in the context of the energy and ecological transition!

Insurance, a sector with a head start in AI

The insurance sector, which is intrinsically quantitative in nature, has benefited from a momentum already underway regarding the integration of AI use cases into its business model. AI has indeed greatly contributed to the operational excellence of insurance, as we observe an acceleration of processing times on many components of its value chain.

AI has boosted the advent of self-service (especially in insurance underwriting and simple claims) via the automation of management actions, in order to refocus insurers’ efforts on actions with high added value. The management of complex or bodily claims becomes more efficient, leading to a better customer experience. Similarly, automating the processing of incoming documents from policyholders or third parties brings significant competitive advantage.

AI also contributes to a better understanding of clients and therefore to a deepening of their segmentation thanks to clustering methods. This applies to several areas, including the determination of tailor-made pricing policies according to the insured’s profile, the adoption of preventive anti-churn measures thanks to the detection of clients likely to leave the portfolio or finally the detection of insurance fraud, which remains a major challenge for the sector, thanks to anomaly detection mechanisms.

While claims management remains a major element of the economic profitability of an insurer, the progressive application of AI use cases affects the entire value chain. This is especially true regarding the pillars of customer experience, namely servicing and customer relations, which are key to improving the Net Promoter Score (NPS) or voice analytics which allows the «augmented» manager to better adapt to the intention and psychological disposition of the caller.

(2) https://www.openaccessgovernment.org/how-automated-meter-readers-are-transforming-the-water-utilities-industry/117799/
(3) https://iot-analytics.com/smart-meter-market-2019-global-penetration-reached-14-percent/

Health: the eldorado of AI innovation

Health is a popular sector for AI applications, and the sanitary crisis has further strengthened this dynamic. Today, one in five AI start-ups choose to invest in e-health.

Whether in the field of radiology, dermatology or oncology, computer vision techniques provide specialists with tools to improve the reliability of their diagnoses. Some AI algorithms enable early prediction of kidney cancer metastases, thus increasing the chances of patient survival between 20% and 80%. AI use cases are proliferating in medical biology, ophthalmology or neurology, while in the field of epidemiology, AI makes it possible to predict the outbreak, amplitude and spread patterns of pandemics.

Nevertheless, the democratization of AI faces major obstacles. The first being the absence of collaborative databases that would consolidate medical data at a large scale. This is a key prerequisite for the training of algorithms but also for the proper processing of patients' medical histories.

The second obstacle or safeguard depending on one's point of view, is that in many countries, the management of medical data faces issues such as the administrative complexity, the confidentiality of medical records or the reluctance of patients to share their medical data.

In the near future, the emergence of collaborative health platforms will universalize medical profiles of consenting patients, and will make them available to their doctors.

Doctors will benefit from AI applications trained on a large number of medical cases and available on a collaborative exchange platform, similar to the business models of certain major online marketplaces.

(4) Net Promoter Score: Indicateur mesurant l'intention de recommandation des clients
(5) https://www.alliancy.fr/france-digitale-cartographie-startups-ia-france

3. AI and R&D intertwined.

Data and statistics are the two cornerstones of scientific research. Today AI capabilities are disrupting research methods, whereas research activities originated the theoretical foundations of AI since Alan Turing's work on the idea of machine intelligence in the 1940s. Seventy years later, AI revolutionizes research in many fields by providing researchers with world-class analytical capabilities while reducing the cost, time and complexity of research.

A theoretical corpus that keeps on developing

Nowadays, the immediate availability of very large volumes of data associated with new algorithmic techniques and computer power has led to major technological breakthroughs: Convolutional Neural Networks, Reinforcement Learning, Transfer Learning, Generative Adversarial Networks (GANs), etc. These breakthroughs act as catalysts for scientific discoveries.
On the other hand, the work on the explicability and interpretability of algorithms facilitates and accelerates the popularization of their use to research communities for which traceability of results is a major concern.

This enthusiasm in the scientific community is reflected in the exponential growth in the number of publications related to artificial intelligence on the arXiv : multiplied by 6 between 2015 and 2020. Similarly, open archives, conferences, and publications dealing with AI are multiplying, with an increase of nearly 14% in the number of publications between 2019 and 2020. This craze has been boosted by international competitiveness. In 2020, China surpassed the United States for the first time in the number of AI-related citations. France is also a fertile ground, with AI-research incentive mechanisms, such as the Research Tax Credit. The country now has nearly 81 R&D laboratories and research centers and more than 13,000 researchers working on AI topics.

AI challenges conventional research

AI allows the consideration of less conventional perspectives: probabilistic programming allows for the generation of hypothetical data that can be compared to experimental data thus broadening the spectrum of data analyzed. In the field of molecular modeling, AI disrupted conventional methods of modeling proteins and molecules. This is what enabled AlphaFold to manage and process significant volumes of data, with over 100 million proteins, to understand protein folding mechanisms on a large scale, which researchers had previously been unable to achieve. The autonomy of learning algorithms is remarkable, enabling the shift from deterministic modeling governed by molecular dynamics equations to ML capable of deducing the behavior of proteins from past behaviors. Therefore, all experimental and analytical chains are impacted by AI. The scaling achieved by AlphaFold would have required decades of research to obtain the same results without AI. Beyond the economic advantages, the gains in time and precision described above have benefits on a medical level, in particular in the field of drug design for which the understanding and modeling of 3D protein structure is crucial.

(6) arXiv : archive ouverte de prépublications électroniques d'articles scientifiques
(7) https://www.intelligence-artificielle.gouv.fr/fr/strategie-nationale/la-strategie-nationale-pour-l-ia
(8) https://www.journaldunet.com/solutions/reseau-social-d-entreprise/1192757-carte-de-france-des-laboratoires-d-intelligence-artificielle/

4. The academic rise of AI.

With a progressively positive public perception, AI has become an important source of economic business and is increasingly implemented within companies. Many people have developed an interest in the applications of AI and are eager to acquire the technical skills to tackle them.

The effervescence of online education

«With more than three million students enrolled since the course launched in October 2011, the most followed MOOC in the history of online education is Machine Learning,» from Stanford University on the platform Coursera. The success of this online course, taught by Andrew Ng, is not trivial. It demonstrates the strong appeal of AI and the new forms of learning that promote and democratize it.

Communities of data scientists as catalysts of this democratization

A specificity of digital revolutions, particularly that of AI, is the considerable influence of online communities contributing to their growth. As a data scientist, it is essential to remain abreast of technological evolutions and the best development practices. There are many online communities, such as Stack Overflow, one of the largest developer communities with 14 million users, where AI enthusiasts can share their code, knowledge, ideas and challenges. Github is another perfect example. As a leading code hosting and versioning platform (with 40 million users and 190 million projects), it has become a de facto clearinghouse for all open source projects, including libraries used in the data science community. «More than 73 percent of data scientists use open source libraries''. Popular AI-specific libraries, such as TensorFlow with 85,500 clones of the repository or PyTorch, play an important role in supporting the developer community in creating AI solutions.

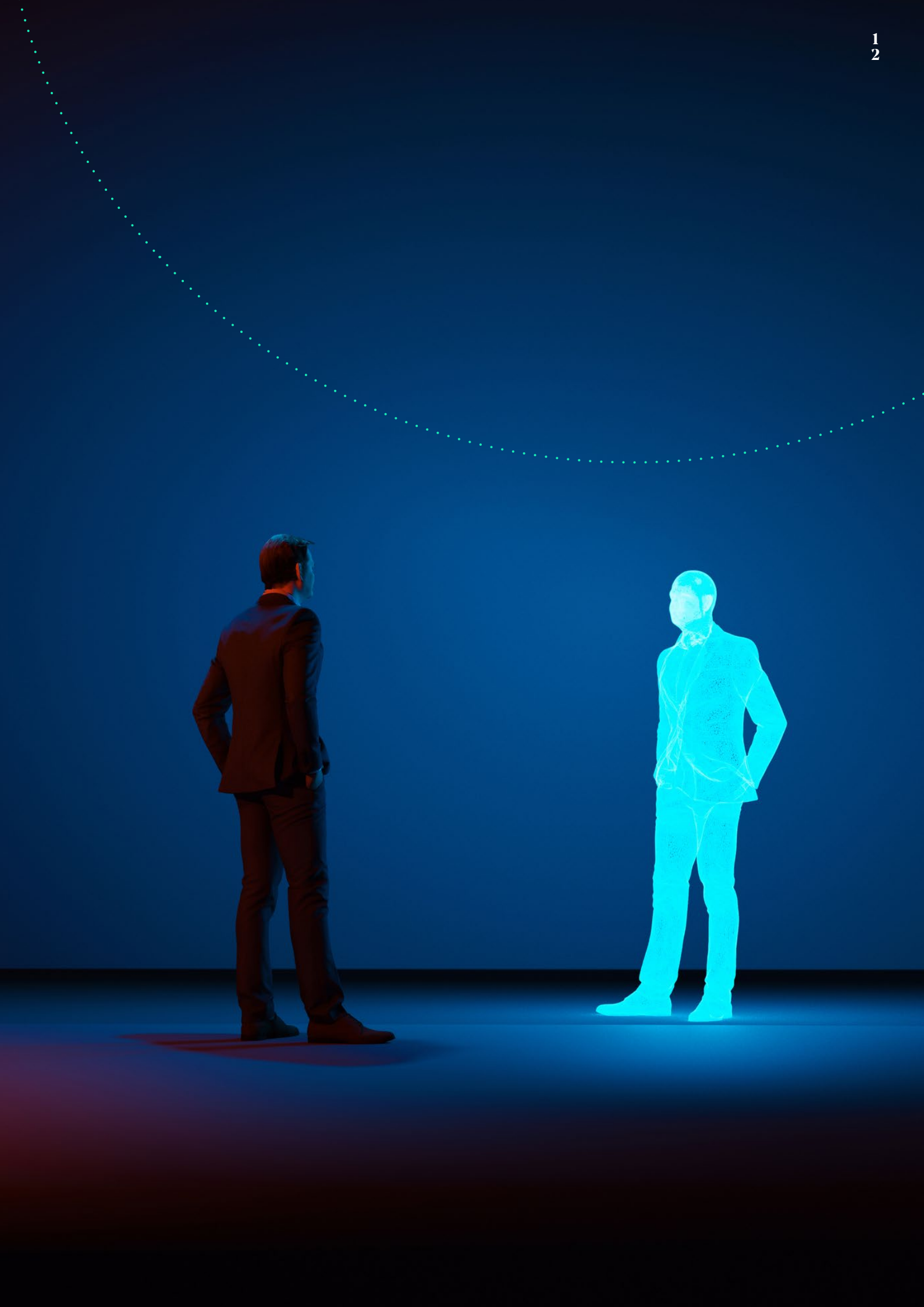
The massive integration of AI in academic curriculum

The proliferation of data use cases and the emergence of new statistical models have been supported by the development of numerous relevant academic courses. These allow specialization in specific fields such as word processing or reinforcement learning. «The number of AI courses has increased by 103% at the undergraduate level and 47% at the master's level over the past 5 years». France now has nearly 18 masters' degrees specialized in AI and more than 1,000 theses on these subjects. These figures are expected to increase: Emmanuel Macron has pledged to «triple the number of people trained in AI within three years» in France.

Companies, key players in training and academic projects

Increasingly, schools are collaborating with companies to create content for their academic programs. Some artificial intelligence courses are designed and taught entirely by corporate speakers. These partnerships between schools and companies are accelerating the professionalization of students and facilitating the recruitment of talent trained in AI.

(9) https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report-_Chapter-4.pdf



5. An ever evolving business-skill reference system.

As roles in the data value chain multiply, the data scientist's job is to maintain focus on their core function: modeling. These evolutions lead to many conceptual changes in the recruitment processes of companies..

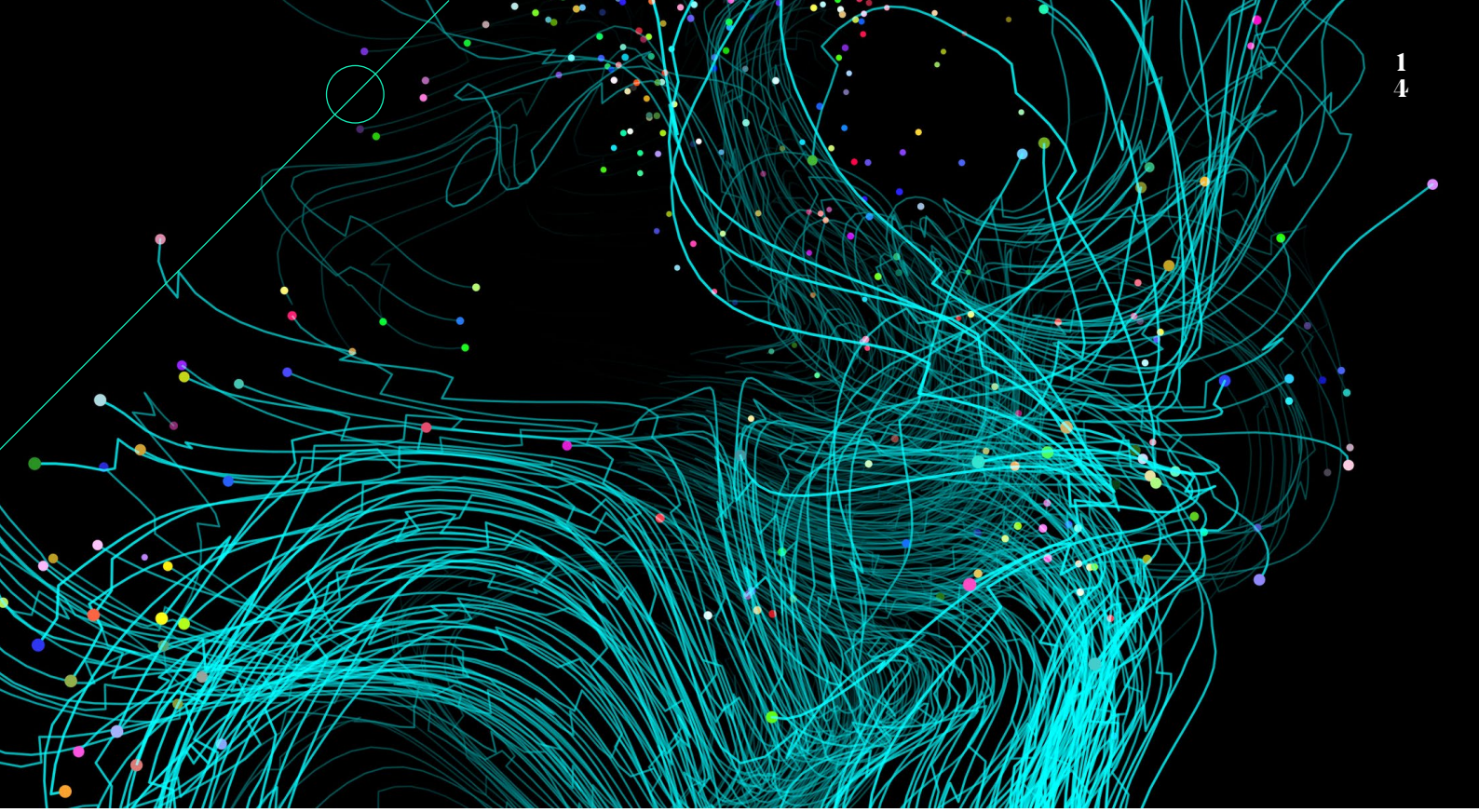
The end of "Data Scientists" ?

In 2017, the Harvard Business review classified data scientist as the «sexiest job in the world». The definition of data scientist given at that time is now being challenged! Previously a Jack of all trades, the job of data scientist is now re-focusing on the heart of the trade, which is modeling. The roles of data specialists are multiplying in the value chain, upstream and downstream of modelisation tasks. Beyond the development of POCs, the complete production of an artificial intelligence project requires specific technical skills such as mass data manipulation, dynamic visualization on a web interface or the use of cloud technologies. An AI solution «product» team is now made up of a wide variety of profiles: data scientist, analyst or engineer, machine learning engineer, cloud specialist, cloud archi- tect, NLP scientists, or even frontend developer.

This evolution is a result of specific scientific and technical requirements, but also of managerial ones. Algorithms are essential, but to succeed in an artificial intelligence project, business stakeholders must be involved throughout the design process, an integration that necessitates astute project management. This implies a strong capacity for collaboration and communication between business and data experts, as well as pedagogy on both sides. To be successful, data science teams need to be supervised and supported by roles such as chief data officer, data quality officer, product owner, or data protection officer. The exact roles and responsibilities of these professions are still being debated, but there is a consensus that they are necessary.

Adaptation of recruitment processes

In the context of fierce competition between companies for AI skills, companies are shifting their recruitment processes towards unconventional methods, in order to increase attractiveness to highly sought-after profiles. In addition to traditional recruitment channels, monitoring both specialized coder forums and data science competitions is becoming a competitive advantage for recruiting the best professionals. A good ranking in such competitions, or significant activity on forums, are now key performance indicators in the eyes of recruiters. Decisive weight is now given to these elements in the recruitment process..



6. Continuous innovation of technologies.

The field of AI has seen one of the fastest evolutions in the market. Innovations in terms of algorithms and modeling as well as in terms of technologies are accelerating. Nevertheless, while the last few years have seen the democratization of deep learning and cloud computing, the next few years could bring algorithmic and technological disruptions, with the emergence of new paradigms (edge AI, low-code, etc.)

The new generation of algorithms: continuity rather than disruption

The new generations of modeling algorithms are not disruptive innovations, but rather improvements of existing models, the most powerful of which are based on deep learning. The latest iteration of deep learning models, transformers, offer great performances and are becoming increasingly prevalent in all areas of AI, from computer vision (image processing) to NLP (natural language processing) and time series.

However, deep learning algorithms generally have anywhere from thousands to millions more parameters than traditional machine learning algorithms, which makes them more demanding in terms of the data and machine resources required to train them. This growing need for computational power and storage capacity puts cloud providers in a prime position to offer cost-effective solutions to businesses, thanks to economies of scale on massive infrastructures.

Outsourcing computational processing and data storage to cloud providers has many advantages for companies , even in cases of sensitive data, for which cloud providers have specific offerings. For example, a GPU (Graphic Processing Unit) must be running at least 70% of the time to be profitable, which is rarely the case for a single company. Furthermore, a complete and expensive technological ecosystem is required to take full advantage of a GPU, which is very costly. Transitioning to cloud resources, where companies only pay-what-they-use and don't have to manage the hardware, is therefore a strategic move for many companies.

Cloud providers are therefore benefiting from a favorable dynamic in a rapidly growing market. Nevertheless, competition remains strong, firstly from services dedicated to artificial

intelligence, but also from data lakes and data warehouses. The latter are beginning to merge to form a single system, accompanied by computation capacities and business intelligence tools. The field of cloud services has recently seen a number of newcomers snatch significant market shares. One of these new players, Snowflake, made a huge IPO in 2021, worth USD 12.4 billion, illustrating the scale of the phenomenon. This represented not only the biggest IPO of the year on the US stock exchange, but also the biggest IPO in history for a software company. Today in France, 70% of the data hosting market is held by Amazon, Microsoft and Google. In response, the French government is seeking to give preference to French and European operators and is implementing a sovereign cloud strategy with the «trusted cloud» label. This new situation in France has led to the emergence of new partnerships, such as the one between Thales and Google.

With a trend of service optimization (less networking, more simplicity), the next few years could see a convergence between cloud providers and SaaS cloud platforms (Snowflake, Databricks, etc.), leading to the emergence of new offerings from cloud providers to fill this gap in their services.

The new paradigms of AI: Edge computing and hybrid models

While deep learning is now one of the most widespread trends in artificial intelligence, other approaches are returning to the forefront. For example, this is the case for probabilistic approaches and expert systems, which were very widespread until the mid-2000s. These different trends are becoming more and more intertwined and operate less and less in silos, and it is probably from a hybrid approach that the next breakthrough innovation will emerge. Such algorithms will undoubtedly be much less resource- and data-intensive, with equivalent or better performance. Replacing big data with smart data, and memory with the ability to abstract concepts, would greatly reduce storage and computing capacity requirements.

Other emerging technologies such as edge AI or federated learning also reduce storage and communication requirements while offering better data security. What these two practices have in common is a profound paradigm shift in the way we use data and train AI models. Traditionally, data from many sources such as sensors or cameras is centralized in a data center, before being processed and used to train a single model. In the case of edge AI or federated learning, the data does not leave the source, but rather the algorithms and models are placed as close as possible to the data source. This paradigm shift offers many advantages, particularly in terms of network latency and data protection, since data never travels through a network.

Low-code: evolution or revolution ?

One of the main obstacles to widespread AI adoption is the shortage of data scientists capable of creating the necessary tools and algorithms. No-code and low-code solutions aim to overcome this problem by providing simple interfaces that can be used, in theory, to build increasingly complex AI systems. Just as no-code web design and user interface tools allow users to create web pages and other interactive systems by simply dragging and dropping graphical elements, low-code AI systems will allow intelligent programs to be created by plugging in different pre-built modules and feeding them with specific data. All this will play a key role in the on-going «democratization» of AI and data technologies.

This paradigm shift in development tools is reflected in the growing market share of no-code and low-code in the highly competitive SaaS (Software as a Service) market. These no-code and low-code tools are emerging as very promising solutions. In 2020, they represented a market share of 15 billion dollars and is expected to reach 112 billion dollars in 2026. This new type of platform flips the development process of an application or service, allowing business experts to be the architects of their own AI solutions. The advantages of this type of tool are multiple, including more permanent added value to the business with faster design times prototyping and deployment. But their main advantage lies in facilitating change management around AI projects, a major challenge for companies in the upcoming years.

In practice, the shift towards these new types of platforms remains a little bit complicated and the support of data experts is still necessary for the time being. It is indeed essential that these new applications can be integrated into the existing ecosystem, have access to proper data and be maintained in case of malfunction. The development of this type of application has the merit of raising and clarifying data governance issues, previously implicit and underlying in AI products. While no-code offers limited functionality and little flexibility in the design of algorithms, low-code facilitates the creation of powerful and more responsive applications, while leaving a certain amount of flexibility to developers.

Beyond algorithmic challenges

If algorithmic performance and technologies used are key success factors in the AI field, the elements upstream and downstream of algorithmic modeling, i.e. data collection and processing on the one hand, and the industrialization of use cases on the other, are just as important. However, most POCS are never deployed on a large scale, due to a lack of production capacity of companies, which requires the synergistic efforts of many departments (project management, adapted technological infrastructure, integration with existing IT systems, etc.).

A second challenge lies in the quality of the data. In many companies today, data is still too disparate and insufficiently standardized to allow industrialization of AI models. In the short and medium term, the greatest value of AI likely lies in the standardization and automated enrichment of data.

Resistance to change is an obstacle to the development of many artificial intelligence projects. For example, the automatic modification of data by an algorithm with results that are difficult to explain is a disturbing idea, even if the error rate is lower than that of existing processes. The acculturation of the data within companies plays a decisive role in the responsibility of CDOs (chief data officers).

(10) D'après les analyses de Forrester Research

7. Public policies and competitiveness around AI.

As AI becomes a lever for transforming society, many governments have adopted proactive policies to support an AI-ecosystem. More than 30 countries have already created national AI strategies to improve their prospects. Nations that take the lead in the development and use of AI will shape the future of this technology and significantly improve their economic competitiveness.

The race for AI: who are the champions today?

So far, according to the ranking compiled by the Center for Data Innovation based on 30 metrics and 6 categories (talent, research, development, hardware, adoption and data), the United States has established itself as a pioneer of AI, but China continues to close the gap in certain major areas (data quality, computer power, research, etc.). Without significant policy changes in the EU (still recovering from the UK's departure), such as the EU changing its regulatory system to be more innovation-friendly, it is likely that the EU will remain behind the US. Similarly, without the U.S. developing and funding a more proactive national AI strategy, China will dethrone the U.S.

The proliferation of national AI programs

Canada was the first country to launch a national Artificial Intelligence strategy in 2017 before being followed by more than thirty countries. The time horizons of the different national AI strategies vary. France sets a 4-year horizon with the National AI Research Program, which precisely lays out the actions to be taken in this period. Other countries are planning their strategies for the next decade, such as the United Arab Emirates with the AI 2031 strategy, which is presented according to certain rankings as being one of the most ambitious national AI programs in the world in this area.

The strategies for implementing AI policies are critical to their success and reflect the ambitions of nations. The emerging trend is to establish dedicated administrations or government institutions specially designed to manage and implement AI strategies and programs, such as the Ministry of Artificial Intelligence in the United Arab Emirates. In other cases, AI mandates are assigned to existing interdepartmental bodies or committees or to public-private partnerships.

Zoom in on the french strategy

After the announcement of 1.5 billion euros invested in AI in 2017 (following the submission of the report by MP Cédric Villani), at the end of 2021, the French government announced a new envelope of 2 billion euros, of which a large part, nearly 50%, will be dedicated to skill development through future institutions of excellence and AI training programs. The emphasis on training aims to strengthen France's competitiveness. As part of the national AIForHumanity policy, interdisciplinary AI institutes (3IA) have emerged in different cities such as Toulouse, Nice and Grenoble. While the R&D centers of Facebook and Google readily recognize the skill level of French talents, French players are concerned with the difficulties in recruiting. Other priorities of this new investment plan: the government wants its AI industry to be positioned in emerging markets, such as embedded AI, edge, trusted technologies or even frugal AI.

To increase the competitiveness of companies, it is not only a matter of transforming research work into «economic opportunities» via the emergence of AI champions, but also the dissemination of concrete uses of these technologies within companies. The Île-de-France region implemented the Pack AI system in 2021, with the aim of integrating data into the ecosystems of companies where the complexities of implementation and appropriation are significant. The "direct return on investment" of AI public policies is difficult to evaluate. However, the government has mapped out, in its 2021 inventory analysis of the AI strategic plan first phase, 81 AI laboratories in 2021 – the highest number among European countries -, 502 start-ups specializing in AI – an increase of 11% compared to 2020 -, 13,459 people working in AI start-ups - for 70,000 jobs indirectly generated. With a hint of patriotism, we observe the rise of 6 French unicorns associated with AI, such as ContentSquare, Shift Technology and Mirakl, although the capital raised is mostly from foreign private investors such as Soft Bank or Advent International.

Finally, in order to further strengthen

(11) https://www.tortoisemedia.com/intelligence/global-ai

the data culture at a regional level, several Hackathons are being organized on a regular basis. The organization of such events show that participants are part of a knowledge management and collaborative work approach. France ranks 6th in terms of organized hackathons with 195 per year in 2019.

(12) https://www.economie.gouv.fr/la-strategie-nationale-pour-lintelligence-artificielle
(13) https://www.maddyness.com/2019/01/15/le-succes-des-hackathons-ne-flechit-pas/

8. Ethical and societal issues : concerns, controversies and opportunities.

AI raises questions. In fact, it relies on very power-hungry structures and contributes significantly to digital pollution. The algorithms race raises concerns on the meaning of human control in light of the «black box» effect of certain algorithms. On a more positive note, AI has the potential to accelerate global efforts to protect the environment and conserve resources by detecting reductions in energy emissions and promoting the development of specific use cases. It could be a key asset in the coming years to address the most pressing environmental and societal challenges.

AI, an asset for ecological issues

The pressure on companies to respond to climate change and reduce their carbon footprint is increasing every year. AI is proving to be a real asset as it offers the ability for companies to better leverage their data to obtain key indicators on different aspects of their carbon footprints. Other AI use cases with high added value for civil protection and the environment are bound to proliferate in the years to come. For example, Computer Vision techniques can improve the monitoring of deforestation and offer the ability for civil protection services to detect forest fires, which are very harmful to the environment. Though the fires of 2021 were among the largest ever observed, caused by increasingly extreme weather conditions, with recurring droughts, AI applications demonstrate strong operational potential in the fight against global warming. For example, drones equipped with computer vision algorithms could locate «hotspots» of fires, allowing firefighters to more effectively target their efforts.

In addition to the numerous AI use cases in the environmental sector, it is clear that the computing power required for AI algorithms has significantly increased in recent years. For instance, DeepMind's AlphaZero Go game has doubled computing power roughly every 3.4 months between 2012 to 2018 –increasing 300,000 times– according to an OpenAI study. Let's recall that, according to Moore's empirical Law, computing power doubles every two years. As a result, machine learning systems are becoming more resource-intensive and generating a significant amount of carbon emissions. A 2019 study found that the training process of an NLP model (transformer with neural architecture) can emit more than 285,000 kg of CO2, which is nearly five times the lifetime emissions of an average car, or the equivalent of about 300 round-trip flights between New York and San Francisco.

Planning and assessing the environmental impact before implementing AI algorithms can help reduce the environmental impact. This involves first choosing well-scaled algorithms and optimizing computation costs when beginning the design process. The choice of hardware on which the algorithm is trained, as well as the choice to use «green» datacenters powered by renewable energy, are other levers that enable the implementation of AI with a low carbon footprint. In the long term, the use of AI will certainly contribute to the fight against climate change, provided that companies developing algorithms account for environmental impact and avoid the oversizing of infrastructure and computational power required for the algorithms.

Ethics and AI, an urgent societal issue

Ethical issues surrounding AI are gaining more and more public attention, with research publications on ethics in this area more than doubling between 2015 and 2020. The drive to create increasingly efficient models has led many companies to prioritize operational gains in recent years, at the expense of interpretability and explainability. Moreover, the use cases related to AI are multiple and diversified but the data used to train models is not always as diversified; many groups are underrepresented or absent from landmark datasets. The proliferation of AI solutions raises many concerns

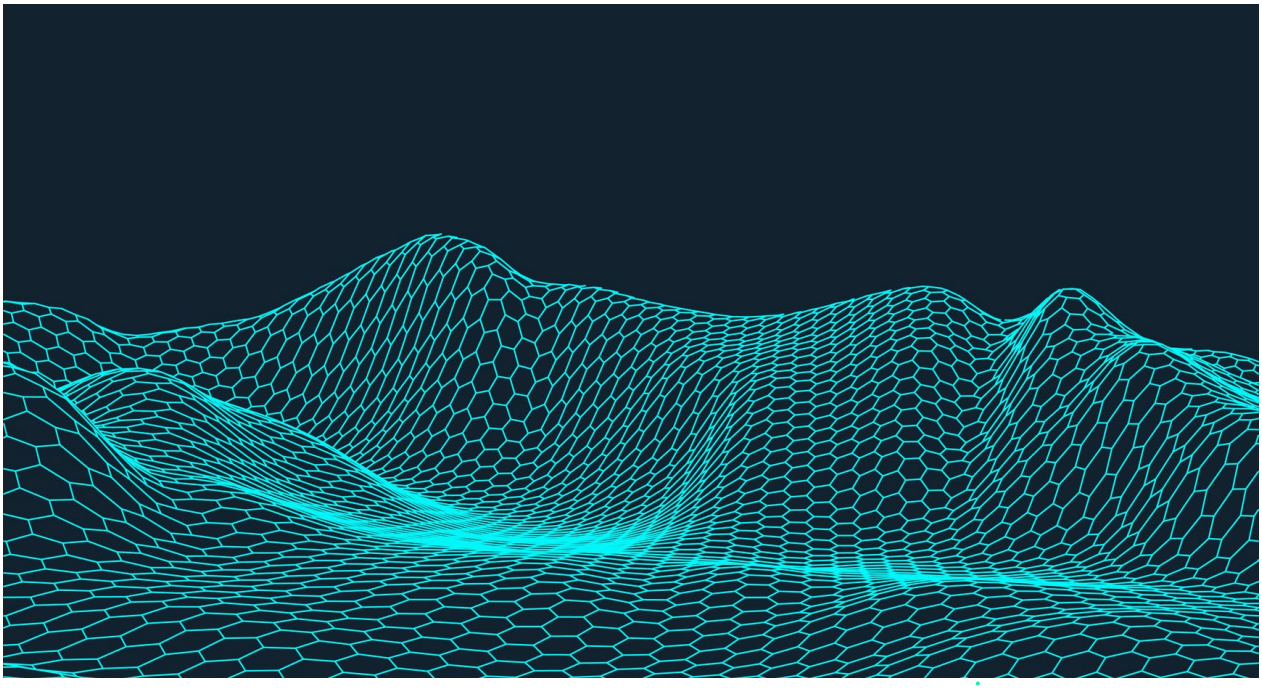
(14) https://arxiv.org/pdf/1906.02243.pdf

about their potential consequences such as privacy intrusion, discrimination, and especially the opacity and racism inherent in various AI decision-making processes. Automated decision-making tools often reproduce human biases based on the race, gender or social class of an individual.

As our societies become more and more dependent on decisions made by algorithms, companies must become more transparent about this decision-making process and find solutions to counter the black box effect of models and data collection processes. This is the case, for example, with neural networks, which are proving to be very effective for voice and image recognition, but whose functioning and results remain difficult to explain. In his report, Cédric Villani recommends in-depth studies to make these algorithms interpretable for regulatory purposes as well as for ethical and performance considerations. Finally, AI can theoretically reduce discriminatory biases in decision-making processes within companies but only if significant changes to the status quo are made. Examples of such decision-making processes include the granting of loans and financing in the banking sector, and hiring in the recruitment sector. Artificial intelligence techniques have the potential to reduce these biases by building adversarial models to highlight disparities. It has never been more important to address existing ethical challenges and develop responsible and equitable AI innovations before deployment.

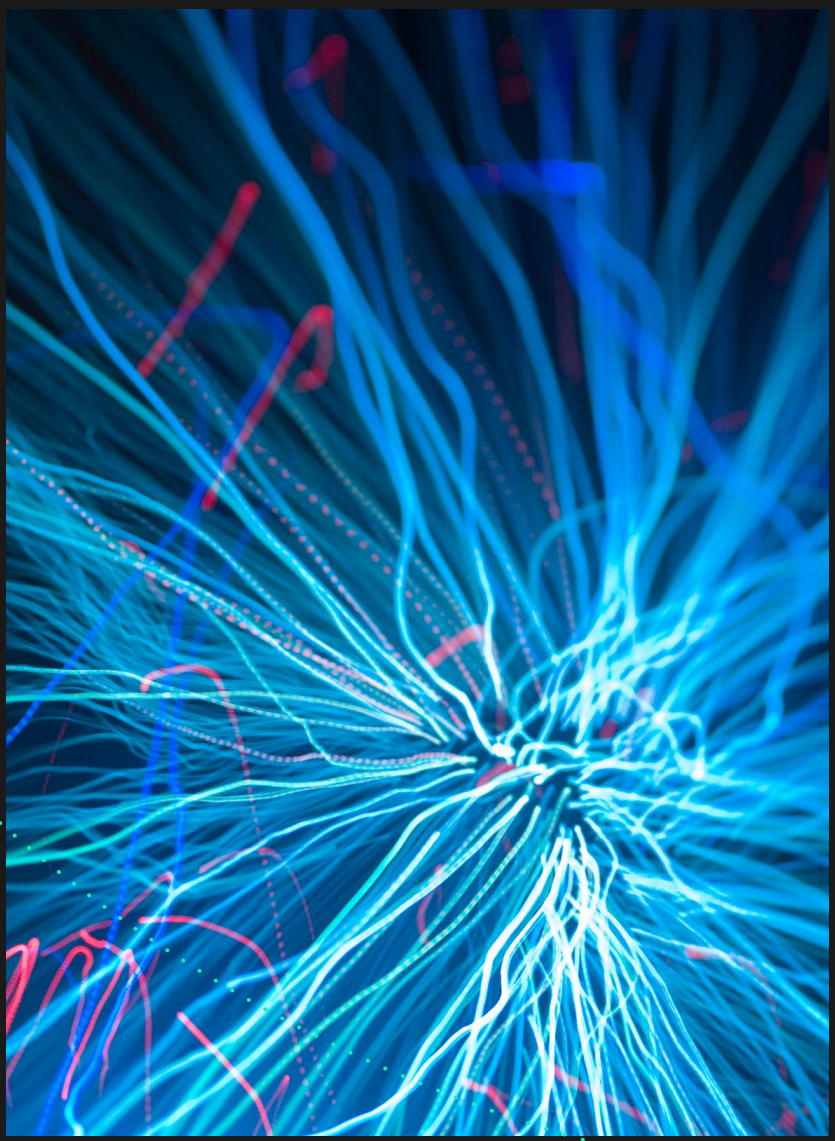
AI in the collective imagination, the new normal ?

The enthusiasm generated by subjects related to artificial intelligence induces a split among the general public, like any disruptive innovation. Between threat and opportunity, the technicality and complexity of the field of artificial intelligence often leads to an erroneous understanding of the stakes, making widespread adoption of AI difficult. The feelings of loss of control over the underlying computing processes, as well as the media coverage of certain use cases that infringe on privacy, are all elements that will require pedagogical efforts to better understand how these algorithms operate. The developing regulatory framework via the General Data Protection Regulation (GDPR) and the AI Act contribute to improving the reputation of AI. The AI Index 2018 reveals that media articles on AI have become 2.5 times more positive between 2016 and 2018. Overall, the public impression of AI will always depend on the purpose of the use cases.



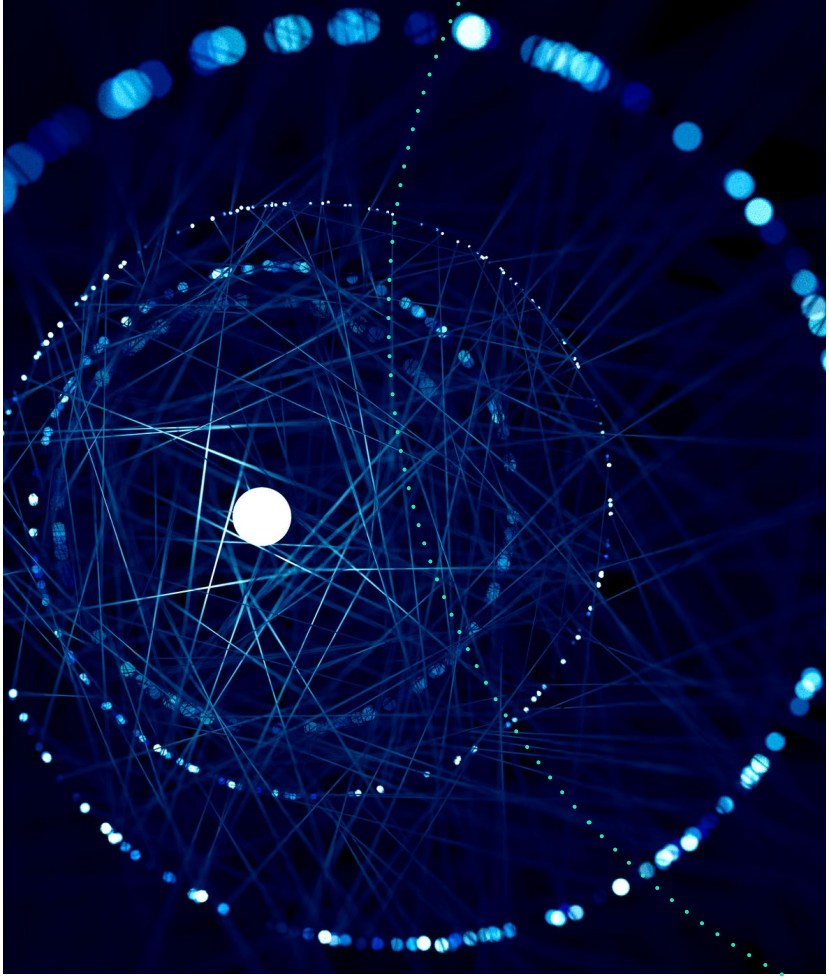
9. AI regulation is maturing but still in its early stages.

The regulation of artificial intelligence is gradually being put in place with the upcoming arrival of the Artificial Intelligence Act, which is in line with the first European regulation around data, the GDPR of May 2018. The AI Act, which is currently being drafted, will be a worldwide reference for a legislative framework for AI, both technically and in terms of its intended purposes. This regulation, which is supposed to regulate and not penalize the development of AI, is designed in consultation with private and public actors in the field of AI.



Towards a «Europeanization» of AI regulation

After having positioned itself as a pioneer in the regulation of personal data with the GDPR, the European Union wants to tackle the whole chain, extending regulation to the processing of all types of data and AI use. This is the aim of the draft regulation called the Artificial Intelligence Act, first published in April 2021. Currently under review of companies, the draft will be submitted to the European Parliament for a vote before being implemented a year and a half later. Given the massive deployment of artificial intelligence in the European ecosystem, this act aims to establish a regulatory and ethical framework for the use of AI, by placing constraints on its techniques and applications. Thus, any implementation of artificial intelligence in the European Union that falls within the material scope of the AI Act will be subject to this legislative framework, regardless of where it is hosted. The legislative grid chosen for this founding act, namely the European Union, is in line with the desire to develop a common policy for the promotion of AI, which is already reflected in a whole series of initiatives. An example of such an initiative is the European Data Portal, which includes more than a million datasets shared between the member countries of the European Union15. This act will enable the pooling of access to data and capitalize on synergies.



Growing legislative pragmatism with similarities to cybersecurity

The AI Act aims to regulate and categorize AI applications according to their level of risk. This allows the protection of citizens against fraudulent practices performed on sensitive data, such as certain real-time biometric identification systems or protection against the unconscious manipulation of behaviors, for example by promoting scandalous and fake news on social media. The risk-based approach has already been adopted in areas such as cybersecurity, with the concept of CCS (Cyber Security Culture) implemented by ENISA (European Network and Information Security Agency). To minimize cyber risk, companies have implemented awareness campaigns and other means to promote a culture of IT security internally and shift behaviors and beliefs towards cyber security.

Similarly, based on risk, the AI Act aims to create an AI culture. Each company involved will have to implement governance and processes in order to prove the mastery of its technologies at all levels from system-design to development. Requirements, such as the systematic drafting of technical documentation serving as references, activity registers or user accessibility requirements, will help prove the compliance of internal systems and devices, particularly in the event of an external IT audit. In addition, the IA Act sets requirements at the supplier, distributor and end-user levels, in order to hold all stakeholders accountable. The inclusion of all actors in the drafting of the text and the specification of roles and responsibilities, allows expanded risk coverage.

The text gives detailed classifications of different AI systems and the associated provisions, which facilitates the proper implementation by the stakeholders by reducing possible misinterpretations. The AI Act will also allow compliant companies to benefit from a better market image and more visibility. This risk classification is in line with the current trajectory of the risk-based approach, which allows for the adaptability of constraints to operational realities. If violations occur, they can lead to record fines ranging from 2% to 6% of the company's global annual turnover.

A disadvantage to international competitiveness

However, given the international competitiveness of the AI market, the aim of the text is sustainable development, not to disadvantage European companies relative to unconstrained international companies. Thus, relaxation of some restrictions are planned in order to avoid penalizing innovations. The European Union plans to set up «regulatory sandboxes'' that will establish a controlled environment for testing innovative technologies for a limited period of time. Companies will be able to access these sandboxes after showing a test plan to the relevant authorities. The European Union wants to promote access to this facility while giving priority to SMEs and start-ups. On the other hand, the regulation is accompanied by a strong budgetary policy to support AI: the Digital Europe and Horizon Europe schemes will devote 1 billion euros each year to Artificial Intelligence projects. In addition, 20% of the European recovery plan must be devoted to the digital transition and AI projects.

(15) https://data.europa.eu/en

10. The closing word by David Martineau.

For Sia Partners, AI has fundamentally changed the game. The needs of our clients, who now have access to terabytes of data thanks to the Internet and the IoT, have evolved. We have chosen to put ourselves in a position to accompany them in the management and exploitation of their data on two fronts :

- Strategy and management consulting, with AI roadmap projects, governance projects and data quality approaches on all sectors covered by Sia Partners. This component has developed quite naturally through the crossroads of our consulting teams and our Data teams.
- The technological component is part of a stronger break within our traditional consulting activities. It is through Heka, our Data and AI ecosystem, that we have chosen to reinvent ourselves in order to attract and grow talent around Data and develop AI software solutions. Heka allows us to have a dual value proposition: Data experts to support our clients' projects at the heart of the Data Labs, as well as industrial AI solutions to further accelerate our interventions and respond to the common problems of many of our clients.

We have built our AI team around a core workforce of 75% Data Scientists, whereas more traditional firms have invested more in Data Analysts. Heka's positioning is therefore closer to the pure players in AI. We have completed this Data Science expertise with profiles of Data Engineers, Devops Engineers, Web Developers, UI/UX Designers. They allow us to embed our algorithmic cores in industrial AI solutions for business users. Today, our team includes more than 200 data experts, in 5 Centers of Excellence in Europe and North America, with an annual growth rate of nearly 100%.

To cultivate our attractiveness and ensure the retention of talent, we invest considerably in R&D. This takes the form of AI accelerators - elementary algorithmic bricks- that we are developing through thematic Labs, such as NLP, Time Series, Computer Vision, etc. They enable our experts to be at the cutting edge of each of these rapidly evolving AI technologies, and to be trained on a continuous basis.

For 20 years, the DNA of Sia Partners' Business Consulting interventions has been «problem-solving» thanks to the sector expertise of our consultants. We use AI to address these challenges differently. This set of expertise and solutions allows us to intervene on very diverse issues. We support a major insurance company on its entire AI strategy and roadmap, just as we work with ski area operators to optimize snowmaking using predictive algorithms and intelligent sensors.

To ensure the proper deployment of these solutions to the standards of our clients' IT teams, we have equipped ourselves with an industrial platform. It ensures rapid and sound developments and guarantees the scaling of our solutions. It is the main tool of our teams, from POC to industrialization. Today it hosts nearly 100 customer platforms, over 1,000 projects and 125,000 pipelines.

Over the next 18 months, we plan to double the size of the team. This will allow us to address clients in new geographies (Middle East, Asia, etc.) and to offer an even wider range of products and services (quantitative and actuarial services, etc.). We will also invest in the development of emerging technologies: Blockchain, Cryptos, Quantum Computing, etc.

Glossary (1/3).

Big Data : The definition of Big Data is as follows: more varied data, arriving in increasingly large volumes and with greater volumes and at a higher speed. This definition is also known as the three V's.
(https://www.oracle.com/fr/big-data/what-is-big-data/)

Cloud : «The cloud» refers to servers that are accessible via the Internet, as well as the software and databases that run on these servers. Cloud servers are located in data centers around the world.
(https://www.cloudflare.com/fr-fr/learning/cloud/what-is-the-cloud/)

Clustering : Clustering is a machine learning method that consists of grouping data points by similarity or distance. It is an unsupervised learning method and a popular technique for statistical analysis of data.
(https://analyticsinsights.io/le-clustering-definition-et-implementations/)

Data-driven : Data Driven, also known as Data-Driven Marketing, is based on an approach that consists of making strategic decisions based on data analysis and interpretation.
(https://www.atinternet.com/glossaire/data-driven/)

Data Lake : A data lake contains data in an unstructured way. There is no hierarchy or organization between data elements. The data is kept in its most raw form and is not processed or analyzed. A data lake accepts and stores all data from different sources and supports all data types.
(https://www.oracle.com/fr/database/data-lake-definition.html)

Data Warehouse : A Data Warehouse is a relational database hosted on a server in a Data Center or in the Cloud. It collects data from various and heterogeneous sources with the main purpose of supporting the analysis and facilitating the decision making process.
(https://www.oracle.com/fr/database/data-warehouse-definition.html)

Deep Learning : Deep learning is a sub-field of machine learning. It consists of algorithms capable of mimicking the human brain through a large network of artificial neurons. The systems are therefore able to learn, predict and decide autonomously.
(https://blog.hubspot.fr/marketing/deep-learning)

Edge Computing : Edge computing is a method of optimization used in cloud computing that consists of processing data at the edge of the network, near the source of the data.
(https://fr.wikipedia.org/wiki/Edge_computing)

Federated learning : Federated learning is a machine learning technique that trains an algorithm on multiple decentralized servers, containing local data samples, without exchanging their data samples..
(https://24pm.com/117-definitions/359-apprentissage-federe)

Hackatons : A hackathon, programming marathon or programmathon is an event during which groups of volunteer developers come together for a given period of time to work on computer programming projects in a collaborative manner.
(https://fr.wikipedia.org/wiki/Hackathon)

Hyper-segmentation : A marketing strategy that consists of developing offers in such a way as to best respond to the needs of the consumer, or even to create custom-made products.
(https://academy.visiplus.com/ressources/definition/hypersegmentation)

AI : The term «Artificial Intelligence», created by John McCarthy is «the set of theories and techniques implemented to create machines capable of simulating human intelligence» according to Larousse dictionary. It is defined by one of its creators, Marvin Lee Minsky, as «the construction of computer programs that perform tasks that are, for the time being, more satisfactorily accomplished by human beings because they require high-level mental processes such as perceptual learning, memory organization and critical reasoning».
(https://fr.wikipedia.org/wiki/Intelligence_artificielle#D%C3%A9finition)

Glossary (2/3).

IPO : Initial Public Offering

Low-code : Low-code development has evolved to take advantage of visual design tools (notably drag-and-drop modeling and point-and-click interface creation) and enable the rapid creation, launch, use and modification of powerful business applications.
(https://appian.com/fr/low-code-basics/low-code-defined.html)

ML : Machine learning is a form of AI that refers to a system that bases its performance on learning from data. It is the combination of a learning algorithm and its data.
(https://www.ibm.com/fr-fr/analytics/machine-learning)

MOOC : Training provided on the Internet and accessible to all.
(https://www.larousse.fr/dictionnaires/francais/MOOC/10910887)

NLP : NLP for Natural Language Processing is a discipline of AI that focuses on the understanding, manipulation and essentially on the understanding, manipulation and generation of natural language by machines.
(https://datascientest.com/introduction-au-nlp-natural-language-processing)

SME : Small and Medium-sized Enterprises

POC : A proof of concept (POC) or demonstration of feasibility is a realization whose purpose is to show the feasibility of a process or an innovation.
(https://fr.wikipedia.org/wiki/Preuve_de_concept)

Reinforcement Learning: : Reinforcement Learning (RL) refers to a class of machine learning problems, of which of machine learning problems, whose goal is to learn, from successive experiments, what to do in order to find the best solution thanks to a system of rewards or penalties.
(https://dataanalyticspost.com/Lexique/apprentissage-par-renforcement/)

RGPD : The EU Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of individuals with regard to the processing of personal data and on the free movement of such data is a European Union regulation that constitutes the reference text on personal data protection. It strengthens and unifies data protection for individuals within the European Union.
(https://fr.wikipedia.org/wiki/R%C3%A8glement_g%C3%A9n%C3%A9ral_sur_la_protection_des_donn%C3%A9es)

SaaS : «Software as a Service» is an application software solution hosted in the cloud and operated outside the organization or company by a third party, also known as a service provider.
(https://www.cegid.com/fr/faq/quest-ce-que-le-saas/)

Scoring : In marketing, scoring is a technique used to assign a score to a customer or prospect. The score obtained generally reflects the probability that an individual will respond to a marketing solicitation or belong to the target.
(https://www.definitions-marketing.com/definition/scoring/)

Système SI : The information system (IS) is an organized set of resources that enables the collection, storage, processing and distribution of information, usually through a computer network.
(https://fr.wikipedia.org/wiki/Syst%C3%A8me_d%27information)

Transformers : A neural network architecture developed that relies on self-attention mechanisms to transform a sequence of input embeddings into a sequence of output embeddings without relying on convolutions or recurrent neural networks. A transformer can be thought of as a stack of self-attention layers. A transformer can include one of the following: an encoder - a decoder - both an encoder and a decoder.

Glossary (3/3).

(https://developers.google.com/machine-learning/glossary?hl=fr#Transformer)

Smart data : Smart Data translates as intelligent data. It is a strategic approach to data processing, to extract from the mass of data to which we have access the ones that are of interest to us, in an ever shorter way.
(https://www.appvizer.fr/magazine/analytique/aide-decision/smart-data)

Your contacts.

Nasri **Badaoui**

Manager
Data Science
PARIS
Nasri.badaoui@sia-partners.com
+33 6 62 62 88 47

Editorial team.

Nasri **Badaoui**

Manager Data Scientist.

Eya **Kalboussi**

Consultant Data Scientist.

Yoann **Russello**

Consultant Data Scientist.

Lisa **Griere**

Consultant Data Scientist.

Hugo **Perrin**

Consultant Data Scientist.

Thanks.

- **Pierre Capelle** — Partner Data Science - Sia Partners
- **Adnan El Bakri** — Founder of Relyfe, Surgeon, author of “predict cancer with AI”
- **Pierre Folk** — Associate Partner Assurance - Sia Partners
- **Jeanne Fourcade** — Ethics and Compliance Specialist - Sia Partners
- **Romain Laurans** — Managing Director Data Science - Sia Partners
- **Pierre Leplatois** — Partner Data Science - Sia Partners
- **Sophie Le Goff** — Assurance and Compliance Partner Sia Partners
- **Thomas Rocafull** — Partner Financial Services - Sia Partners
- **Matthieu Thiboust** — CTO - Sia Partners
- **Philippe Saade** — Specialist AI
- **David Martineau** — Deputy CEO - Sia Partners
- **Matthieu Courtecuisse** — CEO - Sia Partners

- Belgium
- Canada
- France
- United Kingdom
- Germany
- Hong Kong SAR
- Italy
- Ireland
- Japan
- Luxembourg
- Morocco
- Netherlands
- Qatar
- Saudi Arabia
- Singapore
- United Arab Emirates
- United States
- Panama*

* Sia Partners Panama, a member of the Sia Partners network

About Sia Partners.

Sia Partners is a next-generation management consulting firm and pioneer of *Consulting 4.0.* We offer a unique blend of AI and design capabilities, augmenting traditional consulting to deliver superior value to our clients. With expertise in more than 30 sectors and services, we optimize client projects worldwide. Through our *Consulting for Good* approach, we strive for next-level impact by developing innovative CSR solutions for our clients, making sustainability a lever for profitable transformation.

www.sia-partners.com