GETTING COMPANIES ON BOARD WITH RESPONSIBLE ARTIFICIAL INTELLIGENCE

March 2023
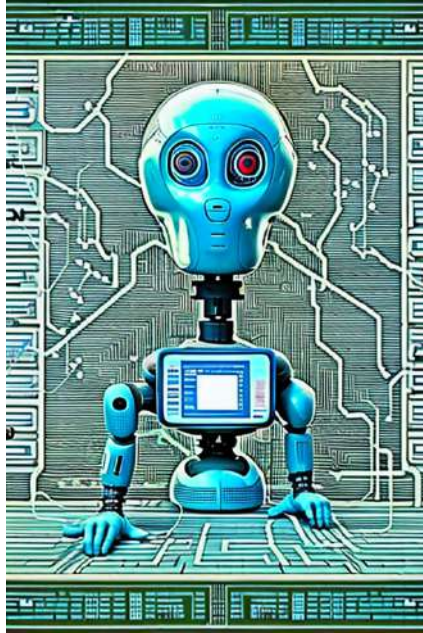
AmCham
AMERICAN CHAMBER
OF COMMERCE IN FRANCE

ESSEC
BUSINESS SCHOOL

Executive summary

Artificial Intelligence (AI) has been qualified by OECD as a General Purpose Technology (GPT), same as electricity and internet in the past. In other words, it is a game changer. It has the potential to revolutionize Small and Medium-sized Enterprises (SMEs) by enhancing their competitiveness. However, successful AI implementations require careful planning, consideration given to human factors, proper data governance, the development of trust in the AI-based systems, and responsible implementation, with ethics and bias reduction, as well as AI's environmental impact, also being key success factors. This is particularly true for SMEs which do not necessarily have the resources of larger organizations. Note that in this white paper, the terms Artificial Intelligence (AI) and Machine Learning (ML) are used interchangeably.

This white paper is the outcome of an eight-months long conversation among AI practitioners, led by ESSEC Metalab and Amcham France. Throughout monthly working sessions, leveraging their past experiences and informed expertise, members of the working group have developed this contribution with the aim to provide concrete reflections, takeaways, and recommendations for SME executives.

Note that in this white paper, we include French ETIs and the largest PMEs under the term "SME", as explained in Chapter 11 - Glossary. The Glossary should be seen as a key resource in this white paper and the reader is invited to having a look early on, as many AI-specific terms that may need clarification are used in this document.

In this contribution, we first discuss the importance of establishing a strategic vision for AI implementation. SMEs are advised to first assess their AI and data maturity score, create a roadmap for implementation, and consider hiring an AI dedicated team. We provide examples of strategy frameworks that apply specifically to AI transformation. We recommend SMEs enter into practice-sharing with startups, other companies, or research organizations to increase their AI literacy, develop an understanding of the ecosystem of resources that's available to them, and get access to successful case studies and learn from those.

Second, we emphasize the importance of the human factor for fostering a collective understanding of AI, at any levels of the organization. Companies must create a vision and agenda for AI that guides the entire organization, while communication, transparency, and understanding will remain crucial for building wide acceptance of AI. The impact of AI on employment should also be considered, bearing in mind that upskilling the existing workforce is vital for successfully putting AI to productive use.

Third, we highlight the importance of acquiring and cultivating the right skills, especially given the global shortage of data scientists and engineers. For that, short-term workarounds are needed, and upskilling of non-AI specialists is encouraged. Cross-disciplinary education and training are also necessary for success, while partnerships with research laboratories and universities can provide fast-track access to high-skilled profiles.

We then underscore the importance of proper data governance and management practices to enable scalability. As the quality and quantity of data impact the outcomes of AI systems, the data lifecycle should be managed with proper governance. Poor data quality is indeed a common pitfall, leading to possibly disastrous outcomes. Additional challenges to data and AI governance in SMEs include a lack of executive sponsorship, poor communication, lack of training and education resources, as well as bias and unfairness.

We insist on the importance of building trust in AI, which is critical for adoption. Fairness, explainability, accountability, transparency, reliability, robustness, privacy, security, human involvement and oversight are all necessary ingredients. Regulators, from EU, China and the US, are actively working on the regulation of AI (i.e. AI Act) and have introduced legal frameworks to ensure a trustworthy implementation of AI across industries.

We highlight the importance for AI projects to be governed by responsible principles. SMEs may face challenges due to a lack of organizational preparedness, but they should take action to reduce the digital carbon footprint of AI. The European Commission, OECD, US Federal Government, China's Ministry of Science and Technology, and the Global Partnership on AI all propose guidelines for responsible AI development which should be considered by SMEs when embarking on AI technologies.

Last, we have a brief look of what SMEs should expect from AI and other disruptive digital technology in the near future, in order for them to keep their eyes on the ball that keeps accelerating.

An appendix has been added for reference on a few possible public funding alternatives available to French SMEs; the upcoming AI Act; and a few additional resources we think our readers may find useful to go beyond this white paper.

We hope this contribution will be regarded by SMEs as a useful set of reflections and recommendations for embarking on the path of successful adoption of AI. By doing so, SMEs can remain competitive in an increasingly complex, resource-limited and fast-moving world.
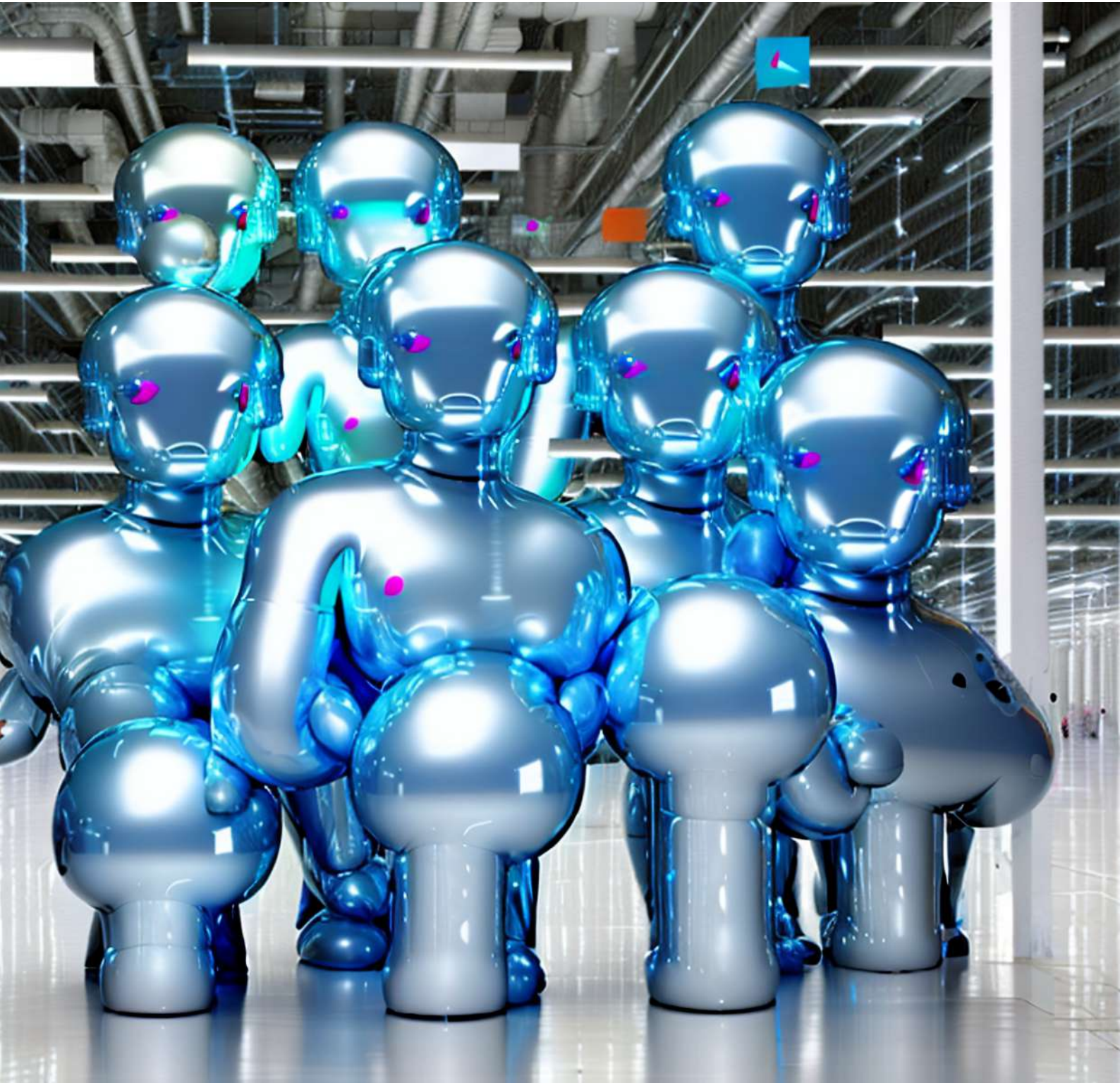
Benoit Bergeret & Julien Malaurent
Co-Directors, ESSEC Metalab for data, technology and society
Co-chairs, AmCham France Working Group on AI Adoption

Contents

CHAPTER 1

WHY THIS WHITE PAPER ?

> *Just as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don't think AI will transform in the next several years.*

Prof. Andrew Ng (Stanford, Google Brain, Baidu)

According to a recent report by OECD (1), SMEs in Europe, when compared to large organizations, face specific barriers to Artificial Intelligence (AI) adoption. Among these, a lack of data culture and awareness about what benefits and returns AI could bring, a need for retraining managers and workers, high sunk costs for internalizing AI, plus a need for engaging complementary investments not properly supported by existing public financing alternatives (2), few evidence and little visibility on the returns on investment, and reputational and legal risks have led to a lag in AI adoption, compared to their US and Chinese counterparts in particular. Because of globalization, this slow uptake risks creating a competitiveness rift, ultimately hurting the European and French economies.

These barriers may be more pronounced in France, where SMEs are characterized by a lower level of digitization compared to larger firms, and a higher degree of regulation and bureaucracy that may hinder innovation, and a lower availability of skilled AI talent and data scientists, despite France hosting about 12% of the available European AI skills.

However, the report also notes that SMEs can benefit from AI in two ways : AI can modify the business environment in which SMEs operate, making it more favorable for them, as well as enabling them to modify their business models and practices, which can lead to increased productivity, larger reach, and the ability to grow.

In the rapidly evolving AI technology landscape, it is therefore crucial for French SMEs to realize the potential of AI but also the consequences of failing to adopt it. Without incorporating AI, French SMEs risk being left behind and losing their competitive edge.

Aware of this state of affairs, the French government has entered into the second phase of its National AI Strategy (3). Initially endowed with nearly 1.5 billion euros over the 2018-2022 period, the first phase of the national AI strategy aimed to position France as one of the world leaders in AI technologies.

On November 8th, 2022 the French government launched the second phase of the National AI Strategy, with a total of €2.22 billion earmarked over the next five years. Training and education in AI accounts for 50% of the effort, and 40% of investments are dedicated to innovation support and economic measures, with a goal of supporting at least 500 SMEs in the adoption and use of AI solutions by 2025, to accelerate their modernization.

1. *"Artificial intelligence: Changing landscape for SMEs", OECD*
2. *"AI and Europe's medium-sized firms: How to overcome an Achilles' heel", Benoit Bergeret, OECD Observer, 2019*
3. *"Stratégie Nationale pour l'Intelligence Artificielle", SNAI (https://www.intelligence-artificielle.gouv.fr/fr)*

In line with this policy, but independently, this white paper has been developed as a tool to help ease SMEs into the universe of AI with guidance aimed at their leadership.

The audience of this document is the executive team in SMEs in France who have not yet embarked on their AI transformation journey.

It should be complemented by leveraging more specific resources such as AI business management executive guidebooks, participation in ecosystem initiatives and communities, and, when applicable, third-party strategy and project consulting resources (as all listed in Appendix C).

Based on real-world inputs from the participants to the AmCham working group that gave it birth, and on years of experience its authors have advocating for and deploying AI in businesses, this white paper aims to demystify AI technologies, and to provide a clear understanding of why French and European companies have been slow to adopt AI. It provides an outline of the obstacles that hinder AI development, and suggests key aspects to carefully consider when thinking about deploying AI in a business.

We hope that it promotes deep and broad corporate introspection, as AI is a technology like no other in its business implications: its transformative potential is tremendous, and it should not be taken lightly.

It is important for readers to note that this contribution does not attempt to provide an exhaustive list of possible business applications of AI, nor does it aim at including an exhaustive catalog of resources leadership can use. Rather, it focuses on providing information, strategic frameworks, and guidelines and best practices to inform business leaders who are embarking on the strategic reflection about AI.

This includes addressing important CEO questions such as: *How can AI be integrated into our business strategy? How to prioritize AI projects? What skills and resources are required for AI implementation? What ethical considerations should be taken into account?* This contribution also points toward useful resources for companies to succeed in implementing AI. As we understand that not all readers may be familiar with the complex concepts and terminologies associated with AI, we have included a glossary to help, educate and inform (chapter 11).

We hope that this white book will serve as an original, fresh, and valuable resource for SME leadership who are embarking on their strategic reflection about AI, and that it will assist them in realizing the benefits of adopting AI while mitigating potential risks.

We are convinced SMEs can develop further competitiveness by leveraging the power of AI to drive growth, innovation, agility and success.

Good luck to everyone in this endeavor!

Opinions expressed are those of the authors and do not represent a consensus across all participants to the working group.

CHAPTER 2

CONTEXT: AI IS THE FUTURE OF THE ECONOMY

TAKEAWAYS

- **AI adoption is a far-reaching undertaking with implications across the European economy and requires upfront investment and costly operational and organizational adjustments.**

- **For SMEs, AI has already brought about higher productivity and efficiency, increased decision-making speed, and automation of manual processes.**

- **Europe is lagging behind other regions in the adoption and development of AI.**

- **Factors contributing to this slow adoption by EU SMEs include imperfect public funding instruments, data sourcing issues, lack of available expertise, and risk-averse corporate cultures.**

- **To support the growth of AI technologies in Europe, the European Commission has taken several measures, including the Digital Europe Programme, Horizon Europe programme, and InvestEU Programme.**

- **Member states have released their own national AI strategies and funding plans.**

AI adoption: a much needed transformational evolution

Recent years have seen rapid advancements in Artificial Intelligence (AI) driven by Machine Learning (ML) algorithms, data-driven statistical methods, improved computing power, data acquisition and storage technologies. AI is widely recognised as a General Purpose Technology (GPT), with the potential to transform entire economies and society.

The adoption of AI is a far-reaching undertaking with implications across the European economy, unlike digital technologies of the past. According to Rogers' Diffusion of Innovations (DOI) theory, new technologies such as AI spread incrementally rather than instantaneously. At the level of an individual company, adopting AI requires defining a strategy, assessing data maturity, harvesting and preprocessing data, then creating and evaluating training models, which then need to be evaluated, deployed, monitored, and maintained. This process takes time. The decision to adopt AI requires upfront investment in infrastructure and in the technology itself, while the potential returns are unknown. Deployment implies costly operational and organizational adjustments which may put some companies at a disadvantage in the early stages.

AI will have important impacts on business culture. Changes to the traditional organization of a company should be expected as an increasing number of workers have access to advanced technologies. Businesses will need to shift management techniques and leadership: until now, AI adoption was the result of management decisions, pushed through the organization. But with new AI tools such as generative AI, as popularized by ChatGPT, the decision to adopt AI for their own use is now often going to be made by employees, which may ultimately help reduce friction that is often observed nowadays.

Key benefits for SMEs

AI is becoming a major force in many industries, though usage in each sector is concentrated among small numbers of early adopters. AI solutions have brought about higher productivity and efficiency.

AI can enable faster decision-making by detecting patterns early. In healthcare, for example, AI can analyze natural language processing to identify potential heart attacks based on speech patterns and unconscious signals. **Increased efficiency can be achieved through the automation of manual processes**.

For example, AI can automatically create machine-readable legal and compliance documents, drastically reducing the time needed for drafting and analysis. AI can also be used for language and speech recognition, enabling businesses to implement chatbots, therefore minimizing the amount of time employees spend on calls. In the primary and secondary sectors, such as agriculture and manufacturing, AI is used to optimize production and automate processes, often with the help of robots. AI has also been applied to the tertiary sector, leading to the development of more advanced and creative applications, such as natural language processing, recommendation engines, sentiment analysis, and even art and design.

Recent McKinsey research suggests that with the adoption of AI, economic growth rates may double and labor productivity may increase by up to 40% by 2035.

The state of AI adoption in France, Europe, and elsewhere

Europe is lagging behind other regions in the adoption and development of AI. According to a 2019 study, France ranks 7th globally in the number of AI startups. A 2020 survey by IBM revealed that 30% of French companies surveyed have actively deployed AI in an operational framework, 36% are exploring but have yet to deploy AI, and 26% have not shown any interest. Adoption rate was as high as 60% among Chinese companies, according to a report published by the China Artificial Intelligence Development Report (CAIDR). The report also noted that small and medium-sized enterprises in China were catching up in AI adoption, with more than 30% of them having started to use AI. And, according to a report by McKinsey & Company, at the same time more than 60% of businesses in the United States have adopted at least one AI application.

In Europe, the US, and India, a substantial proportion of firms using AI operate in the information and communications sector and in professional, scientific, and technical activities; whereas in China, roughly half of the country's firms currently that have adopted AI are in the manufacturing sector, reflecting China's overall dominance in this area. In Japan, there is a more even distribution of AI usage across industries, with over 25% operating in manufacturing and a notable presence in wholesale and retail trade, similar to implementation in the US and Canada.

Despite this, SMEs still lag behind the United States and China in the adoption of AI. Factors contributing to this slow adoption have been identified and include imperfect public funding instruments, data sourcing issues, lack of available expertise, and risk-averse corporate cultures. Research shows that large firms are significantly more likely to adopt AI solutions, and much more quickly - three to five times faster than SMEs.

AI has the potential to redefine the global economic landscape. In Europe, SMEs account for the industrial segment only, over 20 million jobs and over 50% of gross value added, just in the industrial sector (1). For them, missing out on this technology would be as detrimental as ignoring the rise of the internet in the early 21st century and will have drastic consequences for European businesses and for Europe's strategic autonomy. Policymakers must support increased adoption of AI, first by seeking to understand the barriers preventing European firms from adopting AI and ensuring that AI technologies are trustworthy, user-friendly and valuable in day-to-day work.

Key factors driving growth of AI adoption in Europe

Talent availability

The availability of skilled personnel to meet the rising demand of AI adoption is critical in determining a region's competitiveness. While Europe has the largest pool of AI researchers, businesses in the EU are falling behind their American counterparts when it comes to employees with core AI skills or even basic expertise. This disparity highlights the need for effective strategies for AI to reach its full potential in European businesses. Businesses must train personnel for new roles such as Model Owner, MLOps and ModelOps specialists (2).

The education ecosystem must take into account the core competencies to support the large-scale deployment of ML models. For instance, Hub France IA, a nonprofit labeling itself "the operating system of the AI ecosystem in France", has been tasked with the development of a nationwide framework for AI competencies (expected to be published in mid-2023), which will help companies and educational institutions align on the new needs.

Data availability

Data plays a crucial role in the development of AI technologies and is a vital resource for SMEs looking to create new products and services. AI systems are trained using an abundance of data.
The global volume of data is expected to surge from 33 zettabytes in 2018 to 175 zettabytes in 2025, offering the European Union significant opportunities to implement AI systems. Limited access to large data sets, however, poses a challenge for SMEs, which often lack resources available to larger firms. In addition, the digital market in the European Union is fragmented across its member states, hindering the creation of large, unified data sets for AI model training.

---

1. "AI and Europe's medium-sized firms", B.Bergeret in OECD Observer (2019)
2. The Model Owner is the person who takes the responsibility for the usage of a specific ML use case/model within an organization. Each use case is assigned a Model Owner who decides, among others, whether to go to production, considering the state of the development, documentation, performance, and proposed implementation, and monitoring of the models. Model Owners are accountable for the performance, governance, and compliance of ML models. For MLOps and ModelOps, please refer to Chapter 11 - Glossary

Convergence of AI, IoT, and the blockchain

The integration of AI, the Internet of Things (IoT), and blockchain technologies will create new platforms, products, and services, IoT provides a data source, AI processes the data, and blockchain technology records transactions, tracks assets, and provides security. IoT and blockchain will enable AI by providing large datasets for AI to learn from, enabling infrastructure improvements in smart cities, for example, that will enhance the quality of life for residents. Blockchain may also establish open, decentralized markets for data, AI models, and resources.

EU Programmes for AI

To support the growth of AI technologies in Europe, the European Commission has taken several measures. In April 2018, the Commission unveiled its AI strategy, which aims to increase the coordination and cooperation of investments in AI research and innovation (R&I) from both public (EU and national) and private sources. The consortium of funders invested €20 billion from 2018 to 2020 and more than €20 billion annually in the following decade.

Key programs in the 2021-2027 EU budget cycle include: the Digital Europe programme, which focuses on building the EU's digital capabilities, the Horizon Europe programme, a €95 billion R&I initiative, and the InvestEU programme, which combines the European Fund for Strategic Investments and 13 EU financial instruments. With an EU budget guarantee of €75 billion, InvestEU is designed to support investment projects via the EIB Group and national promotional banks. The funds will be allocated to R&I (€10 billion), SMEs (€10 billion), Social Investments and Skills (€3.6 billion), Strategic EU Investments (€31 billion), and Sustainable Infrastructure (€20 billion).

Member states have released their own national AI strategies and funding plans. The German government, for example, pledged to invest €3 billion between 2019 and 2025. The French government allocated €1.5 billion by the end of 2022, with €700 million dedicated to research. Italy has set aside €1 billion until 2025, and the Dutch government will provide €45 million per year until 2023, along with €18 million for the development of supercomputers and €200 million for AI training.

Public initiatives in AI and blockchain from the United States and China

The United States has prioritized funding for AI development at the federal level. In 2018, the US Department of Defense announced a $2 billion, five-year AI technology campaign, called "AI Next," led by the Defense Advanced Research Projects Agency (DARPA). In 2019, President Trump initiated the American AI Initiative to promote US leadership in AI. MIT Technology Review reported that the US government planned to spend $5.9 billion on AI research in 2020 as part of this strategy. The 2020 National Artificial Intelligence Initiative Act allocated another $6.5 billion to AI R&D, education, and standards development.

China has also been a leader in financing AI to promote economic development. In 2017, the Chinese government introduced the New Generation AI Development Plan, which aimed to establish China as the global leader in AI. The plan involved the collaboration of the government and the private sector, promoting investments in AI R&D and initiatives to support its implementation. Lack of available data makes it difficult to determine the total value of these investments, but government spending on AI R&D is estimated around $10 billion per year.
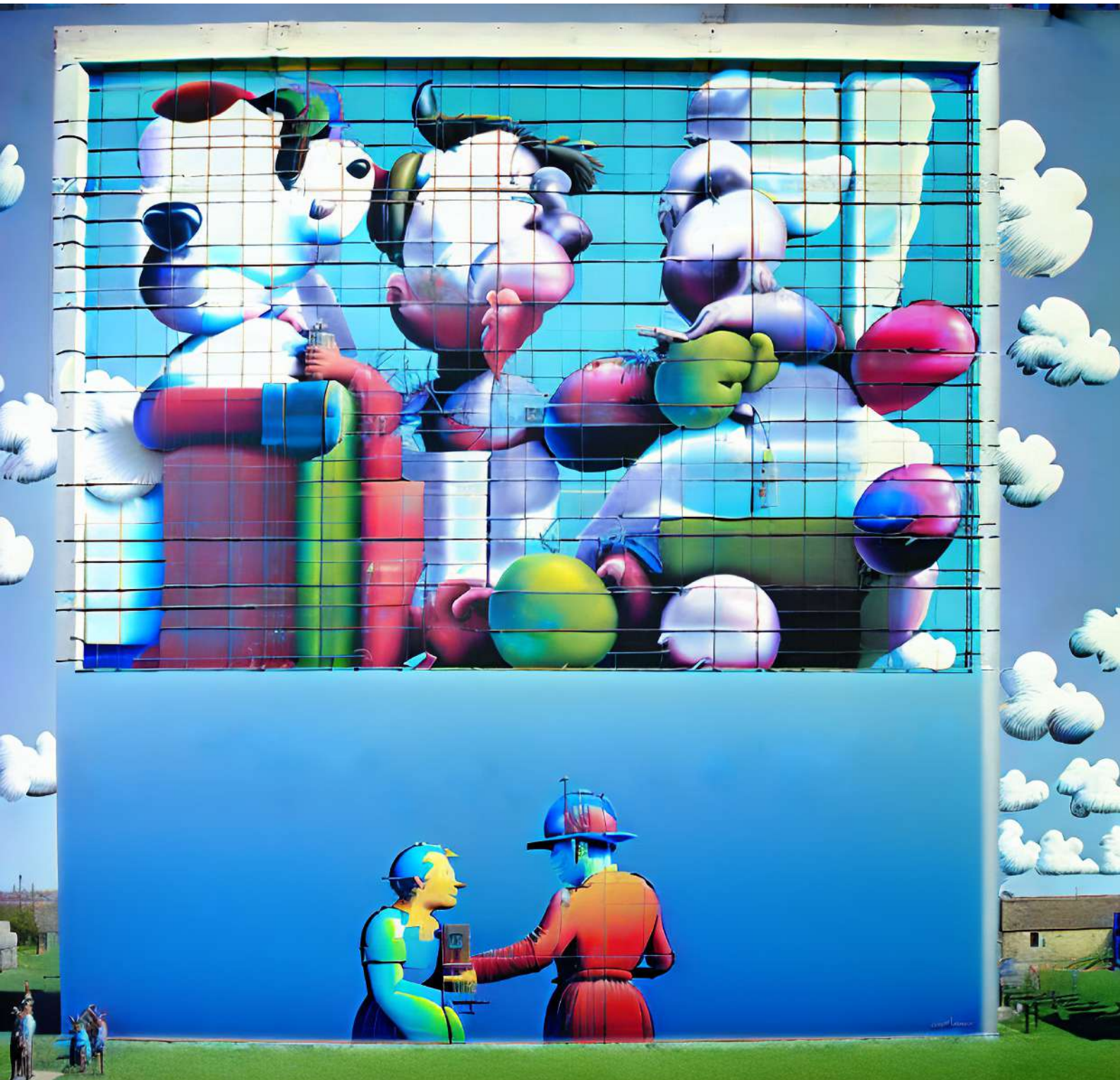
CHAPTER 3

WHERE TO START?

TAKEAWAYS

- **AI adoption strategies should be developed based on the specific business goals and overall strategy of the company.**

- **There is currently no established, universal framework to guide AI adoption. However, AI cannot be implemented without data, and a "data x AI maturity framework" has been developed to assess organizations.**

- **This framework allows companies to gauge where they stand today and develop a comprehensive AI adoption strategy.**

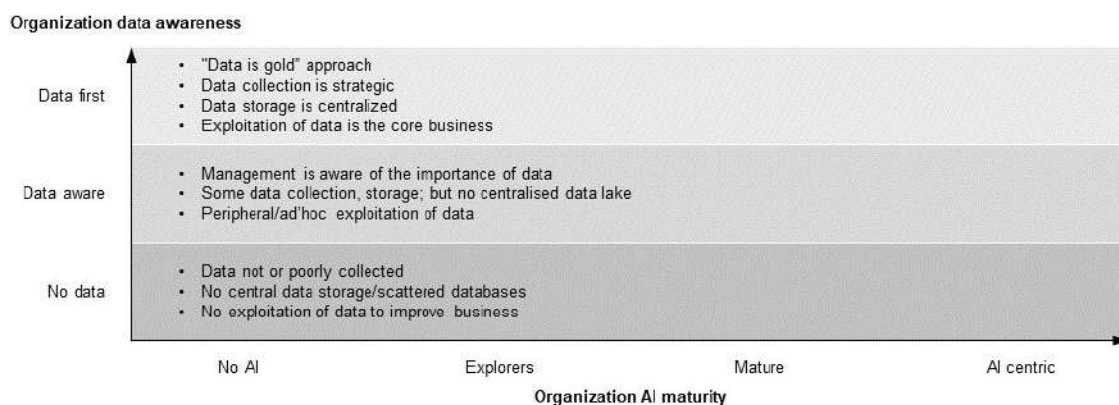- **Not all companies should aim to become AI first businesses.**

This paper is not intended as a practical guide for individual businesses. Rather, our goal is to present frameworks to facilitate leadership in deciding AI adoption strategy. The strategy itself should be developed by individual companies in consideration of their specific business goals and overall strategy.

The data x AI maturity framework

Business frameworks are tools that provide a structured approach for organizations to develop and implement successful, effective strategies. Examples include SWOT analysis, Porter's Five Forces, (which analyzes the competitive forces in an industry and helps organizations understand their competitive position), Business Model Canvas, and Balanced Scorecard.

We present here a framework (1) that has been developed based on years of practical experience and expertise in assisting major companies with their AI adoption strategies. AI cannot be implemented without data. To understand a company's potential for AI adoption we must consider the maturity of the business in terms of data and AI:

Data maturity can be assessed by looking at the existence, governance, exploitation, value extracted, and strategic importance of data for the organization:

Organization data awareness

Data first
- "Data is gold" approach
- Data collection is strategic
- Data storage is centralized
- Exploitation of data is the core business

Data aware
- Management is aware of the importance of data
- Some data collection, storage; but no centralised data lake
- Peripheral/ad'hoc exploitation of data

No data
- Data not or poorly collected
- No central data storage/scattered databases
- No exploitation of data to improve business

No AI   Explorers   Mature   AI centric

Organization AI maturity

© strategies.ai, 2023. Used with permission.

1. This framework is © strategies.ai, Paris, France, offered here for reuse with credit given: "framework © strategies.ai"

AI maturity is best assessed by looking at the way AI shapes processes and investment:

Organization data awareness

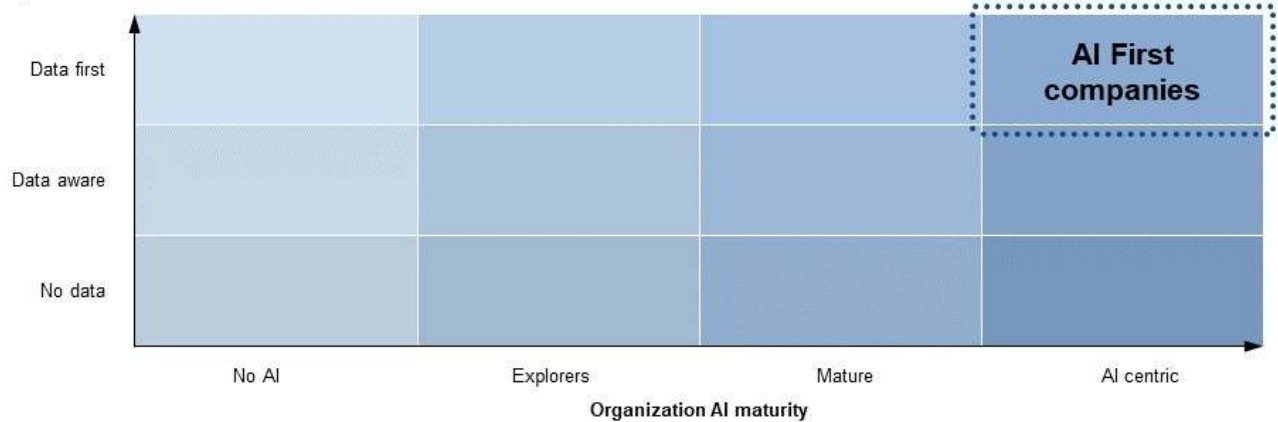|	No AI	Explorers	Mature	AI centric
Data first	• AI? What's that?	• No clear AI strategy • Some level of executive or middle management awareness • AI misunderstood by most, seen as weird or dangerous • AI often a tool for data scientists • Testing and PoC left to the operational teams • Non-core	• Clear AI strategy • C-suite commitment to leveraging AI • AI resources in place • AI and DS work together • AI culture spreading to mid-levels • AI projects in production • Core and non-core	• Potential of AI drives the strategy • C-suite lives and breathes AI • Organization is designed around AI applications • AI is in the DNA of the company • AI fuels the core business
Data aware				
No data				

Organization AI maturity

© strategies.ai, 2023. Used with permission.

In "AI first" companies:

• Organization is designed around data & AI
• Data collection and governance are recognized as strategic and implemented accordingly at all levels of the organization
• Value is derived from exploitation of data, and maximized by the use of AI
• New businesses are created that fully leverage AI.

Organization data awareness

	No AI	Explorers	Mature	AI centric
Data first				**AI First companies**
Data aware				
No data				

Organization AI maturity

© strategies.ai, 2023. Used with permission.

Not all companies are – or should aim to become – "AI First" businesses! In most large European corporations, data is generated internally; but when considering SMEs, the amount of data, and therefore the data maturity, varies a great deal. Some sectors are by nature more data mature than others, such as when core business involves large volumes of data: banks, travel operators, or insurers are pretty much already "data first" companies, whereas manufacturing companies, for example, are at best "data aware".

AI has often been implemented in only certain areas of many businesses, making the majority of those "explorers" on the AI maturity scale.

However, SMEs tend to scatter more broadly across the spectrum for both data and AI maturity:

© strategies.ai, 2023. Used with permission.
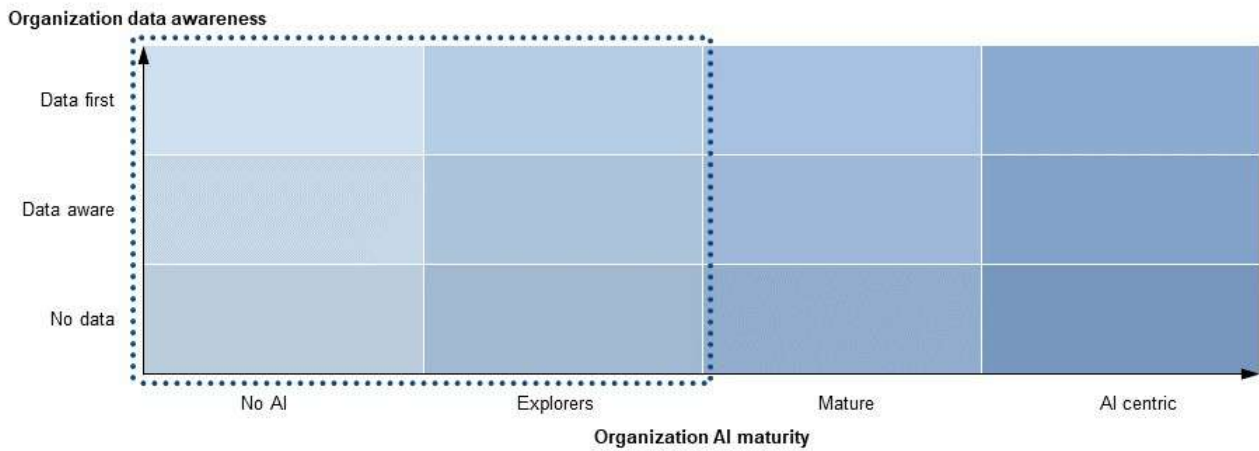
This framework allows companies to gauge where they stand today and where they aim to be in 2-3 years, leading to constructive strategic thinking, ultimately supporting the development of a comprehensive, well integrated AI adoption strategy.

This framework can also be used as a tool for communication among boards, executives, and with company personnel.

Case study: mid-size bank

In 2019 the strategies.ai team collaborated with a mid-size European bank (2,500 M Euro income, 2019) based on this framework. With a data maturity plan already engaged, the bank created their AI strategy:

**Consolidate the foundation**

1. Clearly spell our AI ambition and strategy
2. Build centralized structured data infrastructure (modulo GPRS)
3. Build central Ai expertise center (<> data science; existence independent of strategy)
4. Develop an AI culture throughout the organization (education, community building)
5. Identify first projects with clear KPIs and ROI, going for the core businesses wherever possible (work with outside experts)

© strategies.ai, 2023. Used with permission.
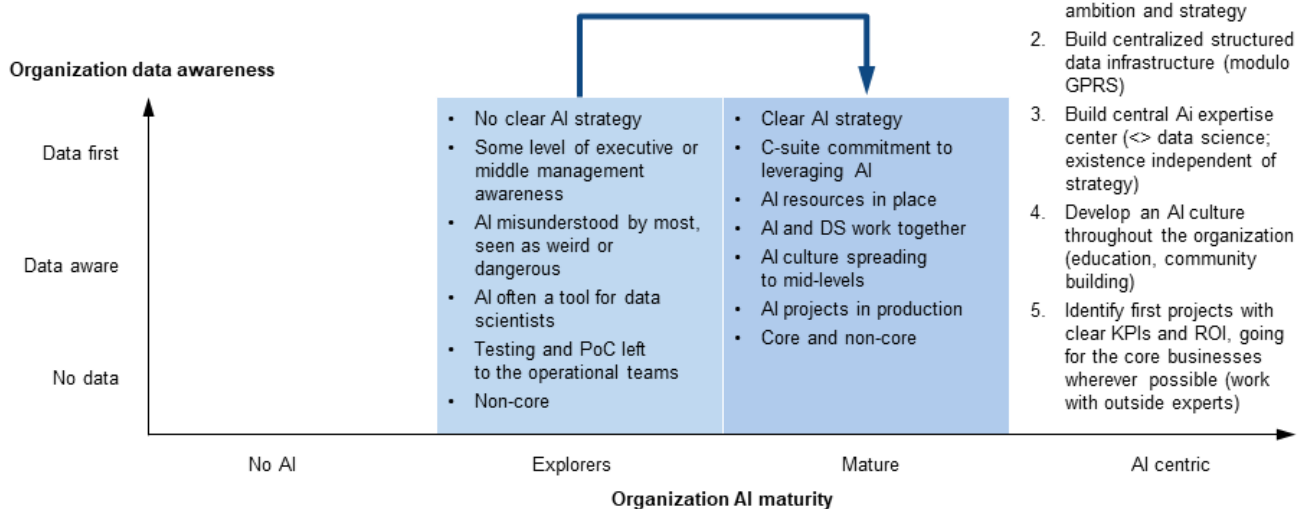
CHAPTER 4

DEVELOPING A VISION
FOR AN EFFICIENT AI STRATEGY
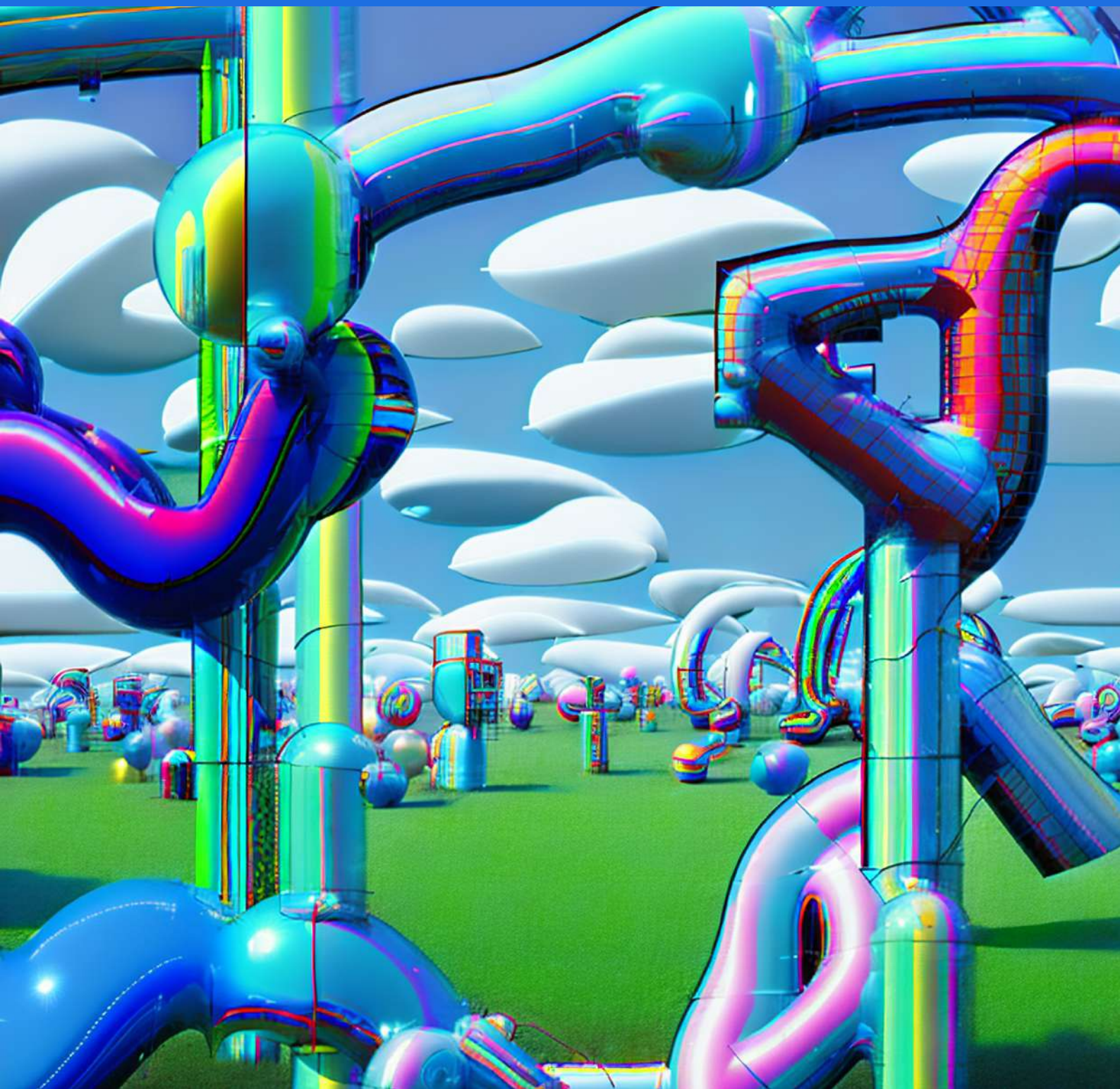
TAKEAWAYS

- **AI adoption can cause a fundamental reconsideration of thinking patterns, operational methods, and business processes.**

- **Business leaders should anticipate the impact of AI and involve all stakeholders in the process. SMEs should establish a clear and comprehensive strategic vision before deploying AI.**

- **Adopting AI requires setting a clear strategy, including an assessment of the company's data maturity, development and implementation of a data governance plan, and data harvesting and preprocessing strategies.**

- **SMEs should use the «Make or Buy» framework to navigate the variety of options available for the implementation of AI.**

- **Utilizing exchange platforms and cross-industry ecosystems can connect SMEs with others in their industry and provide opportunities for collaboration and knowledge sharing, especially on the data front.**

"Start brainstorming about the particular business use cases in your organization that can benefit from massive amounts of data to assist in day-to-day tasks. Remember, AI is good at recognizing patterns, and humans are good at understanding when those patterns have meaning versus when they are spurious correlation."

Ngan Khanh Nguyen - Chief Data Officer at McDonald's

What is at stake?

Artificial Intelligence (AI) and data technologies are not like other digital technologies. Because they can cause a fundamental reconsideration of thinking patterns, operational methods, and business processes, their adoption must be treated as a potential transformation. The impact of AI adoption on people, jobs, organizations, and companies as a whole can destabilize established structures and procedures from the factory floor to the C-Suite.

Traditional approaches to deploying new technologies should be considered with caution, as they have proven unable to address the specific challenges and opportunities of AI. **Business leaders should do their best to anticipate the impact of AI, and all stakeholders should be involved in the process from the start.**

AI technologies can radically transform the characteristics of companies of all sizes by:

- **Driving revenue growth** through new products and **improved customer experience**.
- **Increasing competitiveness** vs. foreign companies benefiting from significant investments in AI.
- **Improving productivity** through cost reduction, automation, and improved efficiency.
- **Opening new markets** through innovation, **seizing new opportunities** that would be beyond reach without data and AI technologies.

This dual potential for top-line and bottom-line impact makes it essential for SMEs to consider AI as a key element of their overall business strategy.

A recent report (1) **from METI in France showed that in 2021, only 9% of French SMEs had adopted AI solutions, with 13% intending to do so in the near future. These SMEs need to establish a clear strategic vision before incorporating AI into their operations.**

The strategic vision will provide direction, purpose, and alignment for all stakeholders. It will allow for improved internal alignment, and will help to foster a culture of collaboration and teamwork, ensuring that all stakeholders, including management, employees, and customers agree with the organization's AI initiatives. It will enhance decision-making, allowing organizations to identify the most promising areas for AI investment, as well as the most appropriate technologies. Finally, a strategic vision will increase the effectiveness of AI implementation. By providing a clear roadmap for organizations to follow, it will help to ensure that all initiatives are properly planned, executed, and measured. Without a well-defined vision, SMEs risk deploying AI solutions that fail to align with their goals and values, leading to wasted resources and missed opportunities.

Adopting AI requires assessing the company's data maturity, developing and implementing a data governance plan, and preparing for data sourcing and management strategies.

> *Often, many leaders hire the team first and then try to figure out what they'll do. At Janssen, we prioritize 2 factors - business importance and data science feasibility. Focusing on the prioritized list of initiatives helps deliver impact, secure buy-in, and build the organizational momentum to do more.*

Najat Khan- Chief Data Science Officer, Janssen

---

1. *"INtelligence Artificielle : L'expédition du dirigeant d'ETI", METI / Kea & Partners / Apax Partners, 2022*

Common pitfalls

Some of the most common pitfalls SMEs may encounter when setting a strategic vision for AI adoption include:

- **Lack of clarity**: Without a clear and specific vision, organizations may struggle to align their AI initiatives with their overall goals and values. This can lead to confusion and lack of direction, making it difficult to prioritize initiatives and measure success.
- **Lack of stakeholder buy-in**: A lack of buy-in from key stakeholders, such as management, employees, and customers, can make it difficult to implement and maintain an AI strategy. Ensuring that all stakeholders are engaged and invested is essential for success.
- **Lack of acculturation to, or understanding of, the transformative potential of AI**.
- **Poor sponsorship at the highest management level**: AI has the potential to drive the business into new directions, so the effort must be well understood at the highest level, or managers will tend to minimize the risk to their performance (and thus their career) by minimizing operational impacts. Some relevant education programs are listed in Chapter 6.
- **Lack of urgency to apply AI to specific business cases**.
- **Lack of prioritization of AI use-cases** to educate and inspire all participants: "show vs. tell".
- **Poor understanding of available resources**: Setting a vision for AI is only the first step, implementing it can be challenging without the necessary resources. Executives often make the mistake of assuming that "AI startups" can handle all of their AI needs. However, this is misguided because startups typically focus on standardized business cases in order to grow their user base and limit changes to their product. On the other hand, most businesses require AI solutions that are tailored to their specific needs and may deviate from the standard features offered by startups' products or platforms. Therefore, it is important for executives to recognize the limitations of AI startups and seek out customized solutions that meet their unique requirements.

- **Focusing too narrowly**: Organizations that focus too narrowly on a specific AI application or technology may miss out on other opportunities. A strategic vision should be broad enough to allow for flexibility and adaptability, allowing organizations to capitalize on new opportunities as they arise.
- **Lack of governance**: Implementing AI systems requires governance and control, to ensure the systems align with the company's values and ethical principles, as well as to ensure compliance with regulations.
- **Not involving experts**: Setting a strategic vision for AI requires a deep understanding of the technology and its potential applications. Organizations should involve experts in the field to ensure that their vision is realistic and achievable.
- **Failure to iterate**: A strategic vision for AI is not a one-time event, organizations should continuously evaluate their vision and make adjustments as needed. Failure to iterate can lead to the vision becoming obsolete and the company falling behind the competition.
- **Limited budget**: Many SMEs have a limited budget, which can make it difficult to set up large projects and accurately estimate their return on investment. This is particularly true for internal development, which requires not only the appropriate management culture and resources, but also the necessary support from middle management.
- **Over-siloed organizations**: Various departments may not understand the full scope of challenges and opportunities presented by AI. This can lead to limited collaboration and coordination, making it difficult to fully realize the potential of AI. In larger groups, the allocation of AI resources is often focused on structuring projects identified by management and closely tied to the core business of the company. This can lead to overseeing the major impact AI can have on other functions such as quality or logistics.

Road to success

SMEs looking to implement AI systems have a variety of solutions available to them. In this section, we provide possible methods and resources to help navigate the development of a relevant AI adoption strategy.

This "road to success" for AI begins with a realistic, documented assessment of the company's current status, and a clear view of the end-goal, including prioritized AI projects on a specific timeline.

As a first step, SMEs should perform **a full assessment and scoping exercise**:

- Step 1: Pre-qualification interviews with the company's sponsors:
  - Assess the relevance of conducting the data and AI maturity assessment .
  - Explain method and expectations.
  - Sign a confidentiality agreement.
- Step 2: Align stakeholders: ensure leaders are on board with the objectives, cost, and potential impacts.
- Step 3: Establish the company's AI and data maturity score using the framework presented in Chapter 3.

- Step 4: Identify projects that can potentially use AI and data technologies using the "Core vs. Peripheral AI" framework presented in Appendix could prove helpful.
- Step 5: Evaluate the potential impacts of the projects on the existing systems, processes, and organizations.
- Step 6: Prioritize projects, define a roadmap.
- Step 7: Create a diagnostic report with implementation recommendations and an action plan.

The ROI of AI

AI projects can be challenging to execute successfully for a variety of reasons, and predicting or managing the return on investment (ROI) can be difficult. Some important factors include the fact that AI is a new technology with nascent use-cases, and the iterative nature of the process means that both ROI and timelines are uncertain. Also, AI transformation implies a level of AI maturity that may not exist in all organizations. However, it is still possible to measure AI ROI.

It is important to go beyond the usual financial ROI and consider the value brought by AI in multiple other areas. AI ROI could for example be analyzed along three categories of returns, as suggested by market research firm Emerj (1) :

- Measurable factors: When evaluating the ROI of an AI project, it is recommended to consider not only **financial results** but also other measurable factors such as **increased efficiencies**, **reduced risk**, and **non-financial measurements like customer satisfaction scores**.
- Strategic ROI is also an important consideration, as it pertains to the **impact of AI on an organization's long-term goals** and digital transformation vision.
- Capability ROI is another important factor that refers to the impact of an AI project on an organization's overall level of AI maturity, including the **skills, resources, and culture** required for successful AI deployment. To build this maturity, organizations must be willing to invest in training, infrastructure, and cultural changes. Educating stakeholders on the importance of building AI maturity is essential for the long-term success of AI projects.

---

1. "The Three Kinds of AI ROI – Emerj's Trinity Model" (https://emerj.com/ai-executive-guides/ai-roi-emerj-trinity/)

At every step of planning an AI project, SMEs can use the "Make or Buy" framework listed in Appendix. This tool is designed to navigate the variety of options available for the implementation of AI:

- Hire an AI team
- Trust existing IT resources with AI developments, under certain conditions
- Use start-ups to build solutions to support function rather than address core business issues
- Use consulting and services companies to drive the process of implementing AI systems by providing guidance and support
- Develop partnerships between companies and universities to gain access to the latest research and expertise in AI
- Establish industry-based partnerships focused on developing shared resources ;
- Identify available AI providers of the relevant type (platform, product startup, consulting firm, applied AI lab)
- Utilize exchange platforms and cross-industry ecosystems connecting SMEs with others in their industry and providing opportunities for collaboration and knowledge sharing
- Establish communities of practice to share information and best practices
- Utilize competitiveness clusters ("Pôles de compétitivité") and industry associations such as Hub France IA that bring together start-ups and large corporations
- Create and share pragmatic "success stories" illustrating the use of AI in solving basic problems in different industries. This will help to communicate the value of AI to middle management and create buy-in for the implementation of AI systems.

SMEs from various sectors have successfully implemented AI

- A **manufacturing** company in the automotive industry is using AI to improve the efficiency of their production lines. The company implemented ML models to predict when equipment would need maintenance, reducing downtime and increasing productivity.
- A **retail** company uses AI to personalize the customer experience on their e-commerce website. By using ML models to analyze customer data, the company is able to generate personalized product recommendations and targeted marketing campaigns.
- A **logistics** company is using AI to optimize delivery routes. By analyzing data from GPS-enabled vehicles, they are able to identify patterns and predict traffic congestion, allowing them to create more efficient routes and reduce delivery times.
- An **healthcare** company uses AI to improve patient outcomes. They implemented ML models to analyze patient data and predict which patients were at high risk of developing certain conditions.
- A **financial services** company is using AI to improve the efficiency of their risk management process. They implemented ML models to analyze financial data and identify patterns that could indicate fraudulent activity.
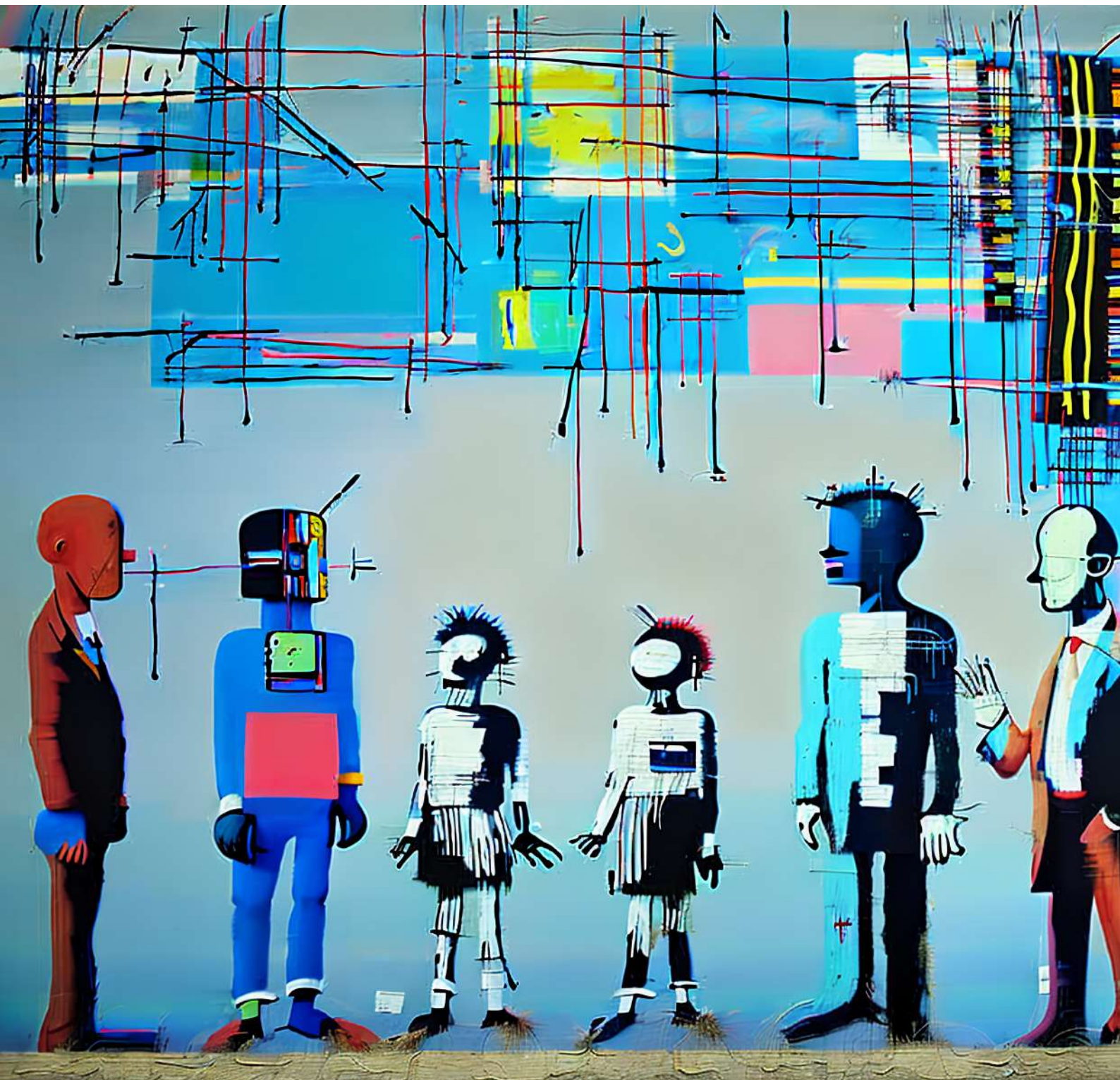
CHAPTER 5

THE HUMAN FACTOR
(AI AS A COLLECTIVE TRANSFORMATIVE PURPOSE)

TAKEAWAYS

- **Increasing collective understanding of AI is essential for successful implementation and leveraging AI for the best possible outcome.**

- **Top management must go through an acculturation on a variety of AI related topics.**

- **Executive leadership must consider the impact of AI on jobs and on operating models, and create a vision and agenda for AI to guide the entire organization.**

- **Establish a culture that encourages collaboration, invention, and testing.**

- **Avoid negative sentiment about AI, and emphasize its potential benefits.**

- **Leverage AI opportunities, while providing training and support for affected employees.**

What is at stake?

SMEs should not view Artificial Intelligence (AI) only from a technical and operational perspective, but rather as **a tool that can have a wide-reaching effect on the organization, its strategies, operations, and stakeholders**. Taking this approach can lead to a clear vision for AI adoption, as well as encourage a culture of collaboration and invention.

Among the most fundamental challenges facing companies considering deploying AI is insufficient knowledge, interest, or understanding among stakeholders. When executives do not fully understand the capabilities and limitations of AI, they can set unrealistic expectations and make decisions that may not optimally benefit the organization. If employees do not understand the strategic importance of data or the potential of AI for innovation, they may refrain from supporting the company's AI efforts. **Collective understanding of AI across all employees of a company is the foundation for successful AI implementation, and is key to leveraging AI for the best possible outcomes**. SMEs must therefore educate their top management on a variety of AI related topics. This might seem daunting, but will enable the rapid transformation thanks to short-loop processes driven by top management.

Company culture can have a huge effect on the democratization of AI within an organization. A hierarchical structure, where decisions are made by a small group of executives (in contrast to a participatory, decentralized approach) can negatively impact the level of employee engagement and empowerment when it comes to AI. Since AI must be used across teams and needs to be developed using an iterative approach, companies with cultures that lack transversal collaboration or fail to embrace testing and learning may struggle to make use of AI effectively. SMEs, which are more agile and able to shift more quickly than large corporations, may have an advantage.

The deployment of AI can sometimes lead to feelings of delegitimization among employees, as automated processes enabled by AI can reduce or even eliminate the need for certain positions, making employees whose jobs were automated feel devalued. AI is also used to make decisions in areas like finance, HR and logistics, leading to increased reliance on algorithms and data, rather than employee experience and expertise. This can leave employees feeling their decisions are no longer valued. Additionally, the introduction of AI can also bring to light disparities in skills between employees, with some more easily able to work with and use AI, while others struggle. This can lead to discrepancies in performance, creating a sense of dissatisfaction among those who find themselves at a disadvantage. AI can also present opportunities, such as improving skills or allowing employees to focus on tasks that require more complexity and creativity.

Businesses must also consider the impact of AI on their operating model. When AI is utilized across a company, local users are granted more power and responsibility. In practice, this has led to IT departments taking on a support role, as opposed to a leadership role as was the case for the introduction of other technologies such as the web. MLOps (see Chapter 11 - Glossary), for instance, challenge organizations to re-evaluate and refine their operating models, as it should be managed on the business side, not on the technical side.

While management must have an unambiguous vision and agenda for AI, all stakeholders must pursue collaboration and clear communication to avoid conflicting objectives on the direction and goals of AI projects.

As the development and adoption of AI technologies rapidly progresses, companies must consider their capacity to attract talents in order to stay competitive. This issue is not limited to individual firms; the transition to AI has potentially far-reaching implications across labor markets. A recent OECD study by Krämer, C. and S. Cazes (2022) suggests that social dialogue can be an effective tool for addressing these changes, fostering inclusive labor markets, and providing support for both workers and businesses. However, this dialogue must take into account the new pressures and practical challenges that come with the adoption of AI, such as the lack of AI-related expertise and resources.

To avoid these outcomes and maximize the potential of AI, companies must ensure that their employees, customers, and other stakeholders understand the benefits associated with the technology, such as increased growth, attractiveness, and competitiveness. And when people are involved in the design and implementation of AI systems, they can help to ensure systems are being used ethically and in a way that benefits all stakeholders.

Common pitfalls

Organizations that fail to capitalize on the potential of AI risk missing out on lucrative opportunities and could face a range of negative outcomes:

- **Incomplete unlocking of value**: When AI programs only generate a small share of the anticipated value, the cost-benefit analysis of any company transitioning to AI is diminished.
- **Widespread rejection of AI**: Misunderstandings and misconceptions about the technology lead to a disadvantage that could cost a company market share and reputational damage.
- **Complete disregard of AI**: Negative sentiment about the technology leads to any solution involving AI being perceived as too risky.
- **Lack of transparency and understanding**: Many stakeholders have preconceived, sometimes negative opinions about AI. Businesses must be transparent about how it is being used and communicate its benefits and limitations broadly. Lack of transparency and misunderstanding can create friction and pushback from stakeholders. Beyond acculturation, understanding and trusting AI depends on clear, accurate information.
- **Misconception of AI's impact on jobs**: As AI becomes increasingly prevalent in the workplace, organizations must consider the possible impact on jobs and provide training and support for employees who may be affected. AI can analyze large amounts of data quickly and accurately, while humans bring creativity, critical thinking, and empathy to decision-making. By combining the strengths of AI and humans, organizations can achieve better results and make more informed decisions: this is what is called "augmented intelligence".
- **Fear of failure**: Some executives can be hesitant to take risks with new technology. This reticence must be addressed in order to create a culture that allows for failure, always a possibility on the road to widespread adoption of AI. It is often said that taking a chance at AI is less risky than not using AI at all, especially when competitors have already adopted it and are reaping the rewards. To further encourage AI adoption, companies should be encouraged to share their stories of failure with the wider community, to accelerate learning and adoption of AI.
- **The belief that AI is a complex and technical subject**, beyond the reach of most businesses, particularly SMEs: This can lead companies to avoid exploring AI, despite the resources available to SMEs at all stages of the AI adoption path. These tools allow companies to build, train, and deploy AI models without in-house technical expertise, making it possible for a wide range of businesses to take advantage of AI.

Road to success

Businesses looking to adopt AI need more than just technology. They must also adopt a holistic approach involving all stakeholders: employees, customers, and collaborators. Companies should focus on building consensus around AI through training, communication and collaboration, as well as more generally creating a supportive and inclusive environment.

When considering the adoption of AI, businesses must contemplate training needs, organizational flexibility, and whether the company has created a culture that values innovation. Managers need to be aware of the potential consequences, such as job displacement, changes to decision-making processes, and employee disempowerment.
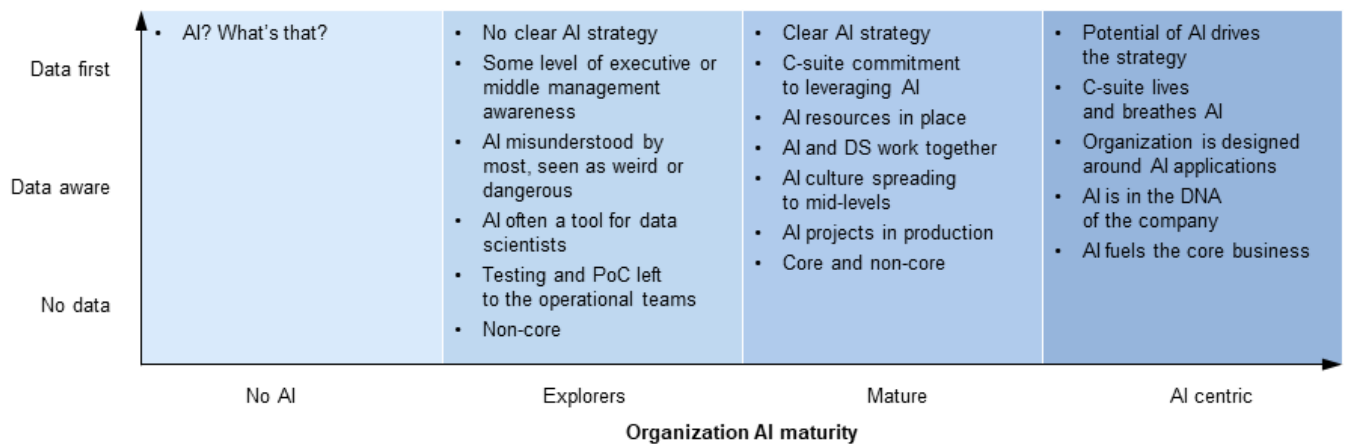
Fortunately, there are **no-code** and **low-code tools** available that can help non-technical people develop AI solutions, such as Vertex AI by Google Cloud and Amazon SageMaker by AWS. In the near future, tools such as ChatGPT may even allow basic programming from text prompts, eliminating the need for employees to learn and use programming environments. **But businesses should not rely solely on these tools, and should instead seek out support for more technical processes that require a higher level of expertise**.

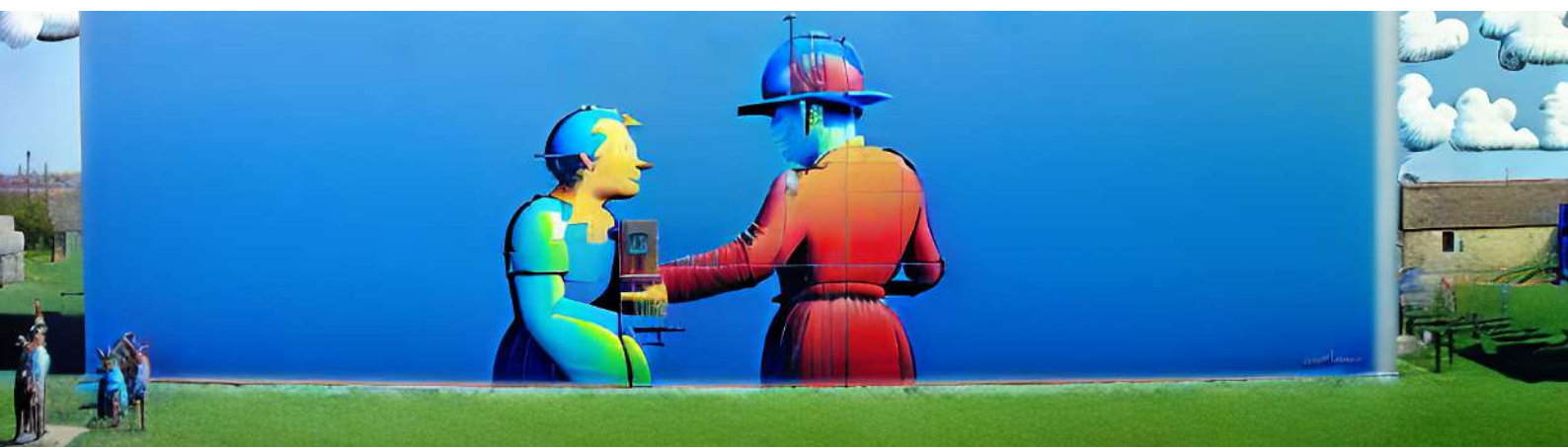To increase success in AI adoption:

- SMEs might consider creating **communities of interest** that bring together different companies, across a given region or industry sector. This can provide a platform for sharing ideas and best practices, and possibly creating data commons for the benefit of all involved.
- Within an organization, employees can exchange ideas and best practices **in groups and forums** dedicated to issues around the adoption of AI. In larger SMEs, transverse AI communities have been proven to help AI adoption by promoting a sense of collaboration and support, and encouraging the sharing of knowledge and expertise. It remains important to secure executive level support, including the necessary budget and resources, to ensure that AI initiatives succeed.
- SMEs can adopt a **hybrid approach**, combining **internal and external resources**, including resources shared with other SMEs. A hybrid model may be particularly effective for SMEs with limited resources as budget and staffing constraints may limit the adoption of AI projects. The cost and effort of adopting AI can be shared across multiple organizations, while still allowing each organization to focus on its core business needs.

A useful tool to understand the human aspects of the readiness of an organization towards AI adoption is to position it on the AI maturity journey introduced in Chapter 4:

Organization data awareness

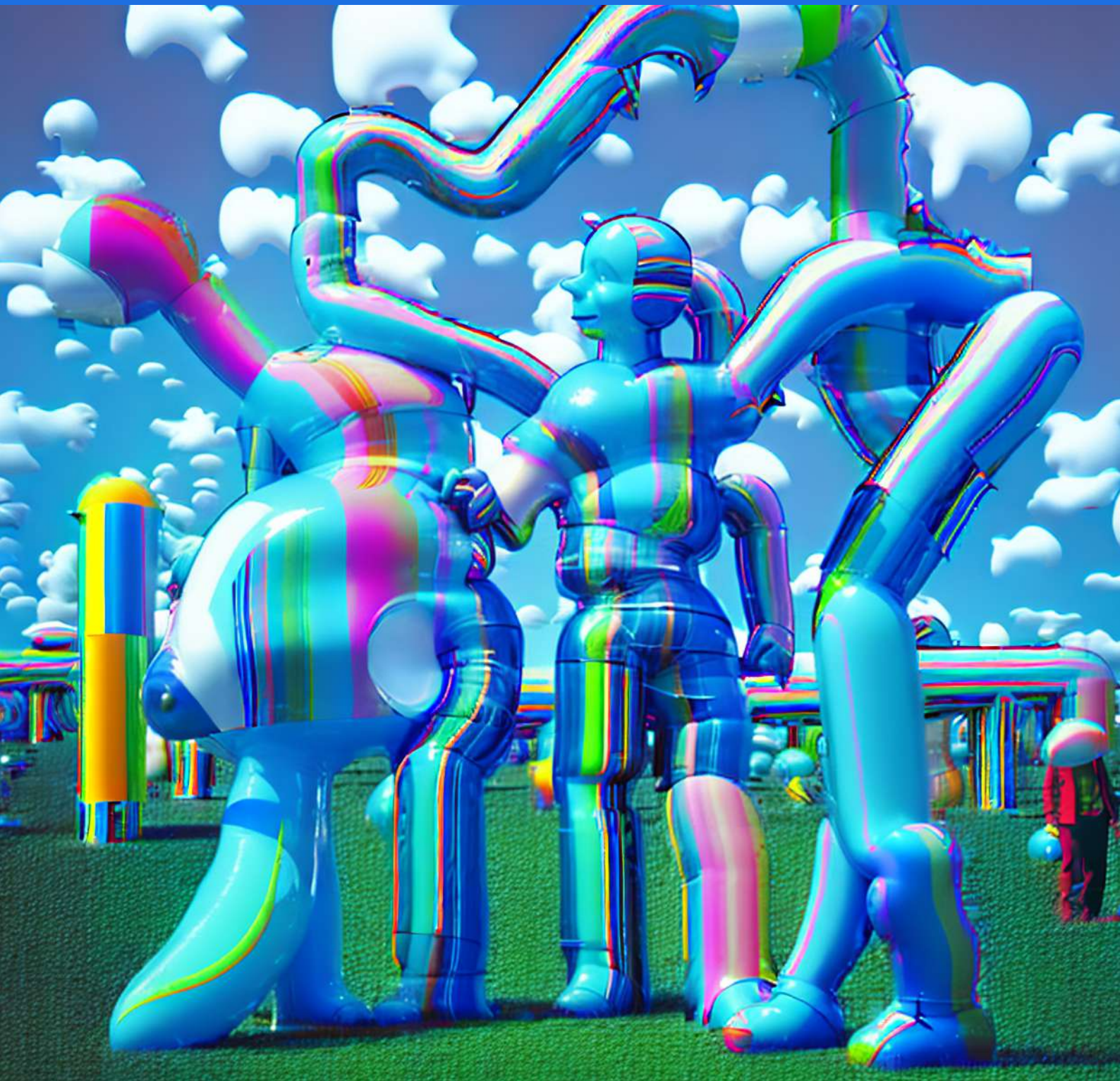|	No AI	Explorers	Mature	AI centric
Data first	• AI? What's that?	• No clear AI strategy • Some level of executive or middle management awareness	• Clear AI strategy • C-suite commitment to leveraging AI	• Potential of AI drives the strategy
Data aware		• AI misunderstood by most, seen as weird or dangerous • AI often a tool for data scientists	• AI resources in place • AI and DS work together • AI culture spreading to mid-levels	• C-suite lives and breathes AI • Organization is designed around AI applications
No data		• Testing and PoC left to the operational teams • Non-core	• AI projects in production • Core and non-core	• AI is in the DNA of the company • AI fuels the core business

Organization AI maturity
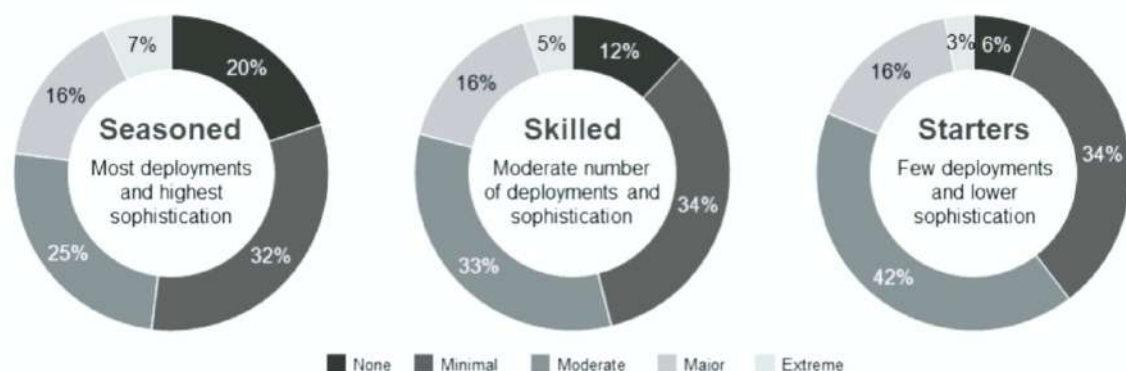
CHAPTER 6

SKILLS MATTER

TAKEAWAYS

- **Companies looking to adopt AI systems face a critical shortage of talents.**

- **For low AI maturity companies, the skills shortage for AI projects is estimated at 60%.**

- **Training high-skilled professionals takes a long time, but formal academic education is not the only way to acquire AI skills.**

- **Upskilling and reskilling existing employees can provide companies with immediate leverage to begin the process of AI adoption.**

- **Companies should diversify their recruiting funnel, work on staff upskilling and retention, train managers on data literacy, and consider deploying low code/no code tools for simple AI use cases.**

- **Commonly demanded AI profiles include Data Analyst, Data Engineer, Data Scientist, Chief Analytics Officer, Data Architect, Chief Data Officer, Data Protection Officer, ML Engineer, Database Administrator, and Data Steward.**

What is at stake ?

Companies looking to adopt Artificial Intelligence (AI) systems face a **shortage of Data Scientists, Data Analysts and Data Engineers**. The number of graduates and professionals in this field is insufficient to meet the needs of businesses.

According to a recent Deloitte Insights survey, low AI maturity companies estimate the skills gap for AI projects at 60%, considerably hampering AI adoption.

How would you rate your company's current skill gap in meeting the needs of your AI projects?

Seasoned — Most deployments and highest sophistication: 7%, 20%, 16%, 25%, 32%

Skilled — Moderate number of deployments and sophistication: 5%, 12%, 16%, 33%, 34%

Starters — Few deployments and lower sophistication: 3%, 6%, 16%, 34%, 42%

Legend: None, Minimal, Moderate, Major, Extreme

Note: We grouped organizations into three segments, based on the number of AI production deployments undertaken and the level of AI expertise across multiple areas
Source: "State of AI in the Enterprise", 3rd Edition © Deloitte Insights

Training high-skilled professionals takes a long time: PhD programs, for example, last 5-7 years. However, formal academic education is not the only way to acquire AI skills. It is now possible – and even desirable, given the rapid pace of technological change – to develop or continue to hone such skills after entering the labor market.

To ease this, the French government recently allocated €781 million to the "massification" of AI education. As part of this push, Hub France IA has been asked to create a nationwide AI competency framework to be used by businesses, education and job market stakeholders. This effort seeks to document the shortage of skills in AI in the short term, and help increase the number of AI specialists in the long term. This report is scheduled to be released mid-2023.

The skills shortage leaves only two alternatives to SMEs: **outsource** or **upskill and reskill** part of their workforce, one of the most significant challenges at hand for them, as it is costly and needs to happen ahead of AI projects delivering measurable ROI.

AI talent in Europe: a snapshot from a recent LinkedIn Economic Graph survey

- The US employs twice as many AI-skilled individuals as the EU, despite having only half the size of the labor force.
- The EU can double the size of its current AI workforce by training and upskilling "near-AI" talent - workers with adequate foundational skills that can be developed with the appropriate training and support.
- AI expertise is concentrated among a small group of industries, typically the leading industry champions within countries. For example, Germany excels in AI for manufacturing and the automotive industry, while finance in the UK, manufacturing in Italy and Sweden, healthcare in Belgium, and telecommunications in Finland follow similar patterns.

- Established and large companies are more likely to be the earliest adopters of AI technology in the EU, while startups and digital natives dominate in the US.
- Like other technology fields, AI is predominantly male-dominated in Europe, with women accounting for only 16% of AI workers, posing significant risks to productivity and bias.
- AI workers tend to be highly specialized. Three AI libraries are required by more than 13% of the jobs advertised on LinkedIn, yet they have a supply below 3%.
- Two-thirds of AI-skilled individuals work in the technology (ICT) sector or within academia.

Common pitfalls

Common pitfalls related to skills and training that companies should be aware of:

- Many companies lack the **necessary in-house expertise** to develop and implement AI and ML solutions. This creates a skills gap that must be filled through hiring or employee training programs.

- Even when companies hire AI and ML experts, their **existing employees may lack the necessary skills to work with these technologies**. Companies need to invest in employee training programs to ensure that their workforce is capable of operating and maintaining AI and ML applications.

- Companies must **identify which employees need to be reskilled or upskilled** to work with AI and ML technologies. This requires a thorough assessment of existing employee skills and capabilities, as well as an understanding of the specific AI and ML applications that will be deployed.

- When upskilling and reskilling, companies must be careful that upskilling and reskilling **does not create a rift between AI-trained employees and their colleagues**, which could be damageable to company culture and employee morale (the feeling of some being "left behind" by some).

- Employee training programs require time, resources, and support. Companies need to provide **adequate training opportunities and support** to ensure that their employees can acquire new skills and apply them effectively in their work.

- AI and ML initiatives **must align with business goals to be successful**, and employee training programs are no exception. Companies need to identify which skills are most critical to their AI and ML initiatives and design training programs that address those needs.

Road to success

In this context, *how can businesses attract and retain high-level candidates? How to ensure that working conditions, compensation, and career perspectives are attractive enough to retain key expertise? What can be done to adapt the workforce in place to the new demands brought by AI? What other resources are available to help SMEs get on the path of AI adoption?*

Cross-disciplinary education and training is key for SMEs to ensure their staff possess the technical skills and an understanding of business stakes and challenges. Businesses must hire technical staff with an understanding of key issues facing the business. At the same time, existing staff who are well versed in the stakes and challenges facing the company can gain a basic understanding of AI, potentially benefiting the organization. Comprehensive and consistent approaches have demonstrated success in this area, such as at ESSEC Business School, where future managers are trained with fundamental AI knowledge to facilitate cooperation with technical teams.

In addition:

- **Upskilling existing staff** can provide companies with immediate leverage to begin the process of AI adoption.
- **Enhancing the "employee experience" and company value proposition** can have a positive impact on attracting and retaining talent.
- Consider deploying **low code / no code tools** for simple AI use cases.
- As noted in Section 8, companies wishing to adopt AI systems but lacking the necessary resources can benefit from **partnerships with applied research laboratories**, such as 3IA cross-disciplinary Institutes and IRT SystemX.
- **Promoting partnerships with universities, engineering schools, or business schools** through the CIFRE doctoral thesis system can give businesses access to high-skilled candidates who may be reluctant to join the corporate world, leveraging the existing "Loi Pacte" policy, which enables researchers to keep their academic status while working for businesses.

How to attract talented AI candidates?

Diversify the recruiting funnel

Most French Universities now offer specific AI programs. Here is a glimpse of some high-quality programs:

- Hybridization of Engineering Schools and Business Schools:
  - Centralesupélec + ESSEC (Bachelor Intelligence Artificielle, Sciences des Données et du Management / Master in Data Science and Business Analytics)
- Engineering schools that have developed data & AI bachelor and master programs:
  - Polytechnique, ENSAI, EPITA
  - Mines ParisTech PSL
  - IMT Nord Europe
- Most French Universities now offer specific AI programs, mostly technical
  - Data pure players companies offer shorter trainings (under a year) which fuel the market with much needed entry level AI candidates: Le Wagon, DataScientest, Simplon
  - Several AI pure play education operations have recently emerged, and will provide the market with a midrange profiles in the coming several years: Alvancity, Data ScienceTech Institute, EPITA IA Institute

Work on staff upskilling and retention

- Train managers to data literacy
  - Examples of training programs for AI non specialists:
    - ESSEC Business School: "Leading with Data-Driven Strategies and Analytics" (online)
    - Mines ParisTech: "IA, De l'idéation à l'industrialisation"
    - Telecom Paris: "Intelligence Artificielle et Science des données pour les managers"
- Train staff to AI: training offered by consulting firms can be tailored to specific requirements, and are well suited to board upskilling / reskilling plans

Hire foreign talents

France offers fast-track visas to highly-skilled professionals, including for AI specialists, designed to boost France's economic attractiveness. The "Passeport Talent" is a multi-year residence permit for non-EU nationals to live and work in France. It applies to high-skilled profiles that will contribute to France's economic attractiveness.

What are the most demanded AI profiles?

According to an April 2022 Kantar/ Essec Business School survey (1), the most demanded AI-related profiles are:

- **Data Analyst**: The Data Analyst is the most sought-after profile. Their technical skills are versatile and extend to programming languages. They also have a strong interest in statistics and can. They evaluate and integrate data, develop dashboards, and automate queries.

1.*"Le futur des métiers de la data vu par les grands groupes français, 2ème édition de l'enquête réalisée pour la French Tech Corporate Community", Kantar Public / ESSEC Metalab / BPCE, December 2022*

- **Data Engineer**: The Data Engineer defines technical solutions within an framework defined by the Data Architect. They maintain the technical applications, integrate the data and ensure its quality, and monitor data flows.
- **Data Scientist**: This term is widely used, although it covers a wide array of skills and competencies. An expert in mathematics, the Data Scientist automates processing through intelligent algorithms.
- **Chief Analytics Officer**: The Chief Analytics Officer is the chief data scientist.
- **Data Architect**: The Data Architect organizes the management of the raw data, optimizing data flows. They design the data architecture, ensure the consistency of the data model, maintain the data dictionary, and can suggest modeling changes when necessary.
- **Chief Data Officer**: The Chief Data Officer (CDO) leads data strategy. The CDO chooses AI solutions in synergy with the CIO and the business units to establish a favorable and secure framework for a company's top decision-makers.
- **Data Protection Officer**: The Data Protection Officer (DPO) is responsible for data governance, and in particular for data protection. The DPO is the person responsible for the implementation of the General Data Protection Regulation (GDPR).
- **Machine Learning (ML) Engineer**: The ML Engineer implements learning models. Their skills are multiple, ranging from statistics and mathematics to modeling and security.
- **Database Administrator**: The Database Administrator (DBA) manages a database, from its implementation to its operation, including configuration and optimisation. While this is not a new role, in the AI adoption process, a DBA must be able to adapt to new architectures.
- **Data Steward**: A pillar of data governance, the Data Steward is the guarantor of the data life cycle. Strong interpersonal skills enable the Data Steward to manage all steps of the data lifecycle in their business area of responsibility.

OECD: The impact of AI in the workplace

A forthcoming global study by OECD (1) about the impact of AI on the workplace in companies from the finance and manufacturing sectors reveals interesting preliminary insights:

- Task change: AI is often implemented in ways that complements human labour.
- Job quantity: In the majority of the companies surveyed, the implementation of AI technologies has no reported impact on the job quantities of workers most affected.
- Skill needs: In a significant portion of companies surveyed, AI technologies prompt changes in skill needs. Most often, workers require new skills, or greater emphasis is placed upon existing skills, requiring retraining.
- Job quality: In manufacturing companies, AI technologies lead to improved physical safety, reduced physical strain, and improved mental well-being on account of reduced culpability for mistakes. In finance, improvements due to more manageable workloads and fewer repetitive tasks are identified. However, in both sectors, there are reports that AI technologies accelerate the pace of work, significantly increasing pressure and stress.

1.Milanez, Anna (2023), *The Impact of AI on the Workplace: Evidence from OECD Case Studies of AI Implementation, OECD Social, Employment and Migration Working Papers, No. 287, OECD Publishing, Paris.*

CHAPTER 7

DATA IS GOLD

TAKEAWAYS

- **Data is a valuable asset and the quality and quantity of the data used to train AI systems directly impacts the outcomes.**

- **It is essential for SMEs to learn to manage the data lifecycle: data acquisition, preparation, exploration, utilization, storage, and maintenance.**

- **Data capital is a newly recognized form of capital created by leveraging data to gain real monetary value.**

- **Data and AI governance require organizational change and new roles within companies.**

- **Poor data quality can lead to incorrect or undesirable outcomes.**

- **Data commons enable data sharing across organizations.**

What is at stake?

The deployment of Artificial Intelligence (AI) in the global economy is a multifaceted challenge that requires not only mainstream adoption and significant capital investment, but also a focus on data. Until recently, data has not been recognized as a valuable asset, particularly in the world of SMEs. **However, data is the lifeblood of current AI systems and its quality and quantity directly impact the outcomes and insights that these systems produce. In other words, poor data leads to poor results, which translates to poor business decisions.**

The data lifecycle

In the context of AI and more specifically ML, the data lifecycle is most often considered to include the stages of data acquisition, preparation, exploration, utilization, storage, and maintenance:

- **Data acquisition**: sourcing reliable and good quality data sources.
- **Preparation**: cleaning, validating, and transforming data to ensure usability and avoid bias and performance issues.
- **Exploration**: analyzing and interpreting data to uncover patterns.
- **Utilization**: creating algorithms to learn patterns and make predictions.
- **Storing**: securely storing and making data accessible. Cloud storage is often the best option.
- **Maintenance**: updating and preserving data quality, ensuring data freshness and regulatory compliance.

The data lifecycle guides businesses in setting up processes and procedures that maximize data quality and thus the accuracy of ML models. Following these steps also enables compliance with data governance policies and regulations.

Data as capital

Data capital is a newly recognized form of capital created by leveraging data to gain real monetary value. It is distinct from traditional financial and human capital and can interact with them in either a cooperative or competitive way. **Data capital has the potential to drastically shift competition and disrupt income distribution between regions and populations, making it a strategic asset for businesses and economies alike**. As such, it is crucial for regional sovereignty.

Poor data, poor AI

Data collection and preparation are essential for successful ML models as the quality of the data used to train the algorithms will directly affect their performance. **As much as 60% of the time spent on Machine Learning (ML) can be dedicated to data cleaning and prepping, because poor, irrelevant or faulty data results in inaccurate predictions.** High-quality data not only improves the performance of the model but also enables it to tackle more complex tasks.

It is essential to run **data quality checks** to ensure the success of a project and prevent potential issues. Although this may take time, neglecting data quality can have a detrimental effect on businesses, leading to inaccurate data input, data quality issues during training, lack of data standardization, outdated data, negative impacts on customer communication, flawed analysis and lost revenue, and potentially even regulatory violations.

To avoid these issues and remain competitive, businesses should prioritize data quality and implement effective data management processes. Doing so proactively is key, as retroactively rectifying mistakes is much more difficult and time consuming.

Data governance, key to AI at scale

Data governance has been around since data has been collected. With the emergence of Enterprise AI, organizations have been encouraged to embrace it. Regulations on data privacy, such as GDPR and CCPA as well as a focus on AI risks and model maintenance have necessitated **a fundamental change in data governance to balance data use with risk protection**.

Scaling AI demands enhanced data governance, requiring **organizational change and new roles within companies**. This is key as **good data governance is a pillar of responsible and sustainable Enterprise AI**. Data governance reduces risk by helping organizations comply with regulations and avoid security issues, which saves money and increases confidence in data, analysis and models. Not only that, companies following clear data governance policies can avoid the risk of negative press and fines associated with data breaches or misuse.

Data governance involves establishing policies, roles, standards, and metrics to manage data effectively. Relying solely on existing data governance strategies can lead to inconsistent data management and a lack of trust in the data. **Without trust in data, confident and accurate decisions cannot be made**.

Traditionally, IT organizations have been responsible for data governance; however, as businesses move towards AI, and as data becomes ubiquitous, AI and **data governance requires a different skill set, including expertise in areas such as data architecture, privacy, integration, and modeling**.

Proper data governance should be managed by business experts who know :

- The definition of the data
- The source(s) of the data
- The value of the data to their business
- The planned use of the data in the context of their business
- Appropriate data use guidelines

In addition, new components of AI and data governance should be added under the data governance umbrella: ModelOps (ML model management) and Responsible AI governance.

ModelOps should define **clear policies, roles, standards, and metrics for data governance** addressing the following questions :

- *Who is responsible for the effective performance of production ML models?*
- *How are models updated and how is model drift addressed?*
- *How are models described and how are versions tracked?*
- *What metrics are used to measure the performance of models and what performance level is acceptable to the business?*
- *How is robustness characterized?*
- *How is the performance of models monitored?*
- *How are models audited and made comprehensible to non-developers?*

Governance strategies should emphasize data oversight and pursue a model of Responsible AI. Despite the popular notion that data science is a neutral and objective field, AI algorithms are not an infallible representation of reality. The dangerous illusion of objectivity can lead to a sense of complacency, undermining the sense of responsibility among teams and individuals involved in AI projects.

AI is not neutral. We must recognize that **it is subject to misuse and that policies and oversight are necessary to ensure its responsible use**. Key considerations for responsible AI governance include:

- Ensuring the data used to train models does not have inherent biases.
- Excluding characteristics such as ethnicity, gender, age, and religion from the training process.
- Detecting and mitigating biases that could harm certain groups and ensuring fairness.
- Ensuring the privacy of personal data.
- Identifying data that has become stale after its intended use has been fulfilled.
- Confirming data collection and storage practices are in line with regulatory standards and company standards.

Common pitfalls

Data and AI governance can be a complex undertaking that requires coordination, discipline, and organizational change. The larger the enterprise, the more challenging the implementation of these programs becomes. Effective AI governance requires a transformation of people and technology.
Despite the clear benefits of an effective AI governance program, organizations may face challenges that hinder their efforts.

Insufficient executive sponsorship

Without **unambiguous and explicit support from senior leaders**, a data and AI governance program is likely to lack enforcement mechanisms. Data scientists, analysts, and business professionals may be inclined to ignore data governance policies and continue existing practices. Top-down oversight is essential in case of non-compliance with data governance policies and the recognition of actions taken towards enhancing data governance.

A culture of ownership is essential for effective data governance

A data governance strategy can only be effective if there is **a culture of ownership and commitment to improving the use and exploitation of data throughout the organization**. This depends on senior sponsorship, communication, education and training, and the use of appropriate tools. As the adage states, "Culture eats strategy for breakfast."

Poor communication

Clear and widespread communication is essential for the success of a data governance program. Without it, employees may be unaware of the policies, standards, roles, and metrics that make up the program. **Education and awareness about data governance policies and standards** is crucial for employees to implement them effectively.

Lack of training and education resources

The acquisition and development of training and education is a critical aspect of data and AI governance, familiarizing individuals with relevant policies and providing practical examples of the importance of data governance.

Lack of relevant data

AI models may not perform accurately with insufficient data, leading to incorrect predictions, poor decision-making, and overall inefficiency. Moreover, AI models that are trained on biased or inadequate data can perpetuate and amplify existing inequalities, discrimination, and harm. To mitigate the consequences of lacking relevant data, **organizations can implement various strategies such as data collection, data sharing, or data augmentation**.

Data commons, where organizations collectively pool data resources, can be a potential solution to the problem of insufficient data. However, businesses often resist cooperating with data commons due to data privacy and security concerns, competition, and the fear of losing control over their data assets. There may also be a lack of trust in data commons governance and ownership structures.

In addition, the cost of developing data sources can be significant, depending on factors such as the amount and complexity of data, the methods used for data collection, storage, and processing, and the resources needed to carry out these tasks. In general, the cost of data acquisition can range from a few percent to more than half of the total cost of an ML project, depending on the scale and scope of the project. For a small-scale project, the cost of data acquisition might be relatively low compared to the cost of developing and training ML models. For large-scale projects, the cost of data acquisition can be substantial, especially if the data is collected from multiple sources, is complex, and high-dimensional.

With the increasing use of synthetic data, companies may be able to create proprietary, usable, and pertinent datasets with good intrinsic quality when data are not otherwise available or accessible, or are subject to specific regulations. Gartner projects more than 60% of data used in AI projects will be synthetic by the end of the decade. This expectation assumes expert knowledge of the domain and the target application to ensure appropriate quality and guarantee approximation of the desired data.

Bias and unfairness

ML models can perpetuate and amplify bias, leading to discrimination and unequal outcomes. Bias in training data can arise from a variety of sources, including demographic biases, human biases, or systemic biases.

When a ML model is trained on biased data, it can develop a skewed "understanding of the world" and make decisions that reflect and reinforce these biases. Bias can affect ML performance in multiple ways:

- Bias can cause systematic errors that skew the results of an algorithm in favor or against an idea. This can lead to inaccurate or unfair predictions that harm individuals or groups.
- Bias can cause overfitting or underfitting problems that reduce the generalization ability of an algorithm.
- Bias can cause ethical or legal issues that violate human rights or social norms. For example, bias can result in discrimination, privacy invasion, manipulation, deception or exploitation of vulnerable groups.

To minimize bias and unfairness in training datasets, it is important to:

- **Collect data from a diverse and representative range of sources** to ensure that the dataset reflects a broad range of perspectives, experiences, and demographic characteristics.
- **Ensure that data is preprocessed** in a way that minimizes bias, such as through data normalization, data augmentation, and data balancing.
- **Incorporate human oversight into the ML process**, such as through the use of annotators or ethical review boards.
- **Regularly monitor and evaluate the performance of ML models**, paying close attention to signs of bias or unfairness.
- **Use fairness metrics** to quantify and monitor the level of bias and unfairness in ML models.

Examples of bias in learning datasets that pass onto ML models predictions

- A Loan granting models: if a loan approval algorithm is trained on data that is biased against women, it may continue to unfairly favor men when evaluating loan applications, perpetuating gender discrimination in access to credit.
- Credit scoring models use personal and financial data to assess the risk of lending money to individuals. However, this data may exclude or disadvantage people who have low income, no credit history or belong to marginalized groups.
- Predictive policing systems that use historical crime data to allocate police resources. However, this data may reflect existing racial and social biases that lead to over-policing certain communities.
- Sentencing and parole models that use criminal records and demographic data to predict the likelihood of re-offending. However, this data may reflect existing biases in the criminal justice system that disproportionately affect people of color.
- COVID-19 diagnosis models that use chest X-rays and other medical data to detect infection. However, this data may be skewed by factors such as geographic location, testing availability and quality of images.
- Face recognition models that use facial images to identify or verify individuals. This data may be biased by factors such as skin tone, gender and age.

Poor data quality

Data quality is an essential factor in training datasets, directly affecting the accuracy and effectiveness of ML models. **Poor quality data can negatively impact model performance and lead to incorrect or undesirable outcomes**. Errors in data entry, measurement, or labeling, as well as missing or inconsistent data can lead to poor data quality. When a ML model is trained on poor quality data, it can develop a flawed understanding of the relationships between input and output variables, leading to inaccurate predictions and decisions.

To avoid the negative consequences of poor quality data, it is important to:
- **Validate data at the point of entry** and during the data cleaning process to identify and correct errors and inconsistencies.
- **Clean and preprocess data** to ensure that it is consistent, complete, and relevant.
- **Use data quality metrics** to quantify the quality of the data and monitor changes over time.
- **Establish rigorous and reliable data collection processes** to minimize the likelihood of errors and inconsistencies.

Road to success

Four keys to defining successful data and AI governance

Implement a dual top-down and bottom-up strategy:
Every data and AI governance program needs executive sponsorship. Without strong support from leadership, companies may not make the right changes (which are often difficult changes) to improve data security, quality, and management. At the same time, individual teams must take collective responsibility for the data they manage and the analysis they produce. Companies must establish a culture of continuous improvement and ownership of data issues. This bottom-up approach can only be achieved in tandem with top-down communications and recognition of teams that have made real improvements and can serve as an example to the rest of the organization.

Balance between governance and enablement:
Governance should enable and support innovation. In many cases, teams need to distinguish between proof-of-concepts, self-service data initiatives, and industrialized data products and recognize the governance needs associated with each category. Exploration and experimentation should be encouraged, but teams also need to make clear decisions about when self-service projects or proof-of-concepts should receive the funding, testing, and assurance to become an industrialized, operationalized solution.

Model management:
As ML and deep learning models become more widespread in the decisions made across industries, model management is becoming a key factor in AI governance. Continuous monitoring, model refreshes, and testing are needed to ensure the performance of models meet the needs of the business.

Transparent and responsible AI:
Decisions made with ML models are facing growing scrutiny. Models make decisions that impact people's lives every day. Both the people impacted and the companies producing models should understand the implications of the decisions models make.

Case study: AI detects tumors faster and more accurately than humans

Together with SAS, Amsterdam UMC is transforming tumor evaluations with AI. It uses computer vision technology and deep learning models to increase the speed and accuracy of chemotherapy response assessments. The process is not only faster but more accurate than when it is conducted by humans.

AI outcomes: The project started by training a deep learning model with data from 52 cancer patients. Every pixel of 1,380 metastases was analyzed and segmented. This taught the system how to instantly identify tumor characteristics and share vital information with doctors, allowing them to more accurately determine whether life saving surgery is viable, or a different treatment strategy should be chosen.

Data sharing: Outside the clinic, the AI platform is also available to more than 1,100 Amsterdam UMC cancer researchers to enhance their research efforts and supports the automatic translation of raw images to objective metrics in a clinical setting.

Explainable AI: For Amsterdam UMC, a critical factor in the ongoing deployment of AI and analytics within clinical settings is to establish an end-to-end auditable and transparent process for decision management in health care. The AI platform provides a summary of how each analysis is performed, making it easier for doctors to track their models and algorithms.

Data commons

As data becomes more central to the operations of modern businesses, **dataset sharing across organizations, in the form of a data common, is a potential game-changer**. Data commons are collaborative repositories of data that are openly available and freely accessible to its member organizations.

Companies can combine datasets and models to create synergistic performance with insights and predictions that would be difficult or impossible to achieve on their own, especially if the data common is created by companies in a given geography or in a given industry segment. Data commons make possible the success, use, and sharing of data to make informed decisions, improve operations, and provide better services, either horizontally (sharing across types of companies in or within a given industry segment), or vertically (sharing with upstream suppliers and downstream customers). Data commons are also beneficial to the wider economy because they facilitate data-driven innovation and collaboration.

Despite the potential benefits of data commons, many businesses resist sharing data, particularly small and medium-sized ones, mostly due to:

Addressing resource scarcity challenges and finding ways to incentivize participation in data commons is critical to realizing the full potential of this emerging trend, especially for SMEs whose resources (specifically data engineering and data management resources) are limited.

- **Fear of losing competitive differentiation**: horizontal sharing, in which companies in the same value chain share data, may lead to a loss of unique insights and competitive advantages. In order to participate in a data common, businesses may need to share sensitive data, which could make it easier for competitors to replicate their strategies.

- The perceived **risk of sharing data with upstream suppliers and downstream customers: vertical sharing** of data with these entities could lead to concerns about data privacy, intellectual property protection, and data misuse. Small and medium-sized businesses, in particular, may not have the resources to adequately protect their data or ensure that it is not misused by others.

- **Cultural barriers to data sharing**: some companies may have a "not invented here" mentality, in other words, they are resistant to accepting ideas or data from outside their own organization, making it difficult for businesses to see the benefits of sharing data and collaborating with others.

- **Difficulty in data governance**, ensuring that the data is accurate, secure, and up-to-date. This requires creating a system to ensure the data is constantly monitored and updated, that the data is kept secure, that it is properly shared and used, in a compliant and ethical way.
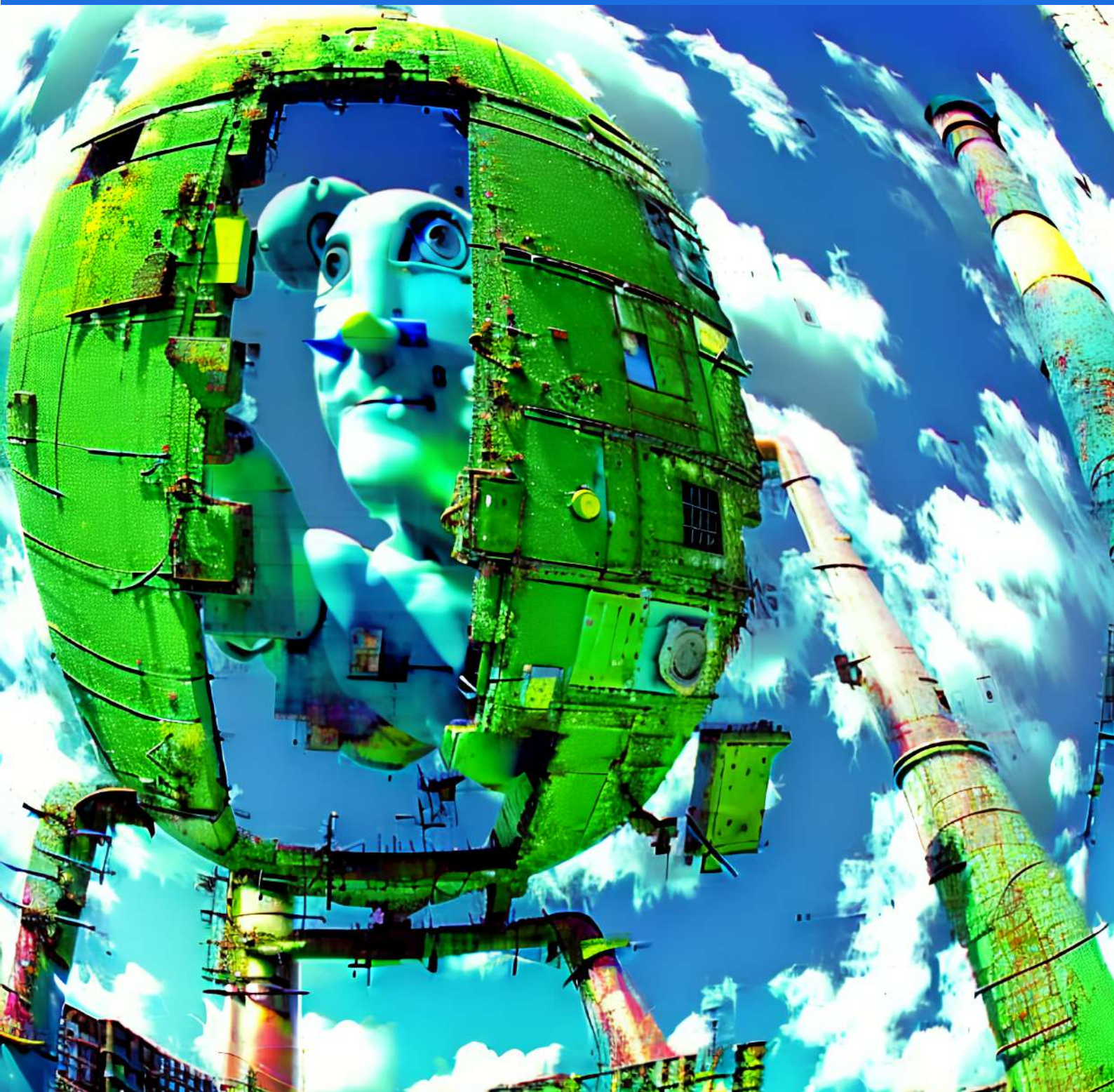
CHAPTER 8

TRUST AS A CONDITION TO AI ADOPTION

TAKEAWAYS

- **Trust in AI is important to foster the adoption of AI and alleviate potential concerns.**

- **Risks associated with AI include privacy and data protection, security, fairness and bias, transparency, and safety and performance.**

- **The European Commission's AI Act seeks to regulate the use of AI systems and introduce a common legal framework for AI across all sectors.**

- **Compliance with the General Data Protection Regulation (GDPR) is a critical aspect of AI trustworthiness.**

- **Businesses should take steps to foster trust in AI, such as implementing security, communicating transparently, providing training and education, and establishing ethical and fair use of AI.**

- **To stay informed on AI regulation in France, readers can join organizations such as Hub France IA, METI, Numeum, AFIA, or CPME.**

What is at stake?

Building trust in Artificial Intelligence (AI) systems is the responsibility of both providers and users. **Experts in AI governance agree that fairness, explainability, accountability, transparency, reliability, robustness, fairness, privacy, security, and human involvement and oversight are essential.** In this chapter, we focus on fairness, bias, transparency, and security; the other key trust-building elements are outlined in Chapter 11. Education and a strong legal framework create the conditions for users to trust AI systems.

To create trust, users must be able to rely on binding rules imposed on AI providers, which is why the EU has developed new norms based on existing work, such as the OECD's "Principles on Artificial Intelligence", that aims at identifying and mitigating the risks of AI, while also avoiding stifling innovation.

Why is trust in AI so important?

To foster the adoption of AI and alleviate the concerns that may hamper its growth, decision-makers, employees, users and regulators must have confidence in AI across multiple dimensions.

Mistrust in AI often revolves around the belief that AI may become uncontrollable and make poor decisions impacting individuals, for example, reproducing or fostering bias; violating privacy by deciphering habits, emotions and thoughts; or misinterpreting signals leading to mistakes. Another common concern is that AI may take people's jobs, a recurring debate recently revived by ChatGPT and its impressive capacity to answer questions and to generate text with minimal prompting.

The OECD Principles on AI

The OECD (Organisation for Economic Co-operation and Development) principles on AI are a set of guidelines for the development and deployment of trustworthy AI. Developed by the OECD in 2019, they are intended to serve as a global standard for AI policy.

The principles consist of five values: inclusive growth; sustainable development; human-centered values; transparency and explainability; and robustness and safety. Human-centered values, for example, include the recommendation that AI systems respect human rights, be designed for user well-being, and incorporate mechanisms for feedback and redress.

While the OECD principles are voluntary and non-binding, they have been endorsed by many governments and organizations around the world as a useful framework for the development and deployment of AI.

What are the risks?

The primary tangible risks associated with AI are **privacy and data protection, security, fairness, bias, transparency, and safety and performance**. AI models that use personal data pose privacy risks and can violate GDPR rules if not implemented with appropriate safeguards. Cyberattacks are a major vulnerability in industries like banking and healthcare, resulting in legal regulations to protect data in these fields such as DORA and NIS 2. AI has also been accused of bias and lack of fairness; historical biases in datasets can perpetuate discrimination. Transparency is needed for people to trust AI; without it, legal issues and non-compliance with data privacy and discrimination laws can arise. Last, AI is a new technology that can operate autonomously giving rise to novel issues of safety and performance.

To combat these risks, institutions such as the European Commission are creating regulatory frameworks such as the proposed AI Act, which includes a classification of AI systems based on the risk they pose to EU citizens, and sanctions against abuses. Appropriate legal remedies increase confidence in AI solutions, such as the $20 million fine imposed by CNIL on Clearview AI for unlawfully processing photographs of individuals.

What regulation exists to safeguard against these risks?

The European Commission's AI Act seeks to regulate the use of AI systems, including those posing high risks for individuals, and introduce a common legal framework for AI across all sectors (except the military and "national security" applications). The proposal regulates providers and professional users of AI systems, but does not confer rights to individuals. While the EC aims to set a global standard for AI regulation, local jurisdictions are developing their own legal frameworks.

The AI Act (currently in draft form) will egulate AI based on risks to the EU citizen, with three risk levels identified:

Unacceptable-risk AI systems	High-risk AI systems	Limited- and minimal-risk AI systems
• Subliminal, manipulative, or exploitative techniques causing harm • Real-time, remote biometric identification systems used in public spaces for law enforcement • All forms of social scoring	• Systems that evaluate consumer creditworthiness • Recruiting or employee-management systems • Systems utilizing biometric identification in non-public spaces • Safety-critical systems (e.g. systems that would put the health of citizens at risk due to failure!) • Any systems used in the administration of justice	• AI chatbots • AI-enabled video and computer games • Spam filters • Inventory-management systems • Customer- and market-segmentation systems • Most other AI systems

Many AI systems, such as spam filters, will be classified as bearing limited or minimal risk and will not be subject to the requirements of the proposed AI Act. Other AI applications, such as deep fakes, will have transparency obligations to ensure users are not deceived. High-risk AI systems, such as those used to evaluate credit-trustworthiness, recruit staff, admit immigrants, or administer justice, will be subject to risk assessment, monitoring, and continued responsibility on the part of their creators, distributors, and operators. Finally, certain AI systems, such as social scoring and real-time bio-metric recognition by law enforcement, are prohibited in the EU, as they represent an unacceptable risk. AI systems used solely for military and national security purposes are exempt from the EU AI Act, although their use could threaten rights such as privacy, freedom of movement, assembly, expression and participation.

AI and GDPR

Because current use of AI systems relies heavily on processing large amounts of data, compliance with the General Data Protection Regulation (GDPR) is a critical aspect of AI trustworthiness. The French "Commission Nationale de l'Informatique et des Libertés" (CNIL) has published resources to help companies adhere to GDPR when using AI. While some GDPR requirements provide direct protection, others offer indirect protection, such as the principle of data quality, which aims to ensure that output is not affected by poor-quality data. Nonetheless, if AI systems are not designed with data protection in mind, some GDPR rules can be a challenge, such as the data minimization principle, which only permits the processing of relevant and necessary personal data.

Common pitfalls

The lack of a clear legislative framework for AI affects security and trust, forcing corporations to be cautious. **Clear guidance from authorities, including the CNIL, is needed for areas of risk such as bias and discrimination.** To this end, the CNIL has created an AI service that provides initial recommendations on learning datasets, but more resources are needed. Given the concern over possible foreign state seizure of data and hacking, data storage and transfer pose a particular problem for trust and security.

A trusted cloud initiative in France

In October 2021, Thales and Google Cloud announced "S3NS", a partnership to develop a locally-compliant trusted cloud offering, helping companies and public institutions to innovate and to accelerate their digital transformation while ensuring autonomy, compliance with sovereignty, and maximum benefits for customers and users. S3NS will launch in the second half of 2024. It will be protected against extraterritorial foreign laws and will comply with the requirements of the Trusted Cloud label from France's Information Systems Security Agency (ANSSI). S3NS will directly operate three data centers in France to ensure data and workload localization.

Road to success

Building trust among employees and stakeholders is crucial for the successful implementation and adoption of AI technology. Businesses can take steps to foster trust in AI deployment:

- **Implement security** to ensure that AI is safe for adopters and users.
- **Categorize risks associated with AI**: performance problems, data leakage, and auditability. While different companies have different risks, creating a typology can help anticipate problems and avoid "false problems".
- **Communicate transparently**, as open communication is key to building trust. Businesses should clearly explain the reasons behind the deployment of AI, the goals they hope to achieve, and the potential benefits and drawbacks for employees and stakeholders.
- **Involve employees and stakeholders** in the deployment process to help build trust and buy-in.
- **Provide training and education on AI** to help employees understand the technology and how it will impact their jobs. This can also address concerns and dispel myths about AI.

- **Ensure ethical and fair use of AI** by establishing clear guidelines and ensuring that the technology is being used fairly. This includes protecting employee privacy and avoiding biased decision-making..
- **The decision-making process of AI should be transparent and explainable**.
- **Regularly monitor and evaluate AI performance** to help to ensure that the technology is delivering on its promises and that any negative impacts are identified and addressed.
- **Create a culture of trust and transparency**. Businesses should encourage open communication and provide a safe space for employees and stakeholders to voice their concerns and ideas.
- **Address fears around sovereignty**, for instance favoring the use of open-source AI tools such as "Scikit-learn", and storing data within national territory.
- **Use abstainer-classifier models** (the model only makes predictions when a high degree of certainty has been computed, and abstains from predicting otherwise) to reduce issues with liability and certification.

CHAPTER 9

RESPONSIBLE AI

TAKEAWAYS

- **Businesses must prioritize trust and ethics in order to make the most of AI.**

- **Responsible AI must be designed ethically and with safeguards to protect human rights, diversity, and a fair society, and be transparent and secure throughout its lifetime.**

- **Companies should strive to create a culture of responsible AI and take into account the environmental impact of AI.**

- **Businesses must evaluate their digital carbon footprint, consider the lifecycle of infrastructure, and use low-carbon energy sources.**

- **Companies must carefully evaluate their AI strategy and be aware of the risks associated with AI.**

- **To deploy responsible AI, businesses must establish ethical frameworks, transparency and traceability, data governance and privacy, stakeholder engagement, responsibility and liability, continuous monitoring and evaluation, skills and capacity building.**

What is at stake?

The implementation of Artificial Intelligence (AI) solutions raises questions that could impede their adoption and even pose risks for organizations and individuals who might benefit (or suffer) from their implementation. While fiction has often portrayed the risks of AI, often in an exaggerated manner, business stakeholders are moving towards the implementation of trustworthy AI that is likely to be widely accepted by the public.

What is responsible AI?

As AI technology advances, responsible principles must govern its development. This includes making sure security protocols are in place and that individuals, communities, and businesses are not put at risk; establishing regulations to ensure AI is not misused or abused; building AI systems on ethical foundations to prevent discriminatory or unfair outcomes; taking into account the diversity of ethics across societies and cultures; and understanding and addressing the environmental impact of AI.

Businesses should strive to use AI for positive outcomes and guarantee sufficient human oversight and control as AI becomes increasingly autonomous. As public expectations move towards favoring businesses with a purpose, ESG (environmental, societal, and governance) principles should be prioritized in the development and implementation of AI. Ethical principles must be applied at all stages of the AI lifecycle, from innovation to development, deployment, and production, with proper human oversight and control.

The lay of the land

In 2018, the European Commission released the work of the High Level Group on AI (HLEG-AI): "Ethics Guidelines for Trustworthy AI", aimed at promoting lawful, ethical, and robust AI systems that respect human dignity, autonomy, privacy, and diversity. The guidelines identify seven key requirements that trustworthy AI should meet: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity and non-discrimination and fairness, societal and environmental well-being, and accountability.

In 2019, 42 countries adopted the OECD Principles on Artificial Intelligence, which state that AI should be designed ethically and with safeguards to protect human rights, diversity, and a fair society, and be transparent and secure throughout its lifetime. These two documents provided the basis for subsequent regional and local regulatory work, including the EU's forthcoming "AI Act".

Outisde of Europe, the US Federal Government has also made AI R&D a priority, focusing on the ethical, legal, and societal implications of AI as well as the safety and security of AI systems. The National AI R&D Strategic Plan (2019) outlines many of the challenges in these areas. China's Ministry of Science and Technology published the "Governance Principles for a New Generation of Artificial Intelligence: Develop Responsible Artificial Intelligence" in 2019, which provides a framework and action guidelines for responsible AI development, aiming to promote a new generation of AI that is safe, secure, reliable, and controllable.

In 2020, 29 countries joined forces to create the Global Partnership on AI (GPAI), hosted by the OECD. This partnership is tasked with bridging the gap between the theory and practice of AI, supporting coordinated research policies and applied activities on shared AI-related priorities. The GPAI's initial mandate focuses on responsible AI.

Self governance example: Google's AI Principles

Google's AI Principles serve as the company's ethical charter. Consistent policies and responsible practices enable the structured governance that allows for scaling the practice of principled AI innovation.

Google believes AI should be socially beneficial, avoid creating or reinforcing unfair bias, be built and tested for safety, be accountable to people, incorporate privacy design principles, uphold high standards of scientific excellence, be made available for uses that accord with these principles.

In addition to the above, Google is committed to not design or deploy AI in the following application areas:

- Technologies that cause or are likely to cause overall harm.
- Weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people.
- Technologies that gather or use information for surveillance violating internationally accepted norms.
- Technologies that contravenes widely accepted principles of international law and human rights.

As the use of AI increases, there is growing concern about whether it aligns with ethical and moral values, and whether it can provide benefits while minimizing risks. AI is based on real-world data and cultural habits, which can lead to the replication of biases and prejudices in algorithms. Even small changes to the data used to train AI systems can lead to discrimination or unfair advantage.

Trust and ethics

As addressed extensively in Chapter 8, **businesses looking to make the most of AI must prioritize trust and ethics**. To earn and maintain the trust of customers, partners, employees and stakeholders, companies should take a responsible approach to AI that accounts for its potential impacts on our lives and society - from personal privacy to job displacement. By creating a culture of responsible AI, businesses can build more sustainable, trustworthy, and ethical systems that better serve the needs and interests of all parties.

AI-specific environmental costs

Large language models training, ChatGPT, Bard, and similar platforms will add hundreds of millions of users to the AI infrastructure. A new society is emerging. But at what cost to the environment? **While digital technology has the potential to vastly reduce resource usage and pollution** (such as through the improvement of process efficiency and making physical supply chains partially obsolete) **its environmental impact continues to grow.** In the US, data centers already account for 2-3% of total electricity consumption and this number could double or triple by 2030. In France, ADEME estimates that digital technology already accounts for 3.5% of greenhouse gas emissions. ChatGPT runs hundreds of parallel learning cycles, analyzing the equivalent of 33 times the total volume of Wikipedia data, resulting in an estimated consumption of 552 tons of CO_2 - the equivalent of 2.5 million kilometers traveled by car. Currently, the digital world accounts for 4% of total greenhouse gas emissions worldwide: 25% attributed to data centers, 28% to network infrastructure, and 47% to equipment. This carbon footprint will only grow larger with the increased adoption of AI.

Large scale deployment of green AI could increase global GDP by 3.1 to 4.4% while simultaneously reducing global greenhouse gas emissions by 1.5 to 4 % by 2030, according to some sources. Europe stands to gain even more, with GDP gains topping $1 trillion.

Green AI: real-world examples

- **ABB** provides AI-powered solutions for building automation, energy management, and smart grids.
- French startup **Agricool** is using AI and ML to optimize indoor farming systems, leading to improved crop yields and reduced water usage. Their platform uses data on plant growth and environmental conditions to optimize lighting, watering, and nutrient delivery.
- **Apple** is using AI and ML to create algorithms that can analyze satellite images of forests and identify areas of deforestation, as well as areas of reforestation. This technology is being used to help identify areas that need to be protected or reforested and can be used to inform policy decisions.
- **DuPont** and **Syngenta** are using AI and ML to create algorithms that can detect diseases in crops and help farmers detect them earlier and more accurately, reducing the amount of pesticides needed.
- **EDF** is using AI and ML to optimize their energy management systems and reduce energy waste. For example, they are using ML algorithms to predict energy demand and optimize the use of renewable energy sources such as wind and solar.
- Google's **DeepMind** AI is being used to improve the energy efficiency of its data centers, cutting down on the amount of electricity needed to power them. This is being done by using ML models to analyze the data from the data centers and identify ways to optimize their energy usage. Microsoft and IBM, are also using AI and ML to create more efficient data centers.
- **IBM** has a partnership with The Freshwater Trust to develop an AI-based system for monitoring and restoring freshwater ecosystems, and a program to develop carbon capture technologies using AI.
- **IEA**, an intergovernmental organization, conducts case studies on how AI can improve building energy management systems
- **Microsoft** partners with the National Geographic Society to monitor and protect wildlife, and an AI for Earth program that provides grants and resources to organizations working on environmental issues.
- **Qucit**, a French startup is using AI to help cities and businesses optimize their transportation and mobility systems. Their platform uses real-time data and predictive algorithms to improve traffic flow and reduce congestion, leading to reduced greenhouse gas emissions and improved air quality.
- **Schneider Electric** leverages AI to monitor and control building energy consumption and performance.
- **The Climate Corporation**, a subsidiary of **Bayer**, uses data-driven insights and predictive modeling to help farmers optimize their planting, watering, and fertilization practices, leading to reduced greenhouse gas emissions and improved crop yields.
- **The Wildlife Conservation Society** is using AI and ML to monitor wildlife populations and combat wildlife trafficking. Their system, the Wildlife Crime Tech Challenge, uses image and acoustic recognition algorithms to identify species and track their movements.
- **Total** is using AI and ML to optimize their energy production and reduce their carbon footprint. For example, they are using ML algorithms to improve the performance of wind turbines and to optimize their solar panel installations.
- **Waste Management Inc.** is using ML algorithms to optimize waste management processes, including sorting recyclable materials and improving waste collection routes.
- **Xcel** uses AI to forecast weather-dependent renewable energy sources such as wind and solar power.
- French agritech startup **Ynsect** is using AI and ML to improve the efficiency of insect-based protein production, which has the potential to reduce the environmental impact of traditional animal farming. Their platform uses data on insect growth and feeding patterns to optimize the production process.

Common pitfalls

The introduction of AI solutions that are fair, accountable, transparent, and ethical has been a major focus of European and US-based organizations. At the same time, over 30 nations are developing strategies to utilize AI to increase their global influence. While these strategies address the economic, environmental, and societal implications of AI, they are based on **culturally specific values, which makes consensus on responsible AI difficult to achieve**. **Climate transition**, for example, has demonstrated the difficulty of establishing international standards, a process that will require the consensus of individuals and businesses. Similarly, **the protection of intellectual property rights** and the potential for AI systems to be weaponized are major concerns that will require debate and consensus building among stakeholders at every level.

Other pitfalls that SMEs may encounter when trying to deploy responsible AI include the unintentional introduction of bias or discrimination into AI systems due to lack of diversity in their teams, or failing to consider the impact of their decisions on different groups of people. SMEs may also struggle to provide transparency and explainability around their AI systems, which can erode trust and limit accountability. They may struggle to navigate the complex legal and regulatory landscape around AI, which can make it difficult to ensure that their systems are compliant with relevant laws and regulations.

Road to success

Companies have actionable means to achieve ESG goals: **evaluating their digital carbon footprint**; **considering the lifecycle of infrastructure**; **developing low-carbon roadmaps**; **raising awareness among employees** about the challenges of sustainable digital development; **using low-carbon energy sources**; and **adopting circular economy processes for the digital realm**.

As part of this effort, because the carbon footprint of an AI system is directly linked to the size of its training dataset, ML methods involving less data are being developed. Techniques such as distillation, transfer learning, and small dataset learning are being developed to improve the environmental impact of the training phase.

In addition, developers are questioning the unlimited use of digital technology, creating models for responsible usage based on simple, low-cost, and recyclable technologies such as repairable hardware and software solutions and slowing down cycles of planned obsolescence.

Since 2022, ADEME has offered a six-module course in sustainable digital development for teachers and educators. At ESSEC Business School, sustainability is integrated into all cross-disciplinary programs related to digital technology.

Responsible AI facilitates hiring

Responsible AI has been identified as a key factor in a growing number of candidates' decisions when considering job offers, according to recent surveys. By enabling AI decision-makers and developers to factor in ethical considerations, organizations can encourage them to find creative ways to achieve the company's objectives while remaining responsible.

The importance of ethical and regulatory challenges

To successfully implement responsible AI, companies must carefully evaluate their AI strategy. Business leaders must be aware of the risks associated with AI. This is especially important because AI is able to learn from past experiences and act autonomously or semi-autonomously, leading to non-compliance with stakeholder, employee, and end-user expectations. AI models must adhere to ESG principles and protect individual rights. To achieve responsible AI implementations, institutions, researchers, and AI platform providers offer a range of tools (such as explainability functions, regulations, and frameworks), but **it is the responsibility of companies to understand and implement these tools**.

The Confiance.AI project in France is developing recommendations for methods and tools to build trustworthy AI applications and define standards of auditability. Impact AI, a "think/do tank" which provides guidelines and best practices to organizations to help them work towards trustworthy and responsible AI, focuses on four topics: ethics, sponsoring, governance model, and protocols & tools. Both projects are part of the French national strategy on AI and are in line with the EU's seven principles for trustworthy AI.

Principles for governance

Public and private leaders must ensure that AI applications are based on principles of transparency and robust governance in compliance with ethical principles and regulations. AI systems must be acceptable to individuals and communities alike.

The key components of responsible AI governance for businesses are:

- **Ethics and values**: A clear ethical framework defines the values and principles that guide the development and use of AI systems.
- **Transparency and traceability**: Companies ensure their AI systems are transparent, with clear documentation on how decisions are made and a record of the data used to train models. This helps build trust with stakeholders and enables companies to detect and address biases or unintended consequences.
- **Data governance and privacy**: Companies establish strong data and model governance policies and practices to ensure the protection of personal data and privacy and the integrity and quality of their data, as addressed in Chapter 9. This includes implementing secure data storage, transparent data usage policies, and regular monitoring and auditing of AI systems.
- **Stakeholder engagement**: Companies engage with a wide range of stakeholders, including employees, customers, regulators, and civil society organizations, to better understand the ethical and social implications of AI.
- **Responsibility and liability**: Companies establish clear mechanisms for holding individuals and organizations accountable for the impacts of AI systems: defining responsibilities and roles, establishing reporting and remediation processes, and ensuring adequate insurance coverage.
- **Continuous monitoring and evaluation**: Companies continuously monitor and evaluate their AI systems to ensure they remain aligned with ethical and social values and that any unintended consequences are detected and addressed.
- **Skills and capacity building**: Companies invest in the development of skills and capacity within their organizations to ensure that AI systems are developed and used in a responsible and ethical manner.

CHAPTER 10

WHAT'S NEXT IN AI FOR BUSINESSES ?
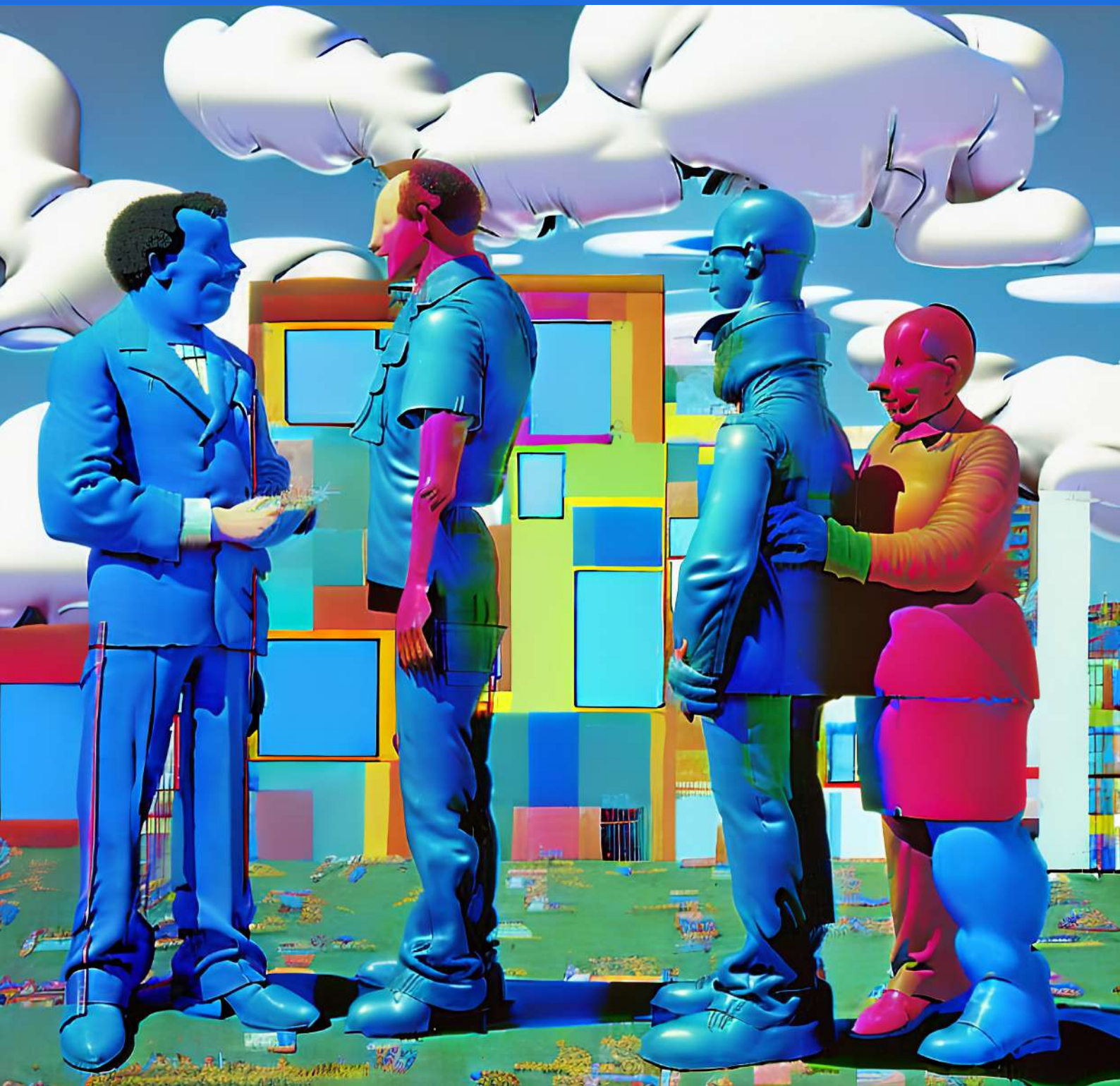
TAKEAWAYS

- **Generative AI systems such as ChatGPT will revolutionize the way businesses use AI, drastically increasing productivity in some business functions. Companies must quickly put in place usage recommendations and controls for ChatGPT and generative AI, to ensure compliance with ethical standards and laws.**

- **Automation of mundane tasks, and other applications such as drug discovery and synthetic data will create further opportunities for small and medium businesses in Europe.**

- **New roles & competencies will emerge in the industry.**

- **Sustainable AI will become a hot topic, with firms scrutinized over their use of AI.**

- **Intellectual property and content consistency will be issues to address. Content authentication and verification will be required.**

- **Disinformation networks will take advantage of generative AI, further sowing confusion and eroding trust across society.**

2022 was a banner year for Artificial Intelligence (AI), culminating in the public release of ChatGPT, a generative AI system developed by OpenAI, in December. Michael Spencer, futurist and writer, qualifies it as "the cambrian explosion of AI" to describe the rapid emergence of generative AI systems that can create new content such as text, images, music and code.

The Cambrian explosion of AI?

The analogy of the rapid development of usages for generative AI since 2022 with the Cambrian explosion is based on the idea that both events represent a sudden and massive diversification of "life forms". The Cambrian explosion was a period of time approximately 538.8 million years ago when many of the major animal groups appeared in the fossil record for the first time. It was characterized by an unprecedented increase in complexity, variety and ecological interactions among animals. Similarly, recent developments in AI are seen as a period of time when many new types of generative AI systems emerged that can create novel usages across various domains and tasks. It is characterized by an unprecedented increase in data, compute power and natural language capabilities among machines. Spencer argued that LLMs (1) are a key component of this generative AI revolution, as they can leverage massive amounts of data and compute power to produce natural language outputs across various domains and tasks.

---

1. *LLMs (for "Large Language Models") are neural network-based ML models that have been trained on massive amounts of text data, using techniques such as unsupervised pretraining, self-attention mechanisms, and transformer architectures. These models can perform tasks, such as text generation, language translation, question answering, sentiment analysis, and text classification, among others, with a high degree of accuracy and fluency.*

2023: The end of the beginning of AI?

The rise of generative AI in 2022 is the most exciting technological development since mobile and cloud technologies emerged over a decade ago. Mobile paved the way for the growth of major consumer internet companies like Uber, Lyft, Instagram, or Snap, while cloud technology facilitated the growth of software-as-a-service (SaaS) companies like Slack, Airtable, or Stripe. In the next few years, we can expect a similar explosion in innovation driven by generative AI, which is poised to play a critical role in shaping the future.

A word of caution: the ChatGPT illusion

ChatGPT was launched in December 2022. In just a few weeks, the system was adopted by more than one million users, the fastest user uptake in the history of digital platforms. Behind ChatGPT's success lies a disturbing truth: millions of users are using a system for purposes it was not developed for. Because they are engaged in conversation with a machine, many ChatGPT users think of the platform as an advanced search engine. But ChatGPT is much more, and much less, than that.

Unlike search engines, ChatGPT in its standalone form is not designed to identify relevant sources of information based on a query, then sort and present them.

Sam Altman ✔
@sama

ChatGPT is incredibly limited, but good enough at some things to create a misleading impression of greatness.

it's a mistake to be relying on it for anything important right now. it's a preview of progress; we have lots of work to do on robustness and truthfulness.

1:11 AM · Dec 11, 2022

3,485 Retweets   722 Quote Tweets   28.7K Likes

On the contrary, ChatGPT has been "trained" by ingesting a "learning set", a frozen set (in 2021) of millions of texts found on the open web and in specialized corpora, in the legal and medical fields, for example. The information that ChatGPT can "bring up" in response to queries is, for many, outdated. The system evaluates the prompts it receives, then constructs a grammatically correct response based on what it has learned from the structure of the texts presented to it during the model learning phase without any clue about the meaning of the prompt or the answer it provides. The information it presents should never be taken seriously without careful verification. The size of the learning dataset, combined with well-implemented probabilistic mathematics, give the illusion of meaning. A statistical serendipity, as it were. **ChatGPT should never be blindly trusted.**

If ChatGPT hints at what search engines could become in the near future (such as Microsoft's integration of GPT3.5 into their Edge browser, or Google's Bard, announced early February 2023), **in its current state, it is only an automatic text generator that produces often correct, but unreliable results**.

These limitations will not prevent the explosive adoption of ChatGPT and other LLM-based, ChatGPT-like solutions by people and businesses, given the tremendous value they already offer, helping authors and writers across a wide spectrum with tasks such as writing summaries, outlines, introductions, conclusions, captions, titles, bullet points and more; providing feedback, suggestions and corrections to improve the quality and clarity of the documents; and adapting to different styles, tones and formats depending on the context and purpose of the documents.

Will LLM-based platforms lower resistance towards AI?

Today, more and more companies are implementing progressive and orderly AI deployment strategies. Some are leaving the initiative for AI deployment in a limited scope to operations; but when it comes to deploying AI solutions to all employees, centralization remains the norm.

The adoption of chatGPT, soon present on every screen, behind every desktop, is going to significantly disrupt these strategies. It's going to be hard to ask people to give up using chatGPT until the right governance systems are in place. We are going to have to adapt, and fast. To move from a top-down, directive model to an attitude of welcoming and integrating these new tools.

We think that **ChatGPT (with its imperfections) will ultimately prove to be a catalyst for the adoption of AI in companies**: because it will make everyone "feel" the value of AI, because it will highlight in a very concrete way the possible productivity gains, ChatGPT could well help to unblock the reluctance to adopt AI solutions that many companies are experiencing.

For this to happen, companies must quickly put in place recommendations for use, controls, and recourse in the event of harmful use. At a minimum, companies should adopt and disseminate usage recommendations for ChatGPT, a generic version of which could look like :

> ## Guidelines for ChatGPT adoption in business
>
> 1. **Define an acceptable scope of use, and prohibit any other use.**
> 2. **Build trust through education: explain the limitations of the tool, remind employees that ChatGPT cannot guarantee the veracity of the information generated, remind the content producer of their responsibility and the need to manually verify all information.**
> 3. **Ensure that there is a resiliency strategy: *"I use ChatGPT, what will I do if I can't access it anymore?"***
> 4. **Enter into a contract with the vendor to reduce the risk of unexpected loss of access to the system.Ensure compliance with ethical standards and laws when using it.**
> 5. **Implement a system to detect content that has been produced by ChatGPT**

In conclusion:

Should businesses embrace ChatGPT? Certainly, for the increase of individual productivity, which will benefit the entire company;

Will Chat GPT eliminate jobs? Probably not in numbers, because helping people become more efficient, more relevant, and more effective adds more value to their work.

In the end, why not see ChatGPT as simply a powerful tool, the natural successor of the word processor and the search engine? A wonderful tool, full of promise, but not to be adopted blindly.

Sustainable AI should be in every mind

There is no escaping the AI explosion. A new society is emerging, but at what cost? In 2023, companies will face pressure to reduce their carbon footprint and minimize their impact on the environment. AI algorithms and their supporting infrastructure, such as cloud networks and edge devices, will require ever more power and resources. While AI can improve resource use and reduce pollution, its environmental impact continues to grow, conflicting with the goals set at COP 25.Europe is a climate leader with ambitious sustainability objectives. AI users and developers must do their part by committing to these goals. Sustainable AI is not only possible, it is necessary and its development must be considered holistically. This requires a collective awareness: in companies, at home, at school, and at universities.

In 2023, AI adoption will surge as models are widely adopted in various applications. "New-age" apps will proliferate and blur the line between reality and science fiction. AI will become mainstream.

Our predictions

This workgroup sees the below as (some of) the main AI trends to watch in 2023 and beyond, loosely sorted from the short term to a longer time horizon:

- **The exponential development of generative AI, beyond images and text**. In January 2023, Google Research announced MusicLM, "a model generating high-fidelity music from text descriptions such as "a calming violin melody backed by a distorted guitar riff." Watch out for other application areas for LLMs.

- **Heightened awareness, education, and ethical concerns across the industry** as AI applications begin to have a real-life impact.

- **Augmented work gets a boost**: Generative AI will find its way into corporations, dramatically increasing productivity, especially when merged with semantic processing systems such as those which power search engines.

- **Generative AI turns the English language into the hottest new coding paradigm**. The future of low code/no code solutions remains to be seen.

"The hottest programming language is now English."

Andrej Karpathy, OpenAI, Tesla, Stanford

- **Image generative AI will enter a difficult phase**, with issues such as intellectual property (content creators and artists opting out of training datasets, for example), and content consistency for video and games production.

- **The emergence of new roles and competencies will require quick adaption**: Model owner, prompt engineer, FinOps, DataOps, ModelOps, MLOps, data stewards.

- **New versions of large language models will dramatically increase in performance but likely won't be larger than current versions** (OpenAI GPT 3.5: 175 Bn parameters, Microsoft/Nvidia's Megatron-Turing: 570 billion). Instead, development efforts will focus on the quality of the training dataset. GPT 4 should be proof of this.

- **DeepMind's AlphaFold, MetaAI and others will push the boundaries of drug discovery** using LLMs, leading to novel therapies such as the mRNA vaccines of 2021-2022.

- **Synthetic data will continue to gather momentum**, especially in data-poor applications and in medicine, where it can be used to simulate large populations of people with diseases, enabling powerful "what if?" scenarios at a fraction of the cost and time required by clinical trials.

- **The AI Act will increase regulatory pressure on both AI producers and AI users**.

- Once the hype fades out, and actual usages start sticking, it will be clear that **the main value of textual generative AI will free text workers and coders from low-level tasks**.

- **Content authentication**, possibly involving the blockchain, will be required.

- **Trust in AI systems will be boosted by technological advances and organizational processes**, leading to better explainability and the operationalization of AI transparency, leading to dramatic improvement in adoption, business performance, and user acceptance.

- **Disinformation networks will take advantage of ChatGPT**, attacking social media and crafting fake websites at a volume we have never seen before.

- Advances in data commons platforms and governance will enable **low-data businesses to safely cooperate with peers in vertical segments** to leverage the power of ML.

- **Sustainable AI will become a hot topic**, with firms scrutinized over their use of AI. Small dataset learning and other data use reduction technologies will emerge.

- **Metamodels**: Deploying ML against data to improve data quality and completeness, assessing data vulnerability, and data duplication will be integrated further up the data pipeline, in master data management, data governance, and extract-transform-load (ETL) processes.

- **The emergence of composite AI**: Hybrid systems that tightly integrate symbolic AI with combinatorial systems will drive the next wave of innovation in AI.

- **The hype over AI "becoming smarter than humans" or "replacing jobs" will continue**: pundits know that fear sells.
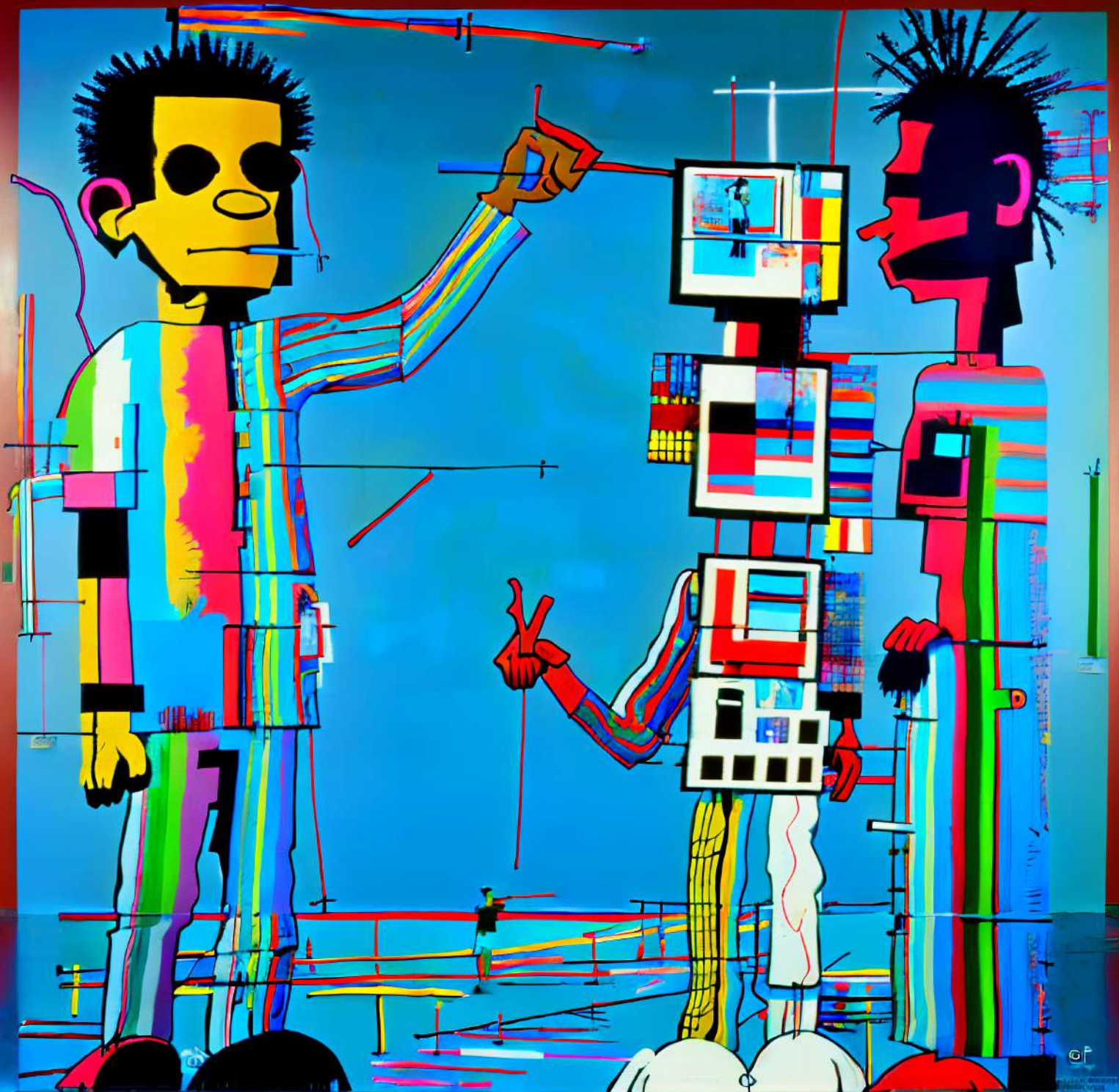
CHAPTER 11
GLOSSARY

AI ("Artificial Intelligence")

AI is the field of computer science that aims at simulating, not duplicating, human intelligence processes in machines. The term was used by the first time in 1956 by John McCarthy as a way to describe the idea of creating machines that could perform tasks that would typically require human intelligence, such as learning, problem-solving, and decision-making. Progress in algorithms and increased computing power have led to the development of machine learning, a subset of AI.

In this document, the terms Artificial Intelligence (AI) and Machine Learning (ML) are used interchangeably.

AI according to Oxford University

"The theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages."

AI according to the OECD

"AI is a general-purpose technology that has the potential to improve the welfare and well-being of people, to contribute to positive, sustainable global economic activity, to increase innovation and productivity, and to help respond to key global challenges. It is deployed in many sectors ranging from production, finance and transport to healthcare and security. Alongside benefits, AI also raises challenges for our societies and economies, notably regarding economic shifts and inequalities, competition, transitions in the labor market, and implications for democracy and human rights."

AI Act ("Proposal for EU regulation of AI")

The AI Act is a law regulating the use of AI proposed by the European Union, the first of its kind by a major regulator. It categorizes AI applications into three risk levels: banned, high-risk, and unregulated. Applications that pose an unacceptable risk, such as government-run social scoring, are banned. High-risk applications, such as tools used to scan CVs, are subject to specific legal requirements. Applications not banned or listed as high-risk are largely unregulated. Like the EU's General Data Protection Regulation (GDPR) passed in 2018, the EU AI Act could become a global standard. The AI Act is examined in more detail in Chapter 8 and Appendix B.

AI Ethics

In the context of AI, ethics refer to the moral principles and values that should guide the development and use of AI. These principles and values are concerned with ensuring that AI is used in ways that are fair, responsible, and beneficial to society. Some key ethical considerations in AI include:
- Fairness: ensuring that AI systems do not discriminate against certain groups of people based on factors such as race, gender, or age.
- Transparency: making sure that the decision-making processes of AI systems are understandable.
- Accountability: holding developers and users of AI systems responsible for their actions and decisions.
- Safety: ensuring that AI systems do not cause harm to humans or the environment.
- Privacy: protecting the personal information and data of individuals from unauthorized access or misuse.

Abstainer classifiers

Abstainer classifiers, also known as reject option classifiers, are ML models that have the ability to abstain from making a prediction on a given input, rather than providing a potentially incorrect classification. This is useful when the model is uncertain about the correct classification, or when the input falls outside the scope of the model's training data. One of the benefits of abstainer classifiers is that they can reduce the number of false positives and false negatives, which can be especially important in applications such as medical diagnosis or fraud detection. By allowing the model to abstain when it is uncertain, the risk of making an incorrect prediction is reduced.

Accountability

In the context of AI models, accountability refers to the ability to attribute responsibility for the outcomes of, or decisions made by, AI systems. A system must document decisions such that negative impacts can be traced back to specific individuals or organizations. This includes identifying and addressing biases in data or algorithms used to train an AI model, and ensuring that the system is transparent and auditable. AI accountability builds on Explainability (defined below) by ensuring mechanisms are in place to attribute responsibility for a model's decisions and outcomes, and by taking action to correct negative impacts. Together, these concepts ensure that AI systems are transparent, auditable, and that the people and institutions that created them are responsible for their outcomes.

Augmented intelligence

Augmented intelligence refers to the use of AI technologies to enhance human intelligence and decision-making capabilities, rather than replace them.In contrast to the traditional use of AI to create machines that can perform tasks that would typically require human intelligence, augmented intelligence combines the strengths of both human and machine intelligence to achieve better results and insights. The term "augmented intelligence" better describes the real life use of AI than "artificial intelligence" because it acknowledges the critical role that human intelligence and human supervision play in data-based machine predictions and decision-making. "Augmented intelligence" demonstrates a more positive and collaborative understanding of the relationship between humans and machine learning, while the term "artificial intelligence" can evoke negative feelings and even fear.

Data common

A data common is a platform that enables sharing and analysis of data sets among different users and organizations. Data commons can benefit small and medium enterprises (SMEs) by providing them access to valuable datasets that they would not be able to assemble on their own, ultimately leading to insights and innovations that can improve their competitiveness and productivity. Setting up data commons often represent challenges such as ensuring contribution from the various stakeholders, ensuring data quality, security, privacy, interoperability and governance.

Data governance

Data governance is the management and control of data within an organization. It includes the establishment of policies, standards, and processes to acquire, store, manage, and use data in line with the organization's objectives. Data governance is essential for accuracy, consistency, privacy, security, ethical use, and legal compliance. It encompasses activities like setting standards, executing data management processes, teaching best practices, and auditing and enforcing compliance. Data governance is key to making sure data is used to achieve an organization's goals and follow its values.

Dataset

In AI, a dataset is a compilation of data used by an ML model, either for learning, or for processing. Typically, it includes numerous samples, each containing input values (also known as features or attributes) and an associated output value (also known as a label or class). For instance, if an algorithm is trained to recognize cats and dogs, the training dataset would feature images of cats and dogs, with labels indicating whether the image is a cat or a dog. This dataset is used for the algorithm to learn to recognize the characteristics of cats and dogs, so it can later make accurate predictions based on new images. The quality of the data and the size of a dataset are critical for the performance of an ML model. A large, high-quality dataset provides more, better data for the algorithm to learn from, leading to more accurate predictions. Conversely, a small or low-quality dataset may not provide enough information for the algorithm to learn effectively, resulting in poor performance.

Enterprise AI

Enterprise AI refers to the integration of AI technologies into an organization's workflows and processes. The deployment of AI has numerous benefits for businesses. It can:

- Automate manual tasks and increase efficiency, freeing up employees to focus on more high-level tasks.
- Provide businesses with valuable insights and data-driven decision making capabilities, improving decision-making accuracy and reducing the risk of human error.
- Help organizations stay ahead of competitors by offering innovative products and services, and enhancing customer experiences.
- Improve risk management and fraud detection, as well as streamline supply chain operations and production processes.
- Support employee training and development and drive innovation within an organization.

ETI ("Entreprises de Taille Intermédiaire")

In France, ETIs are medium-sized businesses with 250-5,000 employees and annual revenue between €50 million and €1.5 billion. ETIs are a vital part of the French economy, stimulating growth, innovation, and employment. They are distinct from small and large businesses: they often have a strong presence in both domestic and international markets, have a long-term vision, and are characterized by their ability to balance long-term investments with short-term profitability. ETIs are typically family-owned, giving them a degree of independence and agility compared to larger firms. The OECD does not have an equivalent category: ETIs are larger than OECD's SMEs, which are defined as having fewer than 250 employees, and either an annual turnover of less than €50 million or an annual balance sheet total of less than €43 million. Additionally, the French ETI definition includes the number of employees, while the OECD definition primarily looks at financial criteria. ETIs can be considered the "lower tier" of the OECD's large enterprise classification. It should be noted that definitions of SMEs, large enterprises, and ETIs vary between countries and regions, so it is vital to consult the local criteria before making any assumptions.

Explainability, explainable AI

Explainability in AI refers to the ability to understand and explain how an AI model makes decisions. There are several challenges that make explainability difficult or even impossible:

- Some AI models, such as deep neural networks, are considered "black-box" models, meaning they are difficult to interpret and understand.
- Certain AI models, like generative models, are not interpretable by design, meaning that it's hard or impossible to understand the reasoning behind the model's decisions.
- AI models can be fooled by input data specifically designed to trick the model, known as adversarial examples, making it hard to trust the explanation of the model's decision-making process without data integrity.
- AI models can make decisions based on complex interactions between multiple features, or even abstract features that humans are not able to interpret or understand.
- Explainability is a necessary condition for accountability because it allows for understanding how the model arrived at a decision and whether the decision is fair, accurate and unbiased. Without explainability, it is difficult to determine if the model is making decisions that align with ethical and legal standards and to identify any issues that may arise from its use. Explainability is an area of active research, and various techniques such as LIME, SHAP, etc. are being developed to improve the interpretability of AI models.

Fairness

In the context of ML, fairness refers to the absence of discrimination or bias in a model's decision-making process. A model should treat all individuals or groups in a similar way, regardless of their protected attributes (such as race, gender, age, etc.). This can be achieved by ensuring that the model's decisions are not influenced by irrelevant factors and that the model's performance is consistent across different groups of individuals. Factors driving fairness in an AI model include the data used to train the model, the model's architecture and algorithms, and the evaluation metrics used to measure the model's performance

Frugal AI

Frugal AI is a term employed to signify the development and utilization of AI in an economical and resource-efficient manner. This method involves the use of simple, inexpensive techniques and technologies to solve difficult problems, rather than complex and expensive solutions. The goal of frugal AI is to make AI accessible and affordable to a wider range of people and organizations, particularly in developing countries and resource-constrained settings. It often requires the utilization of simple algorithms, ML models trained on smaller datasets, low-cost hardware, and other resources like low-cost sensors and smartphones to build and deploy AI systems.

GDPR ("General Data Protection Regulation")

The GDPR is a comprehensive privacy regulation that came into effect on May 25, 2018, in the European Union (EU). It aims to protect the privacy rights of individuals in the EU by regulating how organizations collect, process, store, and share their personal data. Organizations that fail to comply with the GDPR can face significant fines and other penalties, the maximum penalty being 4% of an organization's global revenue or €20 million, whichever is greater. The GDPR has had a significant impact on organizations around the world, as many have had to update their data processing practices and policies to comply with the regulation. The GDPR has also set a precedent for other countries and regions to develop their own privacy regulations, such as the California Consumer Privacy Act (CCPA) in the United States.

Generative AI

Generative AI refers to a type of AI that uses statistical models to generate new data based on patterns it has learned from a training dataset. The data generated can be in the form of images, sounds, text, or other forms of media. The underlying technologies used in generative AI are primarily deep learning and ML, such as Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and Recurrent Neural Networks (RNNs). These algorithms train on large amounts of data and use complex mathematical models to learn the patterns and features of the data. Generative AI provides a way to generate new, diverse, and creative outputs based on patterns learned from existing data. For example, a generative AI model trained on images of faces could generate new, previously unseen faces. A generative AI model trained on musical sequences could create new, original compositions. In product design, generative AI can be used to generate alternative designs for a product based on constraints and design specifications. ChatGPT is a prime example of generative AI trained on a textual dataset.

GPT ("Generative Pre-trained Transformer")

Not to be confused with general purpose technologies, Generative Pre-trained Transformers are a type of language model that can generate human-like text using deep learning techniques. GPTs are built using a neural network architecture called a transformer, which is trained on large amounts of text data using an unsupervised learning approach. GPTs are powerful tools that can be used to create natural-sounding language that resembles human speech. They do this by analyzing vast amounts of text data, and then using that knowledge to generate new text. This technology has a wide range of potential business applications, from creating chatbots that can converse with customers to generating automated reports and summaries.

GPT ("General Purpose Technology")

Not to be mistaken for generative pre-trained transformers, GPT is an innovation that has the potential to revolutionize societies through its impact on economic and social structures. Examples of GPTs include the development of the steam engine in the 18th and 19th centuries, electricity, or the development of the internet in the 20th century. Because GPTs typically aren't the property of a single entity, and have characteristics that make them easy to master by many people, they have the ability to facilitate the development of new technologies, new applications, and even new industries. One of the key characteristics of a GPT is that it has the ability to generate positive feedback loops: as the technology becomes more prevalent and sophisticated, it drives further innovation and adoption. This is because a GPT creates new opportunities and demand for other technologies and industries that are able to take advantage of it. GPTs often have a network effect, meaning the value of the technology increases as more people use it. For example, the value of the telephone increases as more people have phones because it becomes easier to communicate with a larger number of people. AI is considered a GPT

GPAI ("Global Partnership on Artificial Intelligence")

The GPAI is a multilateral and multi-stakeholder initiative that was launched in 2020 by Canada, France, Germany, Italy, Japan, the United Kingdom, the United States, the European Union, Australia, India, Mexico, New Zealand, Singapore, and South Korea. Its goal is to promote responsible and human-centric development and use of AI by fostering international cooperation and collaboration among governments, industry, civil society, and academia. Its goal is to ensure that AI is developed and used in a way that benefits society as a whole. Key areas of focus for GPAI include:

- Ensuring the responsible use of AI.
- Promoting innovation and growth in the AI industry.
- Addressing global challenges such as climate change, healthcare, and the digital divide.
- Supporting the development of AI talent and education programs.

Green AI

Green AI, or sustainable AI, is the development and use of AI in a way that is environmentally friendly and sustainable. This approach involves minimizing the environmental impact of AI systems and aligning them with sustainability goals. Green AI practices include using renewable energy sources to power systems, creating energy-efficient algorithms and models, and using AI to address environmental challenges. Reducing the carbon footprint of AI systems is a key goal, achieved by using renewable energy, optimizing algorithms, and designing energy-efficient AI systems. Green AI is an emerging field that focuses on using AI sustainably and addressing climate change and environmental degradation.

ML ("Machine learning")

Machine learning is a type of AI that uses large datasets to train algorithms for improved performance on a specific task. This type of AI has the ability to "learn" from data, allowing it to make predictions or take actions without explicit programming. There are three main types of ML: supervised, unsupervised, and reinforcement. Supervised learning uses labeled datasets with correct outputs provided for a given input. Unsupervised learning uses unlabeled datasets and the algorithm must discover the underlying structure and patterns on its own. Reinforcement learning involves the algorithm learning by interacting with its environment and receiving rewards or punishments. ML is a powerful tool used in many applications, from computer vision and natural language processing to recommendation systems and autonomous vehicles.

MLOps ("Machine Learning Operations")

MLOps is a new set of practices that allows organizations to manage and optimize the end-to-end delivery and deployment of ML models. It involves collaboration between data scientists, AI experts, and IT operations teams to streamline the development, testing, and deployment of ML models in a scalable, secure, and reproducible manner.

ModelOpS ("Models Operations")

ModelOps is a specific subset of MLOps that focuses on the ongoing management and maintenance of ML models after deployment. This includes tasks such as versioning, documentation, monitoring of model performance, updating models with new data, managing retraining, and troubleshooting. If a business fails to implement proper MLOps and ModelOps, they risk confronting low model accuracy, slow deployment time, lack of reproducibility, and security concerns. In addition, poor model management can lead to a lack of transparency and accountability, potentially resulting in negative business consequences.

NLP ("Natural Language Processing")

NLP is a sub-field of AI that focuses on giving computers the ability to understand and generate human-like language. It involves a range of techniques, such as ML, linguistics, and computer science, to enable computers to comprehend, interpret, and create human language. Examples of natural language processing tasks include sentiment analysis (determining the emotional tone of text), part-of-speech tagging (identifying the parts of speech in a sentence), named entity recognition (identifying named entities such as people, organizations, and locations in text), machine translation (translating text from one language to another), and text summarization (creating a condensed version of a piece of text that captures its key points). NLP is an important, rapidly developing field with practical applications in areas such as chatbots, search engines, and machine translation.

Overfitting - underfitting

Overfitting happens when a model is excessively trained on the training data, resulting in the model capturing noise in the data and fitting the training data too precisely. This can lead to the model being too tailored to the training data, leading to poor generalization when applied to new, unseen data. A high training accuracy but a low test accuracy typically signals overfitting. In contrast, underfitting occurs when a model is too simplistic and fails to capture the underlying patterns in the data. This can result in poor performance on both the training data and new, unseen data. A low training accuracy and test accuracy typically indicate underfitting. To mitigate overfitting, various techniques such as regularization, early stopping, and data augmentation can be employed to prevent the model from overfitting the training data. To address underfitting, techniques such as increasing model complexity, adding more features, or increasing the amount of training data can be used.

Prediction

Predictions made by ML models are a crucial aspect of the technology's use in real-world business decisions. ML models process input data and output a value or class label, reflecting the relationships it has learned between inputs and outputs through training on a dataset. This output is referred to as a "prediction". This ability to make predictions on new, unseen data makes ML a powerful tool for a variety of tasks such as classification, regression, and forecasting. By analyzing vast amounts of data, ML models algorithms can identify patterns and relationships that would otherwise go unnoticed, providing valuable insights to organizations, informing decision-making and driving business outcomes, making ML a crucial tool for many businesses.

Reliability

Reliability, in the context of ML, is the ability of a given model to produce consistent and accurate results regardless of changes in the data used. The reliability of an ML model is a measure of its accuracy and consistency when making predictions on new data. Reliability and robustness are related in that both measure a model's performance, but differ in that reliability measures performance on expected data, while robustness measures how well a model performs on unexpected data.

Robustness

The robustness of an ML model is the ability of the model to produce reliable results even in the presence of unexpected or varying input data. It is different from reliability in that reliability measures the consistency of the model's output, while robustness measures the model's performance in the face of varying inputs.

Sentiment analysis

Sentiment analysis is a process of analyzing and identifying the emotional tone or attitude expressed in a piece of text. ML is used to train models on large amounts of labeled data, while lexicon-based methods rely on sentiment lexicons. Rule-based methods use predefined rules to identify sentiment, and hybrid methods combine two or more techniques to improve accuracy.

Sovereignty

Sovereignty is the power of a nation or government to control and govern data and AI technologies within its borders. This includes regulating and governing the collection, storage, use, and sharing of data, as well as the development and deployment of AI technologies. It also involves protecting citizens' rights and privacy in relation to data and AI in line with the basic functions of governments (law & order, military, and taxation). Nations may be concerned about sovereignty for multiple reasons, such as dependence on foreign technologies which constrain their ability to regulate these technologies, lack of data privacy laws, limited access to data, cybersecurity risks, and economic and strategic issues due to the increasing value of both AI and data. In Europe, sovereignty is also seen as a way to prevent data stored by tech giants from being sold back to the countries that produced it.

Synthetic data

Synthetic data is data that is artificially generated to resemble real-world data, rather than collected from real-world observations. This can be done using a variety of techniques, such as generative models or rule-based systems. The generation and use of synthetic data is a rapidly growing field of research with a variety of potential benefits for businesses, including the ability to generate large amounts of data quickly and efficiently, control and manipulate the characteristics of the data, and comply with data protection regulations such as the GDPR. One of the primary benefits of using synthetic data for a business is the ability to create and test new products and services without risking the privacy or security of real-world data. Synthetic data can be used to train ML models, test new algorithms, and improve the accuracy and robustness of data-driven systems, or augment real-world data, filling in gaps and creating more comprehensive datasets.

Transparency

Transparency of a ML model refers to the ability of the user to understand how the algorithm works, and how it reaches conclusions. This can include information on the data that was used to train the algorithm, the specific techniques and methods employed, and the decision-making process that the algorithm uses.
There are various ways to achieve transparency in ML models, including:
- Model interpretability: the ability to understand the relationships between the input and the output.
- Model visualization: the ability to visualize the model and its decision boundaries.
- Providing information on data: how data was used to train the algorithm.
- Providing documentation: providing users with information on how the algorithm made a decision.
- Providing feedback: providing information based on results to improve AI models and processes.

Transparency is a key building block of trust in AI systems.

Trust

Trust in AI is the confidence and belief in the AI system's ability to act in a reliable, transparent, and responsible manner. Trust is crucial as it affects how people and organizations interact with AI systems, and thus their potential for success. Factors that influence trust in AI include performance and accuracy, explainability, and ethical and social impact. To build trust, AI systems should be designed, developed, and implemented with transparency, explainability, and accountability. This includes providing reasons for decisions, auditing and monitoring, and engaging with stakeholders to understand their needs. Trust in AI is essential to its adoption and use, and has important ethical and social implications. Chapter 8 describes trust in AI systems in more detail.

Upskilling

Upskilling and reskilling refer to the process of training and development of employees to upgrade their skills and knowledge to align with the changing demands of an organization. Upskilling focuses on enhancing existing skills to adapt to new technologies, while reskilling involves acquiring entirely new skills to perform different tasks. Upskilling and reskilling are necessary for businesses adopting AI because the technology requires a new set of skills and expertise for effective implementation. Without these skills, businesses may struggle to make the most of their AI investments and may even hinder their ability to keep up with competitors who have successfully upskilled and reskilled their workforce.

APPENDICES
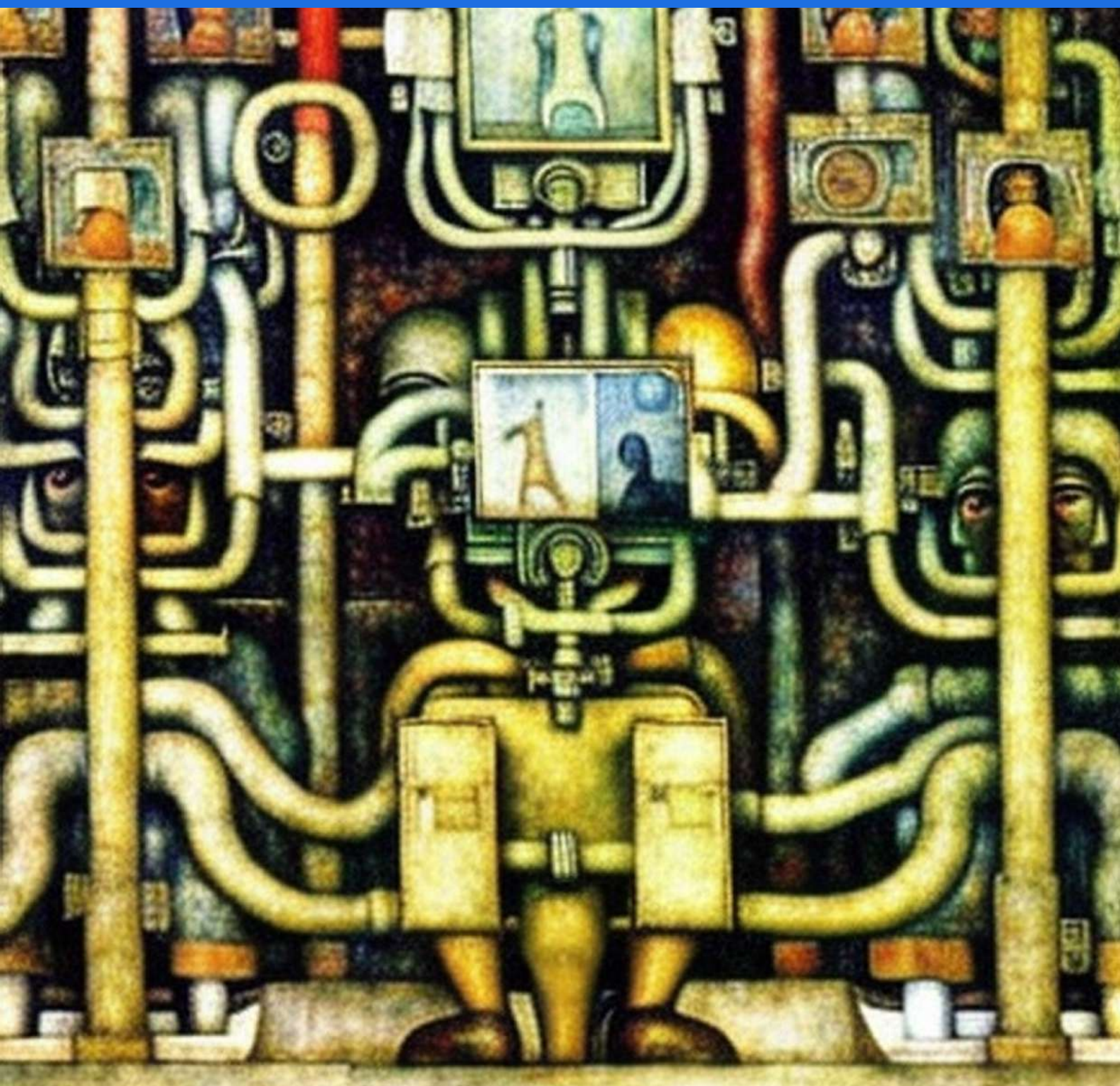
APPENDIX A - Public funding options/support: what is available and who qualifies?

Public funds can play a critical role in an Artificial Intelligence (AI) strategy and European, national, and regional initiatives can support businesses, particularly SMEs, in their efforts to innovate. Public funds can be utilized at two phases of the implementation of an AI strategy. At the very beginning of AI adoption, public funds can help to reduce the financial burden of testing the viability of AI in a business. This can include partial financing of the feasibility study and of the PoC that follows. After AI adoption, public support can finance the "innovation" component of a project, the phase that provides a competitive advantage specific to the company and can be protected by intellectual property laws.

Businesses considering an ambitious transformation project should identify and secure applicable public funding. Some public funding examples accessible to French SMEs include:

- **Pack IA:** The Ile-de-France regional government, for example, has formed a consortium led by the Hub France IA and the IMT Teralab that seeks to put AI at the service of the region's economic players. The consortium offers the Pack IA initiative to subsidize regional businesses, covering 50% of the cost of a company's first AI project over a period of three months.
- **Horizon Europe:** The European Commission offers thematic grants under the Horizon Europe program and innovation grants through its Research, Development and Innovation arm.
- **French national grants:** for developing innovative AI solutions for the competitiveness of the industry are available through national investment bank Bpifrance. R&D tax credit is also available for those SMEs who have hired engineers and PhDs to work on substantial improvements to the state of the art of AI.

Pack IA: the Acorus success story

Beneficiary: Acorus specializes in the renovation of occupied sites; social landlords are its primary clients. The company has five business lines: rehabilitation, refurbishment of housing, routine maintenance, and accessibility.
Staff: 1400 ppl - Turnover: €207M
Level of digital transformation progress: Advanced.
Challenges:
- Clean a product and material database of over 50,000 entries.
- Standardize labels created manually by multiple people.

Solutions:
- Natural language processing.
- Semi-automated creation of business ontology.
- Automated integration of new products in the database.

Before: Data entry recorded by multiple users with very heterogeneous names.
After: Entries registered automatically based on a precise ontology.

"Pack IA was a decisive element in our first experimentation with AI in our business. The preliminary framework and the proposal of a partner are key assets that facilitate the launch of a project by reducing uncertainties. In a few months, we were able to launch our first project and deploy it with an immediate business impact."

Christophe Chalvin

APPENDIX B - Regulatory environment

The AI Act: rules to better protect fundamental rights of EU citizens

In recent years, Artificial Intelligence (AI) has become increasingly popular, posing risks to individual rights and liberties. To take advantage of the benefits of AI while limiting its risks, the European Commission has proposed to set up a regulatory framework for trustworthy AI. After extensive dialogue among stakeholders, the European Commission published a proposal for a regulation on AI, the AI Act, in April 2021. This new regulatory framework will have a major impact on businesses, much like the General Data Protection Regulation (GDPR).

How is AI defined in the proposed AI Act?

The regulatory framework depends first on a clear definition of AI. The EU High-Level Working Group (HLEG) proposed a specific definition (1) in its 2020 white paper, taking into account AI's ability to act autonomously towards achieving a specific goal, whether software-based or embedded within hardware. Albeit quite detailed, it fails to recognize that AI is primarily software, and does not account for future technological improvements.The European Commission has decided to adopt a three-part definition.

According to Article 3 of the proposed AI-Act, AI system means:
- 1) "Software (...)"
- 2) "(...) that is developed with one or more of the techniques and approaches listed in Annex I".
- 3) "(...) and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

Due to its breadth, perceived as excessive by many parties, this definition has been subject to strong criticism. In November 2022, the Council of the EU circulated a compromise proposal that narrows the definition of AI to systems developed through machine learning and logic and knowledge-based approaches.

How to classify AI systems?

The European Commission is taking a risk-based approach that favors technological neutrality, classifying AI systems according to the risks they present to human health, safety and fundamental rights, rather than on the basis of the underlying technology or sector of use. This approach allows regulatory text to remain current and effective, regardless of any future technology developments. AI systems can be classified as (1) prohibited, (2) high-risk, (3) subject to manipulation risk, or (4) without particular risk. The Commission focuses on high-risk AI systems, which are subject to constraints. High-risk AI systems are those incorporated in safety products regulated by other EU acts, or those dedicated to a specific purpose listed in Annex III (education, employment, access to essential services, law enforcement).

1. *"(AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."*

How to determine if an AI system is subject to regulatory framework?

Whether an AI system is regulated by the proposed regulation can be determined by the following:
- AI systems used in the military, law enforcement, or judicial cooperation are outside the scope of the regulation.
- If the AI system corresponds to the definition outlined in the text, it should be allocated into one of the risk categories. AI systems that are prohibited cannot be used; high-risk AI systems entail significant obligations and AI systems with risk of manipulation must meet transparency requirements.
- The status of the company in the supply chain of the AI system will determine applicable obligations, depending on the stakeholder's qualification when putting products on the market or using AI systems with high-risk or risk of manipulation.

How to qualify the role of the stakeholders in the supply chain of the AI system?

The European Commission has regulations covering the entire supply chain of AI systems, applying to all providers and users in the EU or in third countries (authorized representatives will be required for entities outside the EU). Outputs produced by AI systems outside the EU are also covered, even if the data does not originate in the EU.

Several roles are defined in the EU proposed regulation:

- The **provider** is responsible for the most obligations, but other criteria must be taken into consideration, including whether the provider is the developer of the AI system, if they are distributing it under their own name and trademark, and whether it is being made available in the EU free of charge.
- The **importer** is an entity established in the EU markets/sells the AI system, while the distributor makes it available without altering its properties.
- The **distributor** does not provide or import AI systems, but rather makes them available for distribution or use in the European Union market, whether for a fee or for free. It is important to note that the properties of the AI system remain unchanged.
- The **user** is the company or individual that uses the AI system for their professional activity.

If any stakeholder meets the criteria for a provider or modifies the intended purpose of a high-risk AI system, they risk re-qualification. The role of the distributor, importer, user, or third party may evolve over time. If these parties meet the criteria of the provider, alter the intended purpose of a high-risk AI system, or make a substantial modification to such a system, they also risk re-qualification.

What are the main obligations for high-risk AI systems?

Obligations associated with high-risk systems include providing transparency in terms of the algorithms used, maintaining a high level of security and data management quality, and keeping relevant authorities informed throughout the AI system's lifecycle, spanning from design and verification to deployment, commercialization, and monitoring. To meet these obligations, companies must implement an effective governance system that involves multiple departments, such as risk and internal control, legal, information security, and operations.

The main obligations of each stakeholder are the following:

Provider
- Conducts premarket or precommissioning conformity assessment of the AI system and maintains post-market monitoring.
- Implements, documents, and maintains the quality management system.
- Draws up and maintains the technical documentation of the AI system demonstrating compliance with regulations.
- Keeps the logs automatically generated by the AI system.
- Draws up an EU Declaration of Conformity and registers the AI system in the EU database when required.
- Affixes the CE mark on AI systems and provides instructions for use.
- Must act in case of AI system non-compliance.

Distributor and importer
- Ensures the provider has established technical documentation and has carried out conformity assessment of high-risk AI systems.
- Verifies presence of CE marking and instructions for use.
- Refuses to place non-compliant systems on the market.
- Ensures that storage/transport conditions do not jeopardize the compliance of high-risk AI systems.
- Provides appropriate authorities with necessary information and documentation to demonstrate conformity.
- Must act on AI system non-compliance by implementing corrective actions while the importer informs the provider and the market surveillance authority if the AI system presents a risk at a national level, recording this information on the packaging or documentation of the AI system.

User
- Uses AI systems in accordance with instructions for use.
- Ensures that input data is relevant to the intended purpose of the AI system.
- Conducts monitoring of AI systems and informs the provider and distributor of suspected risks, serious incidents or malfunctions.
- Suspends or interrupts the use of AI systems in case of suspected risk.
- Keeps logs generated automatically by the AI systems.
- Complies with obligations under the GDPR if the user also acts as a controller

Focus on the conformity assessment
Providers of high-risk AI systems are obligated to perform a conformity assessment that assesses the risks of the system for the proposed project, demonstrating that the requirements for high-risk AI systems are met. This should be done before the system is placed into service, ideally during the testing and verification stage to check data quality and accuracy. To perform the assessment, the provider can choose between two procedures:
1. Internal checks by the provider, verifying that the quality management system, information in the technical documentation, and the design and development process for the system and its surveillance system conform with the regulation.
2. Evaluation of the quality management system and technical documentation by a third party (notified bodies), which will issue a certificate of conformity valid for 5 years.

High-risk AI systems incorporated into safety components are presumed to be in conformity and are already monitored and subject to conformity assessments under existing legislation.

The provider may not be able to anticipate all uses of the AI system by its customer, especially for standardized products marketed on a large scale. In this case, the provider can perform a general assessment and require customers to complete an assessment for their specific use.

Under the GDPR, the conformity assessment may overlap with the impact assessment, which is an obligation on the data controller, most often the user (not the provider) of an AI system.

What are the main obligations for AI systems with manipulation risks?

Chatbots, deep fakes, emotion recognition systems, or biometric categorisation systems, are subject to less onerous obligations compared to high-risk AI systems. These restrictions are imposed on the provider and the user and seek to ensure transparency to the end-user, making them aware that their interactions, images, and sounds are being processed by an AI system.

Individual persons should be informed as follows:
- Chatbots: the provider should ensure that systems are designed and developed in such a way that the end-user is aware they are interacting with an AI system, unless this is obvious from the context and circumstances of use.
- Emotion: recognition systems or biometric categorization systems: the user is responsible for informing natural persons exposed to the system about how the system works.
- Deep Fakes: the user should disclose that the content has been artificially generated or manipulated.

What are the risks in case of non-compliance?

Like the GDPR, the AI Act determines financial penalties depending on the severity of an infringement. Penalties have been designed to be effective, proportionate, and dissuasive, taking into account the economic viability of small suppliers and start-ups: either a maximum amount in Euros or a percentage of the total worldwide annual turnover for the preceding financial year.

What are the next steps of the legislative process?

The proposed AI Act is currently under discussion with the European Parliament and the Council of the EU acting as co-legislators. The Council adopted its common position ('general approach') on the AI Act on 16 December 2022, including amendments narrowing the scope of the regulation (e.g., definition of AI systems, prohibited AI-systems extended to private actors and social scoring, new exclusions related to law enforcement, new category of risk named 'general purpose AI'). The European Parliament committees have also adopted recommendations and drafted a report. Over 3,000 amendments which are still being discussed. Due to the slow progress of discussions, the final adoption of the text is expected to take place at the end of 2023 or the beginning of 2024.

APPENDIX C - Going beyond this white paper

Practical books (in French)

A few books (in French language) that may be useful to executives wishing to explore adoption further:
- **L'intelligence artificielle pour les nuls** (J.P. Mueller, L. Massaron)
- **L'intelligence artificielle en entreprise: Tout ce que vous devez savoir sur l'IA** (B. Marr)
- **L'intelligence artificielle expliquée à mon boss** (P. Blanc)
- **Guide pratique de l'intelligence artificielle** (S. Roder)

Useful strategy frameworks

Strategy frameworks have been developed that could be useful to SMEs. To cite just a few:
- **Bain & Co.'s "AI Maturity Assessment" Framework** is a step-by-step guide to assess a company's data and AI maturity, not unlike strategies.ai's "Data & Maturity Assessment" framework we chose to mention in chapter 3**.**
- **Strategies.ai's "Core vs. Peripheral AI Application" Framework** helps leadership understand where to apply AI based on criteria that include risk, cost and potential for value and ROI.
- **Strategies.ai's "Make or Buy" framework** sheds light on important considerations to pay attention to when deciding which resources to use for AI projects.
- **Emerj's "Bullseye" Model**, a clear framework for enterprise AI project selection.

Non-commercial communities and organizations in France

Interested readers can also get in touch with non-commercial organizations helping out their members with understanding AI and with onboarding businesses through peer exchanges and experience/expertise sharing. Joining one or more of these groups can allow you to stay informed and involved in current AI developments in France. Among them :

- **Hub France IA** is an organization dedicated to promoting and responsibly implementing AI in France. The Hub France IA runs multiple vertical (industry) and horizontal (function) workgroups such as banking, health, transportation, construction, cybersecurity, environment; ethics, education, regulation, HR, standardization, voice and language. Each workgroup brings together SMEs, large corporations, startups, and academics interested in collaborating on shared, large scale AI initiatives.
- The **Mouvement des Entreprises de Taille Intermédiaire (METI)** pushes for the development of a French Mittelstand. Its members reflect on issues related to AI adoption.
- **Numeum**, a federation of digital service companies, software publishers, platforms, and technology consulting firms, works on AI regulation issues.
- The **Association Française pour l'Intelligence Artificielle (AFIA)** is a research organization working on the scientific and technical questions of AI.
- **The Confédération des Petites et Moyennes Entreprises (CPME)** represents small and medium-sized businesses in France and works on AI regulation.
- **The Innovation Makers Alliance (IMA)** brings together over 6,000 strategic and operational French leaders in technological and digital innovation from 100 large businesses, mid-sized companies, and public administrations to further and develop innovative initiatives.

ACKNOWLEDGMENTS

The American Chamber of Commerce in France and ESSEC Business School would like to thank the following member companies for the time and dedication they put into this publication: **Airbus**, **American Express**, **AbbVie**, **Bain & Company**, **Digital Realty**, **DLA Piper**, **Google**, **Johnson & Johnson**, **La Compagnie, Nvidia**, **SAS**, **Schiller International University**, **The Walt Disney Company**, **Trellix, and the U.S Embassy in Paris.**

AmCham France is grateful to **Benoit Bergeret,** Executive Director of ESSEC Metalab, who raised the topic of responsible Artificial Intelligence as a key topic for businesses, led our working group discussions with conviction and perseverance and whose contribution was key in the writing of this white paper, together with **Julien Malaurent**, Academic Director of ESSEC Metlalab, and Professor of Information Systems at ESSEC Business School, whose experience working with corporations of all sizes was essential.

We are also grateful to the contributors who helped with the writing of this white paper with invaluable insights and comments, including:

- **Fabien Cros**, Data & Measurement Lead, Google
- **Jeanne Dauzier**, Counsel, DLA Piper
- **Walid Farhat**, Chief Customer Advisory Officer, SAS France
- **Thomas Fauvel**, Deep Tech Industry Expert, Choose Paris Region
- **Deborah Guidez**, Strategic Program Manager, ESSEC Business School
- **Denise Lebeau-Marianna**, Partner, DLA Piper
- **Emmanuel Pernot-Leplay**, Senior Consultant Data Protection & Privacy, Deloitte
- **Aude Prebay**, Paris Campus Director, Schiller International University
- **Jean-François Sebastian**, Managing Director, SAS France.

Special thanks to all the persons who supported this effort throughout the process, including AmCham France President, **Bijan Eghbal**, Partner at DLA Piper, **Vincenzo Esposito Vinzi**, ESSEC Business School Dean & President and AmCham Board member, and members of the AmCham team (**Eglé de Richemont**, Managing director, **Zachary Pascaud**, **Laura Delaunay**, **Zahira Benabdallah**, **Raphaël Bon**).

Zahira BENABDALLAH
Membership and Communications Manager, AmCham France
E-mail: zahira.benabdallah@amchamfrance.org

Laura DELAUNAY
Public Affairs & External Relations Manager, AmCham France
E-mail: laura.delaunay@amchamfrance.org

Deborah GUIDEZ
Program manager, ESSEC Metalab for data, technology and society
E-mail: deborah.guidez@essec.edu

**77 rue de Miromesnil
75008 Paris, France**

www.amchamfrance.org

+33 (0) 1 56 43 45 63