Data-driven business management

Extracting value from data across the enterprise

December 2023

Cigref

Data-driven business management

Extracting value from data across the enterprise

December 2023

EDITORIAL

In this digital age, data is often presented as the black gold of the 21st century, constituting inestimable wealth for humanity. However, unlike oil, whose value is widely recognised, data remains largely untapped in our businesses and public administrations. We can draw a parallel between natural resources and data: while oil deposits are diminishing, their exploitation is already well established. On the other hand, the amount of exploitable data is constantly growing, and yet it is only partially transformed into products or services. This reveals an immense potential for creating value for our businesses and public administrations.

Certain tools, such as generative artificial intelligence, recently made available to the general public, are contributing to the growing awareness of the importance of making better use of the data produced by the digitisation of human activities. They also demonstrate the extraordinary power they confer on all those who decide to take advantage of them. Provided they have these tools, adequate computing resources and access to a growing amount of data, every individual and every employee will soon be able to make better decisions within their organisation. It is therefore in the interests of every company and every public authority to harness this power and direct it towards their strategic objectives, in order to derive maximum value from it and control its risks.

But to take full advantage of this windfall, large companies and public administrations alike need to organise themselves around some simple but fundamental principles. Firstly, each business line, embodied by identified managers, must be given responsibility for using and sharing its data, and must be provided with the resources needed to manage it effectively. Secondly, it is essential to facilitate access to data for each player involved, using appropriate tools. Finally, it is essential to establish and impose common rules for the use of data throughout the organisation, to ensure that it is used responsibly and ethically. Data can then be considered and managed as a product of the organisation in its own right, in the same way as its traditional products. This approach fosters a data-centric culture and encourages informed decision-making, in line with strategic objectives.

However, implementing these principles requires a profound transformation in the way organisations organise themselves and the way they work. This report is therefore dedicated to exploring this transformation, addressing its principles, its obstacles and the best practices for overcoming the challenges that lie ahead.

In conclusion, our report aims to raise awareness of the enormous potential for value creation that data represents for our businesses and public administrations. Only by adopting a responsible, ethical and collaborative approach can we truly harness this invaluable resource for the good of all. Enjoy your reading, and bon voyage to the heart of data-driven enterprise management!

Julie POZZI, AF KLM and **Jean-Claude LAROCHE**, ENEDIS

SUMMARY

Today, the use of data remains rather compartmentalised within traditional business lines, while data flows are more often than not cross-functional. However, putting data at the heart of decision-making processes is increasingly meeting the strategic and business challenges of companies and public administrations. For an organisation, this data-driven approach represents a transformation in which all teams have a role to play, as does the executive committee. The executive committee needs to be convinced of the benefits of such a transformation, by demonstrating the value it will create, given the time and organisation required to implement it, and the associated cost. The first part of this report lists the various benefits sought by the executive committee and suggests best practices for validating the benefits of such a data transformation.

Furthermore, the fact that data is rarely shared within a company or public authority due to a lack of trust is even more obvious within its ecosystem (customers, suppliers, partners, vendors). The aim of setting up a data marketplace is to provide a single place to make all data available, whether internal or external to the organisation, so that it can be accessed, shared and re-used.

The Gaia-X association is establishing a framework of trust, with transparent, verifiable and interoperable data infrastructures, to meet the challenge of developing the data economy by facilitating its exchange, as specified in the European data strategy of February 2020. Gaia-X's *raison d'être* is to offer an ecosystem in which data exchanges are based on verifiable and standardised contractual rules that enable scale-up, confidence-building and development based on interoperable and portable architectures.

In order to achieve data-driven management, organisations are transforming themselves based on the **4 fundamental pillars** of the data mesh, which represent the ultimate target to be reached. These 4 pillars are:

1 Organising data into domains;

2 Data as a product;

3 Federated data platform;

4 Federated governance.

For data to become a transformative factor, all employees must be encouraged to integrate it into their reasoning and decision-making. To do this, it is essential to get them on board, train them and get them acculturated, because this is a long-term effort corresponding to a cultural change at the level of the entire company/public administration. What's more, making the right decisions systematically and confidently, in a responsible manner, also means having the right organisation in place, and sharing high-quality data. AI tools, and in particular generative AI tools, which are increasingly being developed in organisations, will reveal all the potential offered by data.

ACKNOWLEDGEMENTS

Our thanks go to **Julie POZZI**, IT Director of Operations Research at **AIR FRANCE-KLM** and **Jean-Claude LAROCHE**, Mission Director at **ENEDIS**, who steered this work, as well as to all the people who took part in and contributed to this Cigref working group (in alphabetical order):

Olivier ACCOLAS - GROUPE EGIS
Franck ATGIE - ENEDIS
Patrick BALDIT - CEA
Laurent BASTIDE - MINISTÈRES SOCIAUX
Arnaud BEAUFORT - MINISTÈRES TRANSITION ÉCOLOGIQUE COHÉSION DES TERRITOIRES TRANSITION ÉNERGÉTIQUE MER
Yannick BELTRAN - KERING
Fahem BEN MESSAOUD - ESSILORLUXOTTICA
Frédéric BERNARD-PAYEN - AIRBUS
Christelle BOURDEU - AIRBUS
Michel CAPEL - FAYAT
Didier CAZETTES - TOTALENERGIES
Evagelos CHAIDACIS - TOTALENERGIES
Thierry CHAMPEROUX - MAIF
Claudio CIMELLI - MINISTÈRE DE L'ÉDUCATION NATIONALE ET DE LA JEUNESSE
Benjamin CLISSON - PLASTIC OMNIUM
Simon CRASE - GIP MDS
Pascal DALLA TORRE - VEOLIA
Franck DALLA-SARTORA - ENEDIS
Henry DE FAGET - ALLIANCE AUTOMOTIVE GROUP
Olivier DJOMBY - ESSILORLUXOTTICA
Harouna DRAMÉ - PLASTIC OMNIUM
Eugenio FERRARIO - LISI GROUP
Sandra FRANCISCO - ACOSS
Xavier FRIGOLA - GROUPE SAVENCIA
Matthieu GARCIA - LABORATOIRES PIERRE FABRE
Nicolas GOUR – WORLDLINE
Kévin GUEZ - ICADE
Stéphane GUILLOT - VALLOUREC
Sandra HOLGADO - ENEDIS
Jean HUOT - VINCI
Mathilde JAMEAU - Rémy Cointreau
Cédric JUBLOT - EIFFAGE
Jean-Christophe LALANNE - AIR FRANCE-KLM
Jean-Claude LAROCHE - ENEDIS
Raynald LASOTA - FRANCE TELEVISIONS Pierre-Emmanuel LE PARCO - AG2R LA MONDIALE
Frédéric LEBOEUF - VINCI Patrick MAHU - POLE EMPLOI Romain MELET - FRANCE-KLM Didier NOier - AIR FRANCE-KLM
FRANCE TELEVISIONS
Pierre-Emmanuel LE PARCO - AG2R LA MONDIALE
Frédéric LEBOEUF - VINCI
Patrick MAHU - POLE EMPLOI
Romain MELET - AIR FRANCE-KLM
Didier NOWAK - HAGER GROUP
François-Mary PERSYN - POLE EMPLOI
Eric PINAULT - PLASTIC OMNIUM
Julie POZZI - AIR FRANCE-KLM
Daniel SCHAERER - SFR
Romain SOULET - LABORATOIRES PIERRE FABRE
Arthur STANGRET - ENEDIS
Frédéric SUTTER - AIRBUS
Louis-Cyrille TRÉBUCHET - FRANCE TELEVISIONS
Sophie TROEL - VEOLIA Max VALLEJOY - VEOLIA VEOLIA
Max VALLEJO - SOCIÉTÉ GÉNÉRALE
Claude VARY - GROUPE ADP
Arnaud VASSEROT - GIP MDS
Taras VELIKOROUSSOV - ORANGE
Laurent VERHOEST - TRANSDEV GROUP
Florent VERRIERE - EDF
Dominique VIENNE - ENEDIS

We would also like to extend our warmest thanks to all those who contributed to the work of our working group:

- Mourad BOUALLAK, Data Division Project Manager at LA POSTE;
- Eugenio FERRARIO, Head of Data at LISI Aerospace;
- Matthieu GARCIA, Company Architect at LABORATOIRES PIERRE FABRE;
- Mick LEVY, Director of Business Innovation at BUSINESS & DECISION (GROUPE ORANGE BUSINESS);
- Julie POZZI, IT Director of Operations Research at AIR FRANCE KLM;
- Grégoire SAISON, Managing Director Tech & Data Advisory at SIA Partners;
- Taras VELIKOROUSSOV, IT&N Financial & Technical Performance director at ORANGE.

This document was compiled and written by Marine DE SURY, Mission Director at Cigref.

TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS

TABLE OF INSERTS

1 INTRODUCTION

Companies and public authorities are convinced of the value of data. Data is an asset for the whole company, not just for one department or business unit. That's why organisations are transforming themselves to make better use of this data, and investing in the construction and implementation of a data strategy to support their overall strategy. See the Cigref report "*Governance and data & analytics architecture: Developing and implementing a data strategy*", published in January 2023.

However, even if companies and public administrations are increasingly taking advantage of data, this development is still confined to certain teams, certain departments or certain business lines. Indeed, very few companies are using data as a lever for reorganising their macro-work processes. Most often, this data remains compartmentalised within traditional business lines, while data flows are more often than not cross-functional. However, companies and public authorities are increasingly looking to place data at the heart of their decision-making (in other words, to become "data-driven"). Given the digitisation of most of their activities, data is becoming a major lever for customer satisfaction, internal performance and the attractiveness of the employer brand. The forty or so members of the working group identified a number of challenges to be met in order to achieve these objectives.

This Cigref report is aimed at all company employees. Its aim is to provide answers and suggest best practices for making data a lever for steering and transforming the entire company or public administration. Before embarking on such a process, it is essential to ensure that it meets a business and strategic challenge with clearly identified benefits, which is a prerequisite for convincing the executive committee and ensuring the success of the transformation. Next, the company needs to organise itself accordingly, anticipate the changes brought about by data and artificial intelligence (AI) tools, and respond to the human resources challenge, because this is a global transformation of the company and not a purely technological issue.

2 DATA-DRIVEN ENTERPRISE: A RESPONSE TO THE BUSINESS CHALLENGE

The organisations taking part in the working group have clearly understood that their transformation to put data at the heart of their processes and decisions only makes sense if it responds to strategic and business challenges, or, in the case of public authorities, to their service offerings. It is a way of maximising benefits and protecting against certain risks. This business or service offering approach requires an overall approach that leads to the use of the data produced and its value for all business activities. This approach is not, therefore, a technological issue, but one of business/public administration transformation, in which all teams have a role to play. This is why, in the remainder of this report, we will use the expression “data transformation”.

Data is present in most businesses and teams, and on a wide variety of subjects. This is why data transformation must be based on certain key success factors:

- **Involving the executive committee** and **Management Boards:** This presupposes that data transformation is seen as a new strategic area or as a source of value serving the main challenges facing the company/public administration.
- **Implement cross-functional projects for teams** under the responsibility of the business lines: data transformation is only worthwhile if it enables the full value to be extracted from the cross-fertilisation of data of various types and origins, both inside and outside the organisation, and if this exploitation of the data is placed under the responsibility of the business lines, who are responsible for demonstrating its value.
- Sharing and giving **visibility to all data projects** so that they can be **prioritised** and the associated IT **budgets protected**: the IT budgets that enable data transformation must not compete with the budgets traditionally used to develop or maintain transactional information systems.

2.1 GAINS SOUGHT BY THE EXECUTIVE COMMITTEE IN DATA TRANSFORMATION

There are several possible triggers for this transformation. Firstly, some companies or public authorities decide to carry out a data transformation in order to **improve** and **optimise** their existing business or service offering. This means using the data available internally to increase turnover, improve operational efficiency and customer satisfaction/experience, and reduce costs. Some business lines have enhanced their application with data from outside their business line, and sometimes even outside the company, as illustrated by the example of Air France KLM below.

Two examples of improvement at Air France-KLM using data from their ecosystem

Predicting the right number of meal trays and consumables on board in our aircraft is crucial to limiting our environmental impact while ensuring customer satisfaction: to do this, our algorithms use all the relevant data to deliver daily predictions for all our flights.

On another level, revenue prediction is essential for optimising an airline's flight schedule. External data sources exist and enable us to enrich our market share models and optimise our offer accordingly.

Julie POZZI, AF KLM

Data also provides an opportunity **to explore new sources of value through innovation**, and to increase sales through new offers, new markets, etc. This approach can be taken initially within the organisation, then through partnerships with its ecosystem - suppliers, customers, vendors, partners, etc. to create new types of services. However, one company emphasises the difficulty of creating these new services with external partners when the business case is uncertain and does not seem promising enough.

Creating new business: an example from Air France KLM

AF KLM's aeronautical maintenance business not only provides repair services for aircraft and their components, but is also positioned in the market for AI solutions to support aircraft operations, in particular with its predictive maintenance offering, which includes an AI component combined with aeronautical expertise.

Julie POZZI, AF KLM

Generally speaking, data transformation enables **practices** to **be made agile** throughout the organisation and to be **aligned to** better meet the organisation's strategy in the service of its customers or beneficiaries. It ensures that the company or public authority can be **managed using reliable and relevant indicators** to rapidly produce the insights/indicators required, in other words becoming “data-driven”. Organisations are currently investing heavily in ensuring that their data is of high quality and accessible via a single, scalable platform. Companies are sometimes organised by region or by business sector, and they generally want to have a **global, consolidated view** by determining the indicators to be considered centrally. In this context, it is important to share and give visibility to all data projects. This enables them to be prioritised and the associated IT budgets to be protected.

Data transformation can also be aimed at **upgrading a** complex legacy **information system that** is slow to change or not very scalable. In this case, it enables better use to be made of services and market potential such as generative AI tools, the cloud, open data, etc.

Another factor triggering **a** data transformation is the need to **meet regulatory requirements**, or to be able to **anticipate them** by using data.

Finally, data transformation helps **to improve** the **image** and reputation of the organisation. It **attracts** and **retains talent** by giving them the desire to develop internally, thereby training a future generation of data-driven business decision-makers.

2.2 MARKET TRENDS

SIA Partners has shared the trends identified by a panel of a dozen international companies from a range of business sectors with regard to their data transformation. The conclusions are as follows.

Most of the panelists have already embarked on a transformation of their organisation to place their data at the heart of their decision-making. At least one panellist is already data-driven: it has integrated data into all aspects of its business, from strategy to operations, and has optimised its business model to derive maximum value from data. A small minority of panellists are only just beginning to evolve with a few sporadic data use case initiatives. The others have already been engaged in this profound transformation for 2-3 years. Some have adapted their enterprise architecture so that they can actively use their data to provide the right tools for understanding and managing their business. These companies have a well-established data culture. In these cases, tools, methods and services have been established throughout the organisation, and initiatives are coordinated across the board.

Finally, a small half of the panel are beginning to collect and manage their data, but are still struggling to use it to make informed decisions, as departments/teams are not yet able to coordinate their data initiatives properly.

We can see that the data transformations of the most advanced companies have been either by regulatory constraints (e.g. the banking sector), or by the possibility of using data directly in their business (e.g. insurance).

Three major levers were identified as stimulating or accelerating data transformation:

- A focus on value: companies that are well advanced in their data transformation are investing massively in use cases with high added value for the business (e.g. innovations, new business opportunities) and are concentrating their resources on projects with the greatest impact. In the most mature companies, the budget allocated to data use cases is around 2.8 times higher than in less mature companies. Finally, the “data champions” have put in place an **agile qualification process to prioritise according to the value provided and the effort involved**.
- A Human Resources strategy adapted to the data transformation: companies that are ahead of the game in their data transformation have generally set up a skills repository and are able to **qualify recruitment needs and plan them according to** their maturity in each use case. They are also putting in place data training programmes to ensure the loyalty of employees specialising in data.

- A partnership strategy: These companies build and orchestrate a network with their third parties to make them real partners and not just simple IT service providers. Finally, these champions rely on cloud services and/or data platforms to accelerate their transformation.

2.3 CONVINCING THE EXECUTIVE COMMITTEE TO IMPLEMENT A SUSTAINABLE DATA TRANSFORMATION

Companies and public authorities are always asking themselves whether it makes sense to undertake a data transformation, because of the time it will take, the organisation required and the associated cost. Is the value derived from predictive maintenance, or from reducing the environmental footprint, to take specific examples, sufficient to initiate and complete the transformation? To convince the executive committee and the business units, you ideally need to be able to **present "cross-silo" use cases, associated with ROI**[1] and proceed in stages.

2.3.1 RAISING AWARENESS

If the executive committee is to become a driving and lasting sponsor of the data transformation, it is important, starting with its maturity on data issues (hence the importance of assessing it beforehand), to raise its awareness with verbatims, feedback and a benchmark presenting what other companies in the same business sector or other companies of the same size are doing. An immersive approach is sometimes proposed to get members personally involved in the various data-related issues, in a fun and engaging way.

2.3.2 SETTING THE AMBITIONS FOR YOUR DATA STRATEGY

Data strategy is part of the organisation's strategy, whether in terms of operational performance or growth. It provides information based on tangible facts, not on beliefs or feelings. It helps to measure the business benefits and the value provided by the use of data. However, it is important to be clear about the ambitions and expectations of the data strategy.

[1] ROI: Return of Investment.

Feedback on the positioning of data within URSSAF

The new management and objectives Agreement (COG in French) for 2023-2027 and the associated IT Master Plan position data as a key strategic pillar in meeting URSSAF's challenges. The programmes relating to customer relations, the fight against fraud and improving reliability are based on the wealth of data in our IT.

The previous COG has already enabled us to implement structuring business repositories: administrative data, contact data and bank details. As part of a data governance approach, the centralisation, documentation and reliability stages have been launched. The cross-referencing of data, via Big Data, to meet the challenges of operating the DSN (Nominative Social Declaration), is now bringing even greater performance to our businesses.

Sandra FRANCISCO, URSSAF Caisse nationale

2.3.3 DRAWING UP A BUSINESS CASE

Once the data ambition has been set, the business case for data transformation needs to validate the approach. To do this, we need to identify the cost drivers and associated benefits, and formalise the data use cases that will serve the organisation's strategic ambition. Even if the “sacrosanct” ROI is difficult to establish, an effort to formalise must be undertaken with the help of the finance and management control teams. The aim is to model both the costs and the expected gains, in line with the use cases, in order to convince the executive committee of the benefits of the transformation. The involvement of management control skills should be a regular feature of the use case cycle, to enable objective arbitration, feedback and improvement of the information feeding into the overall business case.

Finally, some organisations strongly recommend adopting an OKR[2] approach , which enables decision-making procedures to be improved in line with strategy, by monitoring results and making commitments around measurable objectives. So, if an initiative does not feed into the OKRs, it is de-prioritised. Resources are allocated according to factual indicators and criteria and in an agile way. In this case, OKRs ensure:

- Vertical alignment of teams (from the executive committee to operations) in line with the company's strategic objectives;
- Horizontal alignment, as objectives become common. This breaks down the silos between entities;
- Ongoing prioritisation to guide actions.

[2] OKR: Objective and Key Results.

2.4 FROM DATA PLATFORMS TO DATA SPACES

Until now, data has not been shared much because of a lack of trust. This is already true within a single organisation, whether a company or a public authority, but it is even more true between two organisations or within an ecosystem. The first part of this chapter presents the recommended actions to meet this need for trust when implementing data marketplaces (or data platforms) in an organisation. In the second part, we will look at the implementation, as part of Gaia-X, of data spaces in which data can be made available, collected, pooled and shared in complete security and trust. These data spaces also offer analytical, computational and AI capabilities and foster innovation.

2.4.1 DATA MARKETPLACE

Companies are transforming themselves to put data at the heart of their decision-making processes and become data-driven businesses. The challenge is to be able to make the right decisions systematically, in order to modernise, optimise and improve the performance of all the building blocks in the company's value chain, but to do so responsibly and rationally, on the basis of the common asset that is the company's data. The data marketplace meets this need.

The aim of the data marketplace is to make **all the company**'s internal and external **data** available in **a single place**, so that it can be accessed, shared and reused. To this end, stored data is put into context so that it can be consulted. The data marketplace makes it easier to compare and combine different sets of data, enabling higher added-value analyses to be generated. It provides a technical basis for sophisticated processing of large volumes of data.

Companies in the process of implementing a data marketplace have often started by **assessing their maturity.** In fact, the creation of a data marketplace generally accompanies the overall data transformation, and requires an assessment of the current situation.

Maturity assessment, the first stage in the implementation of Pierre Fabre Laboratories' data marketplace

The data maturity assessment was carried out by major business line with the help of a consultancy firm, and the diagnosis was structured around the 7 areas listed below:

- **Strategy**: the level at which the company's strategy is translated into a data analysis strategy, reflecting the company's clearly defined priorities and resources;
- **Organisation and talent**: level of sponsorship, available analytical skills, clarity of governance and interaction with the business;
- **Agility**: evaluation of the agile methodology applied to data analytics, risk appetite and internal collaboration;
- **Technology**: assessment of IT strategy, data analytics platform, cloud solutions and security solutions;
- **Data**: analysis of data quality, availability and governance;
- **Analytics**: study of the ability to generate insights (in other words, a good level of decision after a good analysis) and of the levels of sophistication of analytical models;
- **Adoption and scaling**: Ability to go from "demonstrator" to "product" used by the business. Integrating best practice into decision-making (Data Analytics value).

A small, dedicated team was then set up with the aim of developing employees in these areas to support the Group's transformation.

Matthieu GARCIA, Pierre Fabre Laboratories

The data platform (data marketplace) generally **comprises** a **data portal aimed at** the end user, a **data catalogue** including the lineage[3] **, analytics, AI** and **calculation tools** and, finally, **storage**. A number of key principles have been put forward in the administration of business and government marketplaces:

- The data produced or acquired within the company takes on its full value when it is put to use.
- The teams producing the data are the sole authority in terms of knowledge and expertise concerning the data. It is **their responsibility to share it** through this marketplace.
- Data without context has no value. Contextualisation is made possible by a set of metadata that describe the data and specify its use.
- All employees are involved in the production, acquisition and use of data, so the data marketplace will develop and prosper thanks to a **collective contribution**.

The capabilities of the data marketplace are varied. Data warehousing, together with modelling and cataloguing, ensures that all types and sources of data (structured, semi-structured and unstructured) are integrated and exposed, and that their day-to-day use is accelerated. The aim is to make data accessible to employees by means of simple queries. Virtualisation of all data sources is a useful

[3] Lineage provides information about the origin of the data and its transformation.

technique for implementing the marketplace, ensuring that data is never recopied, replicated or relocated. As far as possible, data should be available in real time.

Ideally, the data marketplace offers managed services such as calculations, regressions, various types of simple algorithms and AI tools. Finally, analytical services are also available to produce dashboards, reports or forecasts linked to the capabilities of the data platform. In the data marketplace, the dataset product sheet includes a number of elements: the name of the product, its description with metadata, the name of its owner, the business glossary, information relating to the context (for example: limits and constraints of the data, history of the data, existence of personal data), but also the topicality of this dataset, its rights, and the quality of each "data" product assessed by its owner. The number of views is also mentioned. The responsibilities assumed by the data owner[4] and the data set owner have been defined according to use and specified in the governance. Sometimes an attribute does not have to be public, so the data set is not shared.

The pillars of Pierre Fabre Laboratories' data marketplace

The Pierre Fabre Laboratories data marketplace is built around 5 pillars:

- **Data stewardship**: data that is poorly informed, erroneous or not monitored is a risk in its use (distorted decision-making);
- **Data marketplace**: for data to be of value, it needs to be cross-referenced, documented and shared;
- **Securing the roles and responsibilities of all** data **stakeholders**: this is vital if we are to regain the trust that has been lost in the use of data. A great deal of work needs to be done in terms of mapping and identifying the right people and skills to ensure the reliability of the system;
- **Data projects**: it is by delivering analyses based on data to inform decision-making that Pierre Fabre laboratories will be able to demonstrate their potential for creating value. For example, before/after the use of data to optimise field rounds;
- **Knowledge and control of data**: it is the employees who will activate the data or not. To exploit the full potential of data, the challenge is to develop a data culture in the way we think and in the way we do things.

Matthieu GARCIA, Pierre Fabre Laboratories

2.4.2 DATA SPACES

Established in 2021, the private not-for-profit association Gaia-X brings together a community of more than 350 members around a single objective: to stimulate the European data economy by enabling the

[4] The Data scientist, Data analyst, Chief data officer, Data engineer and Data Protection Officer (DPO) job descriptions are defined in Cigref's IT job profiles nomenclature. The working group has worked on the data steward profile and has proposed a definition of missions and tasks, which is attached. This sheet will be completed and validated this year so that it can be added to the next version of the Cigref job nomenclature. However, several positions mentioned in this report remain to be defined so that they can also be included. This will only be possible if European organisations are sufficiently mature to find a consensus definition.

creation of “commons”, frameworks, rules, labels and frameworks, for data spaces, in line with the February 2020 “EU Data Strategy”.

Gaia-X's value proposition is to provide a framework enabling data producers, consumers and users to regain control over the use of data and the underlying technology, with **verifiable contractual rules** for accessing and using the data, as well as for using services to analyse and process it. In other words, Gaia-X aims to facilitate the exchange of data by **guaranteeing its protection and security with interoperable and portable services**.

This is why Gaia-X began by characterising **data exchanges** for each data space with contractual, described, verifiable and standardised rules. The Gaia-X trust framework was then established, bringing together all the rules defining the minimum basis required to be part of the Gaia-X ecosystem. These rules guarantee common governance and the basic levels of interoperability between participants in the Gaia-X ecosystem, while leaving users entirely in control of how their data is used.

Finally, Gaia-X has defined and specified **three levels of labels** that guarantee data protection, transparency, security, portability and flexibility for the ecosystem, as well as monitoring at European level. All the cloud providers' solutions are characterised in several catalogues federated according to this label. These catalogues enable users of digital services to compare the different service offerings, because the services are described using the same criteria from one supplier to another, including the respective trust indicators. The Gaia-X catalogue can therefore be used to search, find and select cloud services according to various criteria: Gaia-X Label 1, 2 or 3, type of service, location, certifications.

A number of business sectors, presented on the Hub France Gaia-X website, are currently implementing use cases to demonstrate the value that can actually be derived from the Gaia-X ecosystem. They already include education and skills, energy, mobility and tourism, automotive, aerospace, agriculture, cultural and creative industries, infrastructure and construction, finance, healthcare and industry 4.0, among others.

The association has selected and recognised several industrial or public flagship projects, called lighthouse projects, for their relevance and significant contribution to the development and implementation of the Gaia-X ecosystem. These projects enable the association to understand users' needs, test their requests for developments and validate their industrial implementation.

Example of a Gaia-X lighthouse project for mobility and tourism: EONA-X

EONA-X is one of Gaia-X's lighthouse projects in the field of mobility and tourism (airlines, SNCF, Renault, Apidae, ADP Group, Amadeus, etc.).

One of the current use cases concerns the Paris Olympic Games. The aim is to coordinate and optimise the journey of the 85,000 Olympic Games delegates on their arrival (from the airport to their place of residence over a period of around 15 days) and on their departure (over a period of 3 days). Passengers will be transported by 1,000 buses and light vehicles, and luggage will be transported by lorry. The challenge is to avoid congestion in the vicinity of the airport by massive influxes of vehicles that would hinder normal passenger traffic, and to provide an impeccable welcome and departure for the delegations in order to promote a positive image of France among its emblematic ambassadors, the athletes, journalists and members of the OCOG (Organising Committee for the Olympic Games). A large number of players (public transport, airports, stations, hotels, the Ministry of the Interior, etc.) will be sharing their information thanks to EONA-X (position of buses and lorries in real time, location of delegations within the airport, check-in of deported luggage), which will also feed a digital twin to anticipate problems and find solutions upstream.

This use case is based on the EONA-X association. The founding companies of this association are currently defining the platform's standards and the governance to be put in place. Via this platform, the EONA-X partners offer a catalogue of their data sets. Consumers of this data, who undertake to respect the conditions defined by the owners, can create new services for their users/customers, and optimise operations already in place. Information is exchanged using connectors that can be implemented by the parties involved. Once the negotiation of the exchange contract has been finalised between the parties, the data is exchanged directly from producer to consumer, which both guarantees sovereignty and enables scaling up. The platform's business model is currently being defined: payment based on consumption of the platform's services, or a flat-rate subscription or membership fee.

EONA-X is working on use cases that are generally beneficial to several players on the platform. Real-time data exchange is one of the challenges. For the most advanced use cases, ad hoc governance will be required, for example in the case of a single transport ticket. It is up to the supplier of a dataset to define the conditions for its use, for example a licence, the conditions for free use or those associated with the price offered. Whatever the case of use, the producer of the data remains responsible for the quality and freshness of the information provided.

Charles TELITSINE, ADP Group

To date, Gaia-X is the only initiative that addresses this issue of trust in data sharing, and is aligned with the European strategy. Gaia-X is an association, and its development is the fruit of the collective labours of its members and its team of permanent staff. It is therefore important that businesses and public administrations play an active part in it, so that the Gaia-X infrastructure and framework respond to the challenges they face, as well as to European values.

3 DATA-DRIVEN STEERING: PILLARS AND TRAJECTORIES

Data is a real asset for organisations that know how to exploit it. What's more, it provides factual, rational and objective information for decision-making. However, there are a number of challenges that need to be addressed when transforming data across the enterprise.

The first is to be able to **use data on a company-wide scale**, and even more so on the scale of its ecosystem, which includes suppliers, customers, vendors, partners and so on.

The second challenge is to **facilitate cross-functional data flows between different teams** by arbitrating debates on data ownership and helping to prioritise products or services at scale. For example, customer data can be enriched with operational data to enhance the user/customer experience. The aim is to break down silos and take a "product" approach to data in all its dimensions, right through to regulatory compliance. As we shall see later, when we define this term, the term "data product" enables data to be positioned in relation to all the company's internal and external "customers", and avoids unnecessary duplication. In terms of the principles of data transformation, it should be emphasised that it is not IT's job to manage the data projects of the business lines. A "data product" serves the needs of the business and must therefore be under the responsibility of a business team, which is responsible for the quality of the data delivered for its own activity, or more broadly within the company, as well as the performance of the data processing model. The "data product" is supported by the "data product owner"[5]. Resources need to be allocated within the business to ensure that the model is supervised and evolves over time.

The 3rd challenge is to find **a transformation path that allows the company to progress incrementally in terms of its organisational model and level of maturity for processing data at scale**. The "data mesh" type of organisation, which we will present in this chapter, is one way of meeting this challenge. The data mesh is widely considered to be the target model for companies/public administrations that want to place their data at the heart of their decision-making. However, the idea is not so much to follow the data mesh model in its entirety as to adapt it to the specific characteristics of your organisation. Each transformation of the company in this direction already represents an improvement.

3.1 CURRENT SITUATION

There are three main ways in which data and AI are used in an organisation.

- **Data used to manage**. Data is used to build dashboards and KPIs[6] , and to draw up reports.

[5] As soon as this job has matured, i.e. as soon as European companies have reached a consensus on its missions, tasks, deliverables and KPIs for measuring its performance, it will be included in the Cigref version of the *IT job profiles nomenclature*.

[6] KPI Key Performance Indicator.

- **The data used to monitor**. This mode can be used, for example, to carry out a complete, multi-factorial diagnosis, to apply statistics, to make classifications, to score[7] to help decision-making.
- **Data used extensively, possibly with AI, to enrich the organisation's value proposition**. (For example, a customer visits the website. Knowing their customer data enables us to offer them directly the product that's right for them).

Companies and public administrations are generally siloed, and this is reflected in their IT. The resulting difficulties are a brake on key data-related actions such as:

- Identify useful data;
- Accessing data;
- Cross-referencing data;
- Profiling and quality control of data;
- Managing heterogeneous technologies;
- Preparing the data;
- Sharing results.

To combat these silos, Zhamal DEHGHANI has developed the concept of data mesh, which many companies are interested in. She defines the data mesh as "*a decentralised (distributed and federated), socio-technical approach (it is neither a technology, nor a technique, nor an architecture, it is a complete and global approach) to accessing analytical data for decision-making at scale*".

The data mesh is based on **4 pillars** which are explained in the following paragraphs:

1. Data domains;
2. Data as a product;
3. Self-service data infrastructure;
4. Federated data governance.

3.2 1ST PILLAR OF THE DATA MESH - DATA DOMAIN: ORGANISING DATA INTO DOMAINS TO BRING POWER TO THE BUSINESS LINES

There are several possible ways of grouping data into domains: by business line, by function in the organisation chart, by product, etc. The aim is to create a breakdown of the organisation's information assets that reflects its strategic priorities and major activities. The aim is to create a division of the organisation's information assets that reflects its strategic priorities and major activities. It is essential that this division by data domain **reflects the business reality** of the company or **services** for the public administration. This division must be into **independent domains** attached to business lines that **are responsible for the data within their perimeter**.

[7] Scoring is a technique used to assign a score to a customer or prospect. The score obtained generally reflects the probability that an individual will respond to a marketing solicitation or belong to the desired target. It therefore measures the appetite for the potential offer.

The organisation by domain makes it easier to identify use cases and therefore to choose the most relevant ones. On a completely different level, it also enhances the formalisation of business processes by integrating and listing the data used in these processes.

The domains are operationally responsible for creating, distributing and deploying these “data” products. Data domain managers need to be appointed to set up the organisation, and given responsibility not only for analytical data but also for operational data, from its generation through to its final use. This autonomy enables end-to-end data quality issues to be resolved at scale. Managers set priorities, define the data they want to access and maintain a common vocabulary. The domain is responsible for building and managing the data products. Each domain creates and distributes its data products, governs them and coordinates with the other domains.

For example, a company has chosen an organisation divided into 4 business areas, as close as possible to its core businesses: research & development / production / business services / services and solutions. Within these areas, it has created sub-areas. Another has chosen to set up data offices by region to be closer to the business reality. To do this, it has adapted the organisation of its IT accordingly.

The obligation for the various domains to expose their data to the entire company or public administration is an imperative condition for the effectiveness of this organisational model. It avoids falling back into new silos. The “products” and “governance” pillars, described below, also make this possible.

A strong commitment on the part of management seems to be a prerequisite. From one team to another within the same entity, the degree of maturity may be very different and the willingness to share inhomogeneous. The level at which decisions are made about priorities, within each domain and between domains, needs to be formally specified, bearing in mind that the general interest of the whole organisation must remain an absolute priority.

Finally, it is important to agree on a glossary of data-related professions/missions/roles to avoid the same name or concept being defined differently from one entity to another. This provides a foundation for the common good.

3.3 2ND PILLAR OF THE DATA MESH: DATA AS A PRODUCT

This 2nd pillar is highly structuring. **Data becomes a product designed to serve the domain's internal or external customers**. It must therefore be officially **consumable**, **validated and ready to use**. This implies documenting the “data” product with a “usage” vision and not just a technical one. The data can be raw or processed and of several types: a product for final use or a product used to produce other products. In fact, the “data” product can be enriched with other operational data to enhance the customer experience. Offering data as a product gives it visibility and avoids duplication or the multiplication of more or less similar copies. However, the notion of product needs to be clearly defined, as it is likely to cover a wide spectrum with quite different meanings. Like any product, **the “data” product must be promoted and evolve with consumer needs**.

Illustration of the "data" product concept

Imagine that data was water... and that we wanted to present it as a product.

A **container would** then be needed: the bottle for the water, the data set for the data. We also want to present the **product usage information** (label for water, metadata and documentation for data), do the **marketing** and **distribute** the products (via data marketplaces and APIs). By considering data as a product, we are not simply making data available, but covering the entire value chain from development to distribution. "Data" products therefore embody not only the data itself, but also all the elements required for its proper use and governance.

The concept of the "data" product is central to the data mesh and, more broadly, to the data-driven transformation of businesses. It allows us to review the methodology for constructing and using data by standardising the approach.

Mick LEVY, at BUSINESS & DECISION

Figure 1: illustration of the "data" product concept in the case where the data is water. Source Business & Decision - Data mesh white paper: Spiderman guides you towards the data-driven company

The "data" product meets the following 6 characteristics: it must be discoverable (i.e. visible), addressable, documented, reliable, interoperable and secure.

- **Discoverable**: the product is referenced in a data catalogue or marketplace. Metadata makes it easy to explore and identify.
- **Addressable**: each "data" product is located at a permanent and unique address that guarantees its continued use by the consumer domains, regardless of how it changes over time and in accordance with the access policy.
- **Documented**: the data is documented by the domain in a federated catalogue to ensure clear and unambiguous understanding and interpretation by consumers (location, data freshness, semantics, life cycle, data model).

- **Reliable**: the quality of the data is continuously measured and monitored by the producing domains (communication of quality indicators) to ensure that the products are reliable and that there is a high level of confidence in their use.
- **Interoperable**: products are based on common standards, making them easier to make available, re-use and cross-reference.
- **Secure**: as a strategic asset, data is protected according to its level of sensitivity and authorisation (access rights, authentication, encryption, etc.).

The products are systematically exposed to everyone, with simple, standard, centralised access via APIs and data marketplaces. In this way, the “data” product positions the data for all the company's internal “customers” and avoids unnecessary duplication.

This pillar should be implemented when the lifecycle of a “data” product or the expression of a business need is defined. This product must have all the attributes that enable it to be reused by other entities in the organisation.

To start building an initial product, some participants suggested starting either from existing raw data or from needs expressed for identified uses. These two approaches - bottom up by looking for valuable experiments, or top down by starting from a business need that is insufficiently met - are mutually enriching and help to define the most useful “data” products.

3.4 3RD PILLAR OF THE DATA MESH: FEDERATED DATA PLATFORM

This pillar proposes a common data platform enabling users to build their data processing on a common foundation and with tools guaranteeing interoperability. This platform must offer centralised capabilities that are useful to everyone, starting with a catalogue of data; it therefore creates the means to ensure user **autonomy** through common for using the data. This data platform, made available on request to the domains, is standardised, interoperable and federated. The choice of tools meets users' needs, whether for data visualisation and management, to facilitate decision-making, or for the creation and implementation of new products and services. State-of-the-art tools are made available to simplify usage and facilitate access to data.

Functions for acquiring, storing, transforming, accessing and consuming data are deployed using common tools (API Management, Data Catalog, authorisation management, etc.). The platform must be compatible with all types of technology in order to be able to evolve and meet needs. The models used to ensure data quality, processing, security and sharing in the most automated way possible must be replicable. It is important for resources to be allocated to the business on a self-service basis, so that machine capacities are allocated according to product access needs. The functions of this platform must be federated in order to control costs.

Ideally, the federated data platform will fulfil several functions:

- Implement the rules defined by Governance;
- Find out about existing products (documentation, quality of service, etc.);
- Develop and make available new products rapidly;
- Supervise and manage the production (RUN) of services using the “data” product;

- Manage detailed data access authorisations (identity management, traceability).

The platform provides rapid access to referenced data thanks to its visibility, and encourages the virtuous circle of developing "data" products that have the qualities required to meet the organisation's needs.

This federated data platform is not an exclusively technical subject. Its functions and quality determine the perception, within the company, of the value of a single data resource that can be easily used by the business.

The platform must be able to manage identity management (IAM) and guarantee solid RUNs with well-defined and monitored service levels. So that all employees, all potential users, can use it on a self-service basis, the platform must be easy to use, with its functionalities clearly explained. For this reason, training courses should be organised to encourage its use.

To implement it, you need to choose the most appropriate technologies and define a technically coherent framework, including an API hub where it is easy to create and add new APIs.

As siloing generally stems from a lack of trust in the organisation, it is very important that the platform enables access rights to be managed, and that information is only distributed to those who need to be able to access it, depending on its sensitivity.

APIsation strategy

An API, or application programming interface, is a standardised set of classes, methods or functions forming a front end through which software offers services to other software. Each API is a "doorway" to the IS: it provides access to data by transmitting it, without allowing access to the company's information system. APIs therefore offer the opportunity to "open up" the information system, to expose new services or processes and to access them. They also make it possible to exploit data in real time, particularly data from the Internet of Things (IoT). The increasing use of APIs means that they need to be managed, the associated traffic administered and monitored. This is why, faced with growing requirements for traffic management and analysis of exposed APIs, organisations very often call on an API management system.

There are a number of reasons why organisations should API their information systems:

- Increased safety and traceability requirements;
- Increase in inter-application data exchanges, particularly with SaaS applications;
- Data that comes from several places within the organisation: from the ERP system[8] , from the CRM system[9] , or from other sources. The complexity of the data justifies the construction of a catalogue within the organisation.

[8] ERP stands for Enterprise Resource Planning. It is a software package that enables "all of a company's processes to be managed by integrating all of its functions".

[9] Customer Relationship Management is a type of software that focuses on managing customer relationships.

LISI Aerospace's API strategy

LISI Aerospace has defined its strategy around 3 pillars:

- **Security**: security in traceability, tracking and APIs. Customers need to be able to be disconnected in the event of cyber-attacks on their side, for example;
- **API catalogue**: APIs providing access to data need to be catalogued and managed so that business users can test them themselves and check that they are suitable;
- **Skills**: the organisation is setting up training courses for its business units to develop their skills and development practices, and to enable them to adopt the right reflexes, such as giving priority to APIs (known as the "API first" strategy).

In order to meet the challenges of security, cataloguing and acculturation, LISI Aerospace has decided to equip itself with an API Management platform. Analytics and calculation tools are also available. The platform consists of 3 building blocks:

- API gateway
- API publisher is dedicated to administrators to help them develop APIs.
- API Portal, which displays the catalogue with all the data. Everyone in the organisation can see the data available, even if they do not necessarily have access to all of it. The APIs are stored in a data centre.

LISI Aerospace has put KPIs on each API so that users can rate them on different aspects, for example, the interest of the API.

APIsation, which meets the dual challenge of openness and security, is becoming more widespread within LISI Aerospace, with around twenty applications currently in production for 5 backends[10] . 4 applications are offered as SaaS. APIsation requires development work and technical knowledge. However, this technological subject is of interest to the business lines. LISI Aerospace recommends setting up a cross-functional organisation with identified resources to design, implement/build and test APIs. The business lines must be involved in order to ensure data quality and also to carry out tests quickly. The catalogue comes with around a hundred resources on display. The portal makes APIs easy to use, even for business users. The APIs currently receive several million calls a month, and there is real control, including a real view of how the APIs are being used and trends. The product vision of APIs has necessitated appropriate training to increase the skills of our teams. Today, the exposure of an API can move into the test/production phase in two days.

Eugenio FERRARIO, at LISI Aerospace

[10] In computing, a BackEnd (sometimes also called a background) is a term used to designate the output stage of software that is intended to produce a result. It is opposed to the FrontEnd, which is the visible part of the iceberg.

3.5 4TH PILLAR OF THE DATA MESH: DATA GOVERNANCE

This 4th pillar aims to ensure the functional exploitation of data assets and "data" products, and to maintain consistency in the operation of domains, while guaranteeing their interoperability. Thus, at enterprise level, the decision to create a single data catalogue, with common practices, formats, rules and standards such as access management, is a data governance decision.

Federated governance ensures that the **use of data** can be **monitored and controlled**. It is the user who is responsible for how the data is used. Governance makes it possible to know who uses what and how. It involves creating rules and providing a framework for the use of data by the domains, for example by specifying who is in charge of which "data" product, letting people know when to stop "data" products that are not being used, etc. The governance model is distributed across the domains and the leadership functions are federated.

> **Governance needed at Air France KLM**
>
> The need for governance and an operating model, the principles of which have been validated by the Air France executive committee, has become apparent in order to clarify the organisation (roles, bodies, processes) and talent management (training, career management). The business lines need to identify their data officers, data analysts and data stewards. At the same time, they need a strike force of data-savvy technical staff. Finally, most business lines consume data produced by other teams. Solid governance at corporate level is therefore essential.
>
> **Julie POZZI** at AF KLM

In a federated model, roles and responsibilities are shared between the central entity and the local entities. Governance rules are defined at central level, and relays in each entity ensure that they are applied.

Some participants noted the difficulty of working when the links are not managerial. Cross-functional relationships do not simplify choices. Management support is essential to decide who governs what. A federation management team or "data office" greatly facilitates the definition and dissemination of rules for the use of "data" products and the federated platform. The management team is headed by a CDO (Chief Data Officer) who reports to the executive committee.

Federated governance also makes it possible to orchestrate the data transformation, to break this transformation down into stages and to monitor its progress on an ongoing basis. Its aim is to provide an opportunity to create value by combining data, to democratise the use of data in confidence and to create a distributed "power & duty" model for a network of players who accompany and support the process.

A number of good practices were shared by participants:

- Set objectives for the staff member who assumes the role of data product owner within a domain and measure the use made of their data, whether by staff in their domain or by those in other domains;
- Identifying and managing the risks associated with data use, in order to build confidence and increase its exploitation;
- Audit and regulate dependencies and redundancies, as well as the organisation's semantics (definitions, data cataloguing) in order to “clean house” by deleting what is unnecessary.

3.6 POSSIBLE CHANGES TO THE ORGANISATION

Companies are transforming themselves by adopting different organisational models, depending on their level of maturity on the subject of data, and on their strategy. The target is the “data mesh” organisation described above, but no company has completely reached this stage. Initially, data capabilities reside essentially in a dedicated Business Intelligence department, managed within IT, then initiatives are piloted by the business lines with the support of IT, then they are coordinated at group level either by a CDO (Chief Data Officer), or by a central data department (or office). Finally, mature companies arrive either at a hybrid model with several CDOs distributed by business line, or at a hub-and-spoke organisation in which the Business Lines become autonomous mini hubs in their own right, before reaching the ultimate target: the data mesh described above.

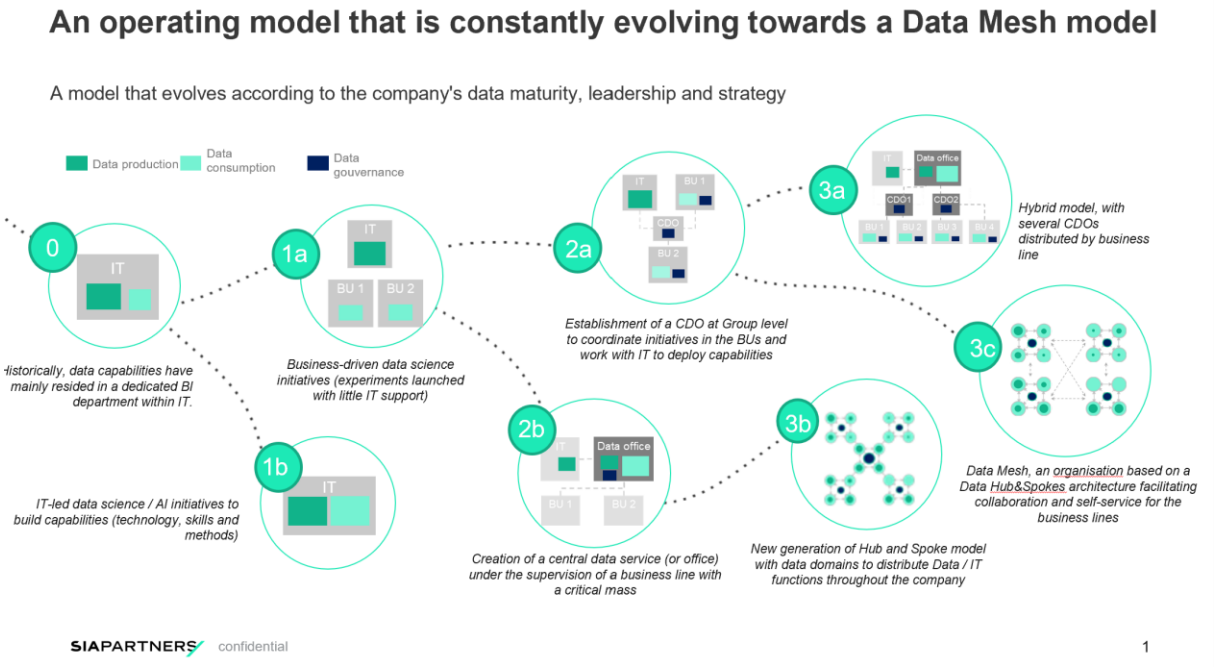

Figure 2: Organisational evolution trajectories. Source SIA Partners

4 HR AND ENVIRONMENTAL CHALLENGES OF DATA TRANSFORMATION

For data to become a transformative factor and be placed at the heart of management, **it is essential to get all employees on board**. **Employees** need to **be encouraged to use data in their reasoning and decision-making.** The aim of data transformation is to acculturate **all** employees to the use of data, its challenges and its best practices, by professionalising, democratising and supporting its complex and/or cross-functional uses. This is an effort that needs to be maintained over the long term to achieve a cultural change throughout the company or public administration.

Resistance to change varies widely. For example, the adoption of AI tools proves more difficult when business units fear a negative impact on the scope of certain jobs, particularly for operational functions. Employees identify the tasks that will be taken away from them, without necessarily seeing the new opportunities that are opening up at the same time. Supporting the data transformation process helps to address the issues of concern and to give meaning to the prospect by sharing the target, which is to “drive the business with data”. Indeed, it is important to explain to employees the reasons for the data transformation, what is at stake for the company or public administration and how to get there. Communicating on these subjects is essential if they are to have visibility and buy into the objectives.

4.1 LEARNING ABOUT DATA AND AI: TOOLS, SUPPORT AND COACHING

To develop a corporate culture based on the knowledge and use of data, organisations rely on communication and leadership teams who usually work in 3 areas: formats, communication channels and customised support.

Communication formats: tools are adapted to needs and made available according to the desired communication strategy: newsletters, presentations on data, seminars, podcasts, articles, video reports, etc. Data is collected at all levels, and particularly at operational level, which is why it is so important to explain the issues surrounding data (quality, completeness, etc.).

Communication channels: a variety of communication channels are used to meet different needs and share information in a targeted way, whether with a specific category of employees or more widely: internal social networks, tools to encourage teamwork, animation of communities of the curious, novices or trained, forums and portals.

Tailor-made support: specific communications are defined according to need and an action plan, supported by appropriate communications tools, is put in place.

The Air France KLM data factory

To support its data strategy, Air France is launching the Data Factory in September 2022, bringing together resources from IT, Business Lines, Transformation, HR and Communications. The data factory team is deliberately multi-disciplinary and very small, because the aim is not to "do things for people" but to **democratise the use of data**. Data transformation means moving from the centralised use of data by IT teams to open, direct and autonomous use by business units. The real challenge in this kind of transformation is to help business units to transform themselves so that they can use data in their own applications / solutions / products / services.

The Data Factory offers three services to accelerate data transformation:

- Professionalization and acculturation: the aim is on the one hand to develop the data culture and its challenges at AF KLM and on the other hand to develop the skills associated with data. Training and retraining courses are offered.
- The Dataviz Café, or "data visualisation" café: the aim is to make data accessible to everyone by providing local support for accessing data and creating dashboards. There are always at least two data engineers at the Dataviz café to welcome employees with questions about our data assets: "Who can I contact to get access to this data?" Or requests for support in creating dashboards: "How do I use PowerBI to cross-reference these types of data?" Finally, you can view a list of the company's existing KPIs or dashboards, so you don't have to reinvent the wheel!
- Supporting and accelerating data projects. This is done by setting up demonstrators or PoCs. The aim of this support is to accelerate the implementation of complex or cross-functional data projects, prioritised and led by the business lines right through to production, by providing them with skills and a proven methodology.

Julie POZZI at AF KLM

4.2 RESPONDING TO ENVIRONMENTAL CHALLENGES AND EUROPEAN REGULATIONS

A number of companies and public authorities are looking to use data transformation to measure **their overall CSR** (Corporate Social Responsibility) **impact**, respond to sustainable development and environmental issues such as the decarbonisation of IT strategy or the greening of data, and gain better control of their environmental impact. Indeed, in the absence of collective maturity in the ecosystem of organisations (suppliers, customers, vendors, partners, etc.) on environmental issues, organisations are rather powerless to control and manage the use of resources.

The global vision enabled by data transformation is therefore a key stage in testing actions and gaining maturity. Cigref's "Numérique Responsible" (Responsible Digital) Circle has shared its responses to

environmental and sustainable development issues in various reports[11] . The Cigref "IT for Green" working group is currently working on measuring the "net" value between the positive and negative impact of digital technology, based on feedback from participating companies.

As part of the "European Data Strategy" developed by the EU, a number of regulations aim to provide a framework for the role of digital services and markets, and to facilitate data transfers while ensuring the trust and security of European organisations and citizens. Some of these regulations have already been finalised, and all of them will require organisations to adapt and do a great deal of work. For example, the Data Act aims to maximise the value of data in the economy by facilitating the availability and sharing of data between businesses, consumers and public bodies. Complementing the Digital Governance Act, it specifies who can create value from data and under what conditions. It's still early days in terms of how organisations will articulate and integrate these regulations, but the impact on data-driven management is clear, and will require choices and adjustments.

4.3 IDENTIFYING DATA-RELATED ROLES, POSITIONS AND MISSIONS

In order to determine who should be responsible for the data, a company has studied its cycle, from its production to its use, including the processes that feed it. In most cases, this requires detailed knowledge of the business. This is why, in line with the recommendations of the data mesh, a number of organisations have decided to give the business lines **responsibility for data sets,** given their crucial role throughout their lifecycle. Data is an **asset** for a company or public authority, and forms part of its **heritage**. It therefore needs to be governed, protected, made accessible and upgraded in the same way as any other strategic corporate asset. Here again, in line with the recommendations of the data mesh, a number of companies are advising that federated data governance should be put in place, managed at company level, and that **responsibility for the most strategic data** should be assigned **to members of the executive committee**.

The roles, missions and professions involved in data are varied. For example, around the **structuring** and **governance** of data management activities, there are data owners[12] , data officers[13] and data stewards[14] , who are responsible for ensuring the **protection**, **accessibility** and **quality** of data. In some companies, the recognised data professions were previously often centralised in IT, such as data engineers and data scientists. But there were few or no data analysts or data business analysts, and their data role was not recognised. It was necessary to formalise the need for data analysts, data stewards, etc. in all business lines. The challenge is to define and promote data roles and positions within the business lines. It is essential to professionalise data-related activities, to size up requirements and to structure data-related jobs and the associated career paths.

[11] Cigref, in partnership with The Shift Project, Sobriété numérique: une démarche d'entreprise responsable, 2020.

Cigref, in partnership with INR, Sobriété numérique: Piloter l'empreinte environnementale du numérique par la mesure, 2021.

Cigref, Politique RSE au sein de l'IT : contributions positives de la DSI à la politique RSE de l'entreprise, 2022.

Cigref, Tool for assessing CSR criteria in IT projects, 2023.

[12] As soon as this job has matured, i.e. as soon as European companies have reached a consensus on its missions, tasks, deliverables and KPIs for measuring its performance, it will be included in the Cigref version of the IT job profiles nomenclature.

[13] See its definition in the IT job profiles nomenclature.

[14] See the appendix to the report.

The data mesh is also creating new roles and functions. It is essential that the executive committee commits to staffing these key data positions with the right people, playing the game of mobility between departments and promoting these positions as accelerators. The CDO (Chief Data Officer) is responsible for implementing and leading data governance within the organisation, and relies on the domain managers, the data domain owners. In general, governance initiatives are taken by the CDO and the Data Council, if the organisation has set one up, in consultation with the Data Stewards of each entity. It is worth noting the importance of data functions/roles, the scope of which is refined with the data mesh, but which differ from one organisation to another.

Companies and public authorities are looking to identify all the employees who carry out data-related roles, missions or professions. This makes it possible to identify in-house talent and skills and to develop them, as well as leading the various data communities, sharing best practice and communicating more effectively with these employees. It also enables us to identify the skills that are lacking or that need to be strengthened, in order to develop the attractiveness of these missions, coordinate recruitment and mobility, and set up training programmes at Group level. Some data-related jobs are already defined in Cigref's *IT job profiles nomenclature*, while others, now sufficiently mature, will be able to be characterised in the next edition in 2023-2024. Finally, some will not yet be defined because there is no consensus on them.

4.4 TRAINING

Companies and public administrations are facing a shortage of data and AI profiles. What's more, training for these new roles and missions does not necessarily exist on the market, or if it does, it is rare and therefore in high demand. This is why companies are not hesitating to create new training courses themselves, or to build courses based on identified needs, and offer them to their own talent. These training courses enable employees to develop their skills in the new “data & AI” functions and promote the retraining, as part of internal mobility, of employees interested in roles/missions in the data field. This approach is also a good way of retaining talent.

Courses are designed to meet the internal needs of data product owners, data analysts, data engineers and data scientists. Some courses are designed for internal profiles and offer training leading to certification. In these cases, the company works in collaboration with schools or organisations specialising in professional training. In one of the collaborations set up by a participant in the working group, the training company is obliged to select future students through interviews and tests, as the number of applicants generally far exceeds the number of places available. However, those who fail the course and are highly motivated are invited to take further training to improve their weak points, as identified with them, so that they can attend the next session. In most cases, employees who take the certification course have a job waiting for them when they leave.

Training offer to support employees in the field - example of a large group

For several months now, a company has been offering a range of services designed to support all employees in the field in their development in data/IA functions, as well as developing the autonomy of business lines in data/AI projects.

This offer is based on 2 actions:

Bringing the transformation to all the Group's field staff, with the following training courses:

- Evolution of corporate culture with the data/IA dimension;
- Training on business lines, skills and data/AI tools.

Supporting change for employees in the field, who are affected by data/AI projects in the branches and subsidiaries.

- Acculturation and skills upgrading for employees affected;
- Advice on the change management plan and associated deliverables;
- Co-construction on issues of scaling up.

Large company member of Cigref

A number of companies host and train apprenticeship students in data-related professions, with the most motivated of them being offered a job once they have obtained their diploma.

It is therefore important for organisations to **communicate clearly about the various training courses available,** both in-house and externally, so that employees can easily identify them. There are often several categories of training offered, depending on the objectives and target audience:

- There are clear routes to becoming a data product owner or data analyst, for example;
- Training in tools and technologies to support the digitisation of businesses, for example training in Python or machine learning;
- Training courses to help you become acculturated to data and AI. These cover the essential knowledge needed to gain autonomy over data, or to understand AI, as well as the way in which the latter enriches activities in order to develop its uses.

This 3rd training category meets a number of needs in the context of supporting the data transformation of companies or public administrations. For example, teams launching data and AI projects need to understand the semantics and governance involved. Then there are entities that produce data but don't know what to do with it, such as HR or Finance teams. Training offers them the opportunity to understand what these teams can do with the data, and therefore to express a more precise need. Training in how to identify use cases and implement them in the business unit meets this need for support. Some units have now become autonomous and only call on the data team for expert advice.

Some data/AI projects have a major impact on the jobs of certain employees. In order to support teams during their deployment, one company is offering a “change management” training course based on 4 stages:

- Analysis of the context of the change: inventory of the players involved, impact of the change and analysis of the socio-dynamics;
- The change management plan: training plan, communication plan, specific support plan for each target;
- Support and assessment of the change: evaluation of the change, action plan for its implementation, feedback and exchange of best practice;
- Consolidation of the data/AI ecosystem: assistance in building the data ecosystem, action plan to increase skills, data function integration plan, integration of best practices.

5 CONCLUSION

Data-driven business or public administration means being able to make the right decisions systematically and confidently, in a responsible and rational way, based on the assets represented by the organisation's internal or external data. This means having the right organisation in place, training all employees in the various issues at stake, and sharing high-quality data. That's why, before committing themselves, companies and public administrations are assessing the benefits of this data transformation, taking into account the time required, the organisation to be put in place to achieve it and the associated cost.

Certain teams or business areas, such as HR and Finance, are good candidates for starting to implement data-driven business management because they have well-structured data. However, before embarking on the process, the value provided by such a transformation needs to be assessed and matched to strategic and business challenges, or, in the case of public administrations, to their service offerings, in order to convince the executive committee of the benefits of such a transformation, which takes time and requires substantial resources.

For some years now, a number of companies and public administrations that are organising themselves to be data-driven have been participating in the development of European data spaces within the secure and trusted Gaia-X ecosystem. The challenge is twofold: to find use cases that create value for several of the stakeholders in the data space, and to put in place governance that is both flexible and structuring so that new ones can continue to be developed on a large scale. As Gaia-X is currently the only initiative aligned with the European strategy that meets this requirement for trust in data sharing, it is essential that businesses and public administrations play an active role in ensuring that its infrastructure and framework meet their needs and European values. Gaia-X is in fact an association, and what it produces and proposes is the fruit of the collective labours of its members and its team of permanent staff.

In the future, artificial intelligence tools, and increasingly generative AI tools, are likely to change the game. Firstly, by acting as a super assistant (or co-pilot) to give super powers to employees, but above all because they reveal all the potential offered by data, once it is massively exploited, to extract value from it. This is one of the topics that the Cigref working group on AI and generative AI will be addressing this year for the 2023-2024 cycle.

6 APPENDIX: THE JOB OF DATA STEWARD

At a meeting of the working group, we identified the tasks and duties of the data steward, a job that has emerged with the data-driven management of the company. After proposing a definition of the job, we listed the various activities below.

6.1 MISSION OF THE DATA STEWARD

The data steward is at the frontier between business processes and uses (including lineage) and IT implementation.

The data steward is a key player in the governance of the company's data. They ensure the quality of the data (i.e. how long it is kept and how it is purged). He or she describes the data and monitors its quality, with the aim of improving its usability in order to extract value from it. It contributes to the harmonisation of data management practices across the organisation.

He may act on behalf of the data owner.

It guides the consumers/users of the data and can play the role of referent in understanding the data.

6.2 TASKS OF THE DATA STEWARD

Data stewards carry out the following activities:

- **Ensuring data accessibility**:
  - Creating and providing users with the necessary access interfaces to enable them, depending on the level of confidentiality authorised, to access data via a dynamic, connected data catalogue.

- **Monitoring data quality and compliance**:
  - Ensuring compliance with standards defining expected levels of quality (relevance, reliability, security and durability) when data is created, processed, stored, accessed or deleted;
  - Ensuring that the level of data quality is compatible with the use cases that require it, and improving it if necessary.

- **Continuous quality improvement process**:
  - Document data (metadata, compliance information, business glossary, data model, master data, business rules, etc.);
  - Mastering the use of the data within the scope of his/her remit by supporting the manufacturing/production teams;
  - Ensuring that data consumption and exploitation by the various business processes are consistent with the intended use of the data;

- Gathering the needs expressed by consumers, as well as requests for improvements (particularly in terms of data integrity, quality and harmonisation in a specific area) and problems encountered;
- Prioritising and passing on requests and points for attention to the Product Owner.

- **Data governance (within its remit)**:
  - Contribute to the development of the Data Management policy
  - Identify the players in the data lifecycle and their responsibilities (from capture to exploitation/consumption);
  - Express definitions, standards and necessary data strategies;
  - Formalise the standard business processes involved in creating and maintaining data domains;
  - Ensuring that data policies and good data management practices are applied by all players;
  - Apply defined data sharing policies;
  - Contribute to regulatory monitoring of data;
  - Orchestrate data governance by administering the Data Lake.

Note: the data steward does not define the rules, but implements them. He is both a technical and a business manager of the data.

ABOUT CIGREF

Serving the economic growth and competitiveness of our members, large French companies and public administrations, users of digital solutions and services, through digital success.
Cigref is a network of major French companies and public administrations whose mission is to develop its members' capacity to integrate and master digital technologies. Through the quality of its thinking and the representativeness of its members, it is a unifying force in the digital society. Cigref was founded in 1970 as a not-for-profit association under the law of 1901.

To achieve its mission, Cigref relies on three areas of expertise that make it unique.

Membership
Cigref embodies the collective voice of France's major companies and public authorities on digital issues. Its members share their experiences of using technologies within working groups to bring out the best practices.

Intelligence
Cigref participates in collective discussions on the economic and societal challenges of information technologies. Founded nearly 50 years ago, Cigref is one of the oldest digital associations in France, and draws its legitimacy from both its history and its mastery of technical issues, the foundation of skills and know-how that underpin digital technology.

Influence
Cigref promotes and respects the legitimate interests of its member companies. As an independent forum for exchange and production between practitioners and stakeholders, it is a benchmark recognised by its entire ecosystem.

CONTACT US

www.Cigref.fr
21 av. de Messine, 75008 Paris
+33 1 56 59 70 00
Cigref@Cigref.fr


Cigref
SUCCEED
WITH DIGITAL