SEMANTIC PLATFORM
FOR THE DEPLOYMENT OF AI AND DATA ON A LARGE SCALE IN THE COMPANY

INTRODUCTION TO THE TRAIDA FRAMEWORK (TRANSFORMATIVE AI AND DATA SOLUTIONS)

March 04, 2024

This report in English is the original version of the text, also available in French on our website: www.engage-meta.com.

Keywords: business system, systemic approach, semantic platform, knowledge graph-oriented database, data governance, AI governance, responsible AI.

In this report, we refer to the "business system" as the overall system that enables the company to carry out its activities. It includes the information system (applications, databases, automated processes, infrastructure), the organization (actors, roles, non-computerized knowledge and processes, values, human resource management policy) and AI algorithms.

INTRODUCTION

This report presents a comprehensive approach to the large-scale implementation of artificial intelligence (AI) within businesses, emphasizing both strategic and technical perspectives. Our focus is on establishing an architecture and framework tailored for the expansive deployment of AI throughout an organization's entire business system, rather than delving into vertical AI applications. This white paper aims to serve as a foundational guide for executives and decision-makers, even those without a technical background, to initiate more operational tasks like strategic business framing of responsible AI, understanding its impact on team attitudes, training requirements, software tooling, and identifying initial application cases. While we concentrate on the architecture of the business system, our goal is to enhance your understanding of IT domain and grasp the key trends shaping the evolution of your business in the era of AI, equipping you with the knowledge to align your business system with the transformative power of AI and its data.

AUTHORS

The main author is **Pierre Bonnet**, an expert in enterprise architecture, data governance, and supporting companies in integrating new technologies. He is also the founder of the Engage-Meta community, a think tank focused on developing practices to enhance knowledge accumulation, as well as data and AI practices. **Professor Nguyễn Xuân Hoài**, an expert in artificial intelligence and Head of AI Academy Vietnam is the co-author.

- Pierre Bonnet – HLFL Consulting – pierre.bonnet@hlfl-consulting.com
- Nguyễn Xuân Hoài – AI Academy - nxhoai@aiacademy.edu.vn , (https://aiacademy.edu.vn)

Contents

AI CONTRIBUTIONS AND RISKS

WIDESPREAD USE OF AI

AI consists of a set of computer tools that are rapidly improving and becoming increasingly user-friendly. **Its entry barrier was broken with the release of ChatGPT in November 2022**. Today, one does not need to be a computer scientist to use AI. Its contributions are exponential and integrate into our daily lives as well as those of businesses. In this context, it is natural for decision-makers to worry about the future of their business if they were to miss the coming AI wave. Being outclassed by competitors better equipped with artificial intelligence is an immediate existential risk for all companies.

POTENTIAL MASS PROBLEMS

However, as Milton Friedman said, "there is no free lunch." The bright side of the exponential contributions of AI also has a dark side. These are the exponential risks generated by its misuse. As long as the company deploys these technologies on a few vertical applications, the malfunctions and unreliability of AI are compensated for by the organization. **Human intelligence still knows how to cope with AI failures, but only up to a certain threshold**. There comes a point where these disconnected vertical AI applications generate a mass of problems that the organization can no longer solve. In this situation, the company will turn around to reduce its AI usage and take the risk of being outclassed by smarter market players. It will then suffer an exponential growth of problems that harm employees motivation. It is likely that in the absence of anticipation of a large-scale AI deployment, technical and human chaos will ensue. **If your initial successes in AI mask your preparation for its large-scale deployment, you are in the situation of jumping out of a plane without a parachute while telling yourself "so far so good"**. Of course, this preparation is not easy and requires a lot of expertise. That's why we wrote this introductory paper. We propose measures for a large-scale deployment of AI. This is an architecture for “responsible AI” and a framework called TRAIDA (Transformative AI and Data solutions). They provide best practices to succeed in a progressive and operational scaling up.

AI BARRIERS

POOR QUALITY OF DATA FROM SILOS

Large companies have complex IT systems with numerous applications and heterogeneous databases. These systems are verticalized by functional and organizational silos. The quality defects are then multiple: duplication of data that generates errors; business processes that collide with the boundaries of silos; and complex, costly, and fragile technical integration solutions (ETL, EAI, MDM, Data Virtualization). Few companies undertake the task of overhauling their silos to develop applications and clean databases organized by business domain. Even though the IT industry offers technologies to achieve this, such as Service-Oriented Architecture (SOA) and Data Mesh, their implementation remains delicate. This kind of overhaul imposes a transformation over several years and mobilizes significant budgets with skills that are difficult to find. Faced with the weight of this project and to avoid questioning all the silos, solutions that preserve the existing ones are favored by companies. This is where tools for master data management, customer data integration, data virtualization and operational data store come into play. They involve implementing layers of data aggregation and cross-silo repositories. They then unify, more or less well, the data to increase the relevance of their governance and quality. The company thus retains its silos and partially improves the management of its data with imperfect solutions that preserve the existing system. **Yet, with AI, this approximate governance of data is no longer acceptable**. In the absence of quality data across the entire business system, AI systems do not function properly, and it is not always possible to compensate for the errors they generate. Moreover, in the enterprise context, it has been found in practice that **application of AI in silo or vertical business operations does not produce high return on investment**. To maximize the benefit of AI for business transformation, each enterprise should apply AI deep and wide across multiple business operations. In the rest of this paper, you will see how the "**semantic platform**" offers a solution to deeply improve the quality of data across the entire business system, without imposing an overhaul of the silos.

LACK OF INDIVIDUAL AND COLLECTIVE KNOWLEDGE ACCUMULATION

For IT experts, the acceleration and breadth of information technology no longer allow a single individual to know everything (a). The field of expertise of the IT engineer is thus reduced to a few vertical tools demanded by the market. It is normal for schools to train for these needs, however **it is necessary to place greater emphasis on cross-disciplinary skills such as data and process modeling, enterprise architecture, or the governance of complex systems**. Although they are being taught less and less, these disciplines are essential for the successful completion of transformation projects with AI and data. These projects are by nature cross-sectional within companies, with significant implications for system architecture and governance. **In terms of businesses, knowledge is poorly formalized** (b). IT solutions are interested in the structured data necessary for software operation and not enough in the knowledge of the individuals. The majority of knowledge remains in the brains of the actors with little collective sharing. The reluctance to write hampers the formalization of knowledge, which is essential for training AI systems.

LACK OF ALIGNMENT BETWEEN BUSINESS AND TECHNOLOGY

This fragmentation of IT knowledge (a) and the lack of formalization of business knowledge (b) do not facilitate the alignment between business and technology. **There is a need for more experienced architects with a broad IT culture, as well as for more business experts capable of formally documenting knowledge**. There is also a lack of practices for documenting and tracking the alignment of business with technology, particularly for transformation projects.

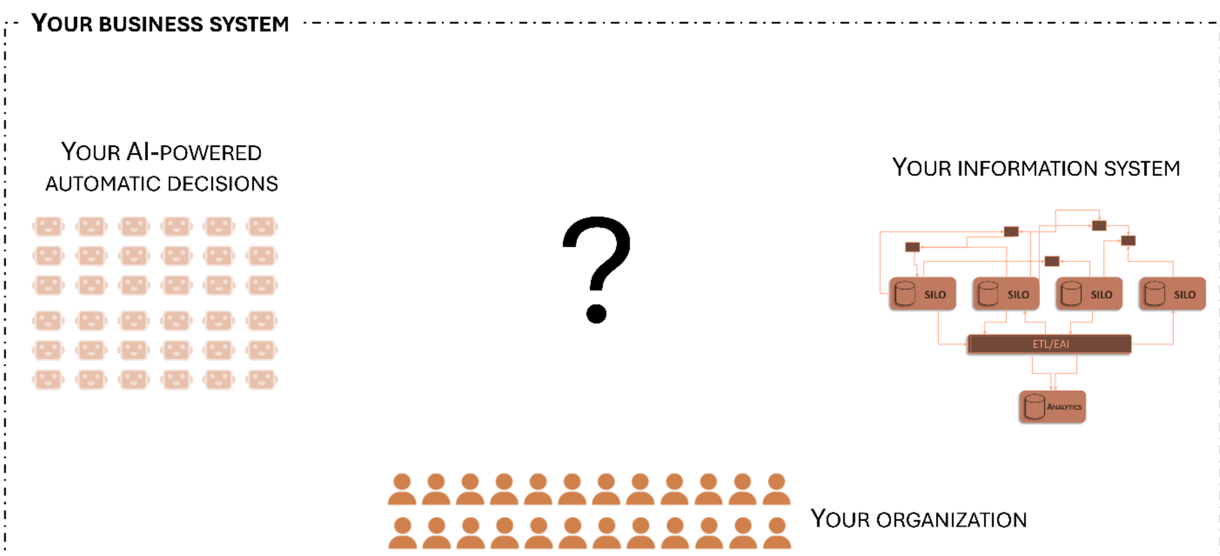
SEMANTIC PLATFORM

RESPONSIBLE AI

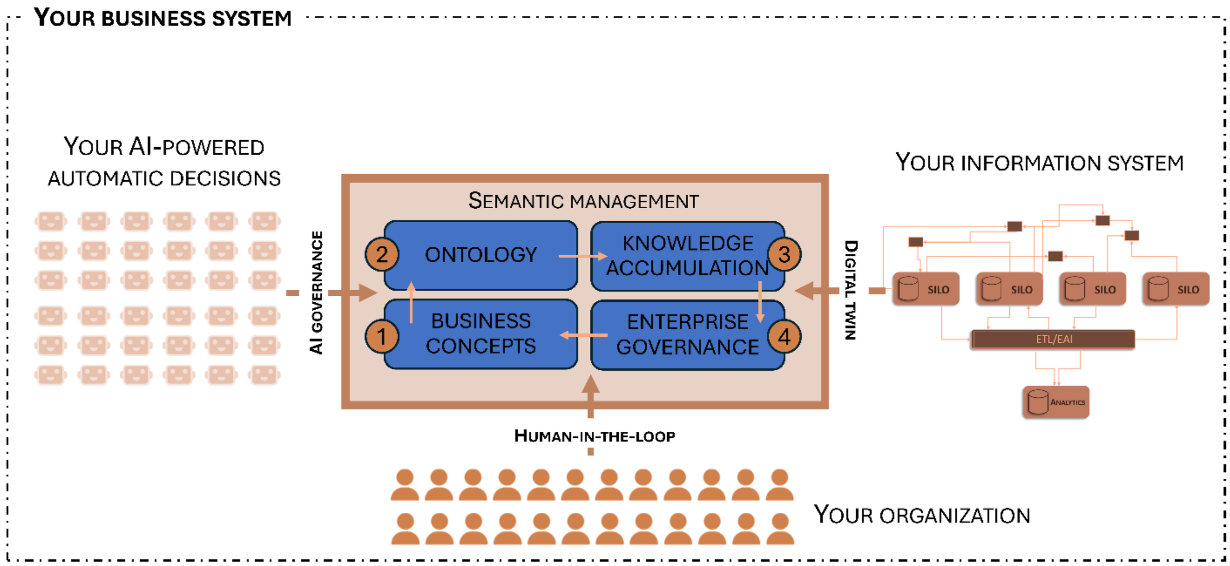
The barriers we have just described can be lifted to target a responsible AI provided that:

1. The quality defects in data are fully resolved without necessitating a systematic overhaul of silos;
2. The individual and collective knowledge of the actors are capitalized and used to train AI systems;
3. Business and technical alignment is formalized.

To achieve this, it is necessary on the one hand to have an architecture to integrate the information system, the organization, and AI at the scale of the enterprise; and on the other hand, a framework to accumulate the practices necessary for managing transformation projects with data and AI. This architecture is based on a "semantic platform"; and the framework for accumulating practices is based on TRAIDA (Transformative AI and Data solutions). Therefore, to successfully deploy AI on a large scale, these three domains of the business system must be coordinated (a,b,c):

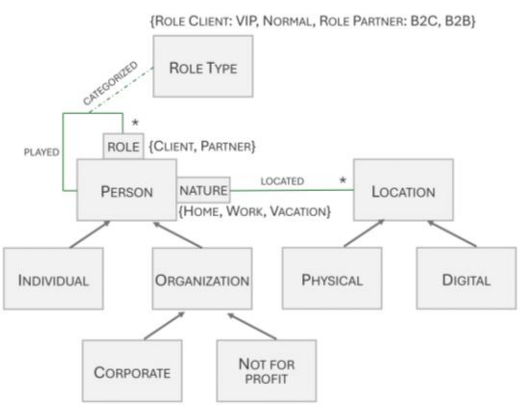


Firstly, on the right side of the figure is the **information system (a)**, illustrated here with its silos. It represents applications, databases, files, and automated processes. On the left side of the figure is **artificial intelligence (b)**, represented as a portfolio of intelligent algorithms and tools capable of making automated decisions. AI's contribution lies in its ability to enhance the automation of decision-making. To function, AI needs data sources from the information system and must interact with the third domain shown in this figure, namely the **organization (c)**. The question mark at the center of the illustration cannot remain unanswered by an infrastructure that covers all AI use cases. Without it, there would be connections in all directions between the three areas (a, b, c), with risks to the quality and traceability of results. The spaghetti plate that often exists with data management within silos would then risk being even amplified at the AI level. Therefore, a central infrastructure must be proposed to coordinate the entire system. Let's look at its principles in the next figure.



Our vision is based on a "**semantic platform**" that ensures the coordination function of the three areas formed by the information system, the organization, and artificial intelligence. This platform provides a glossary of business concepts, which are the key data elements and business objects used for the execution of operations across entire enterprise. They must be defined in a unified manner (enterprise glossary) and are used to construct the ontology from which AI algorithms draw data.

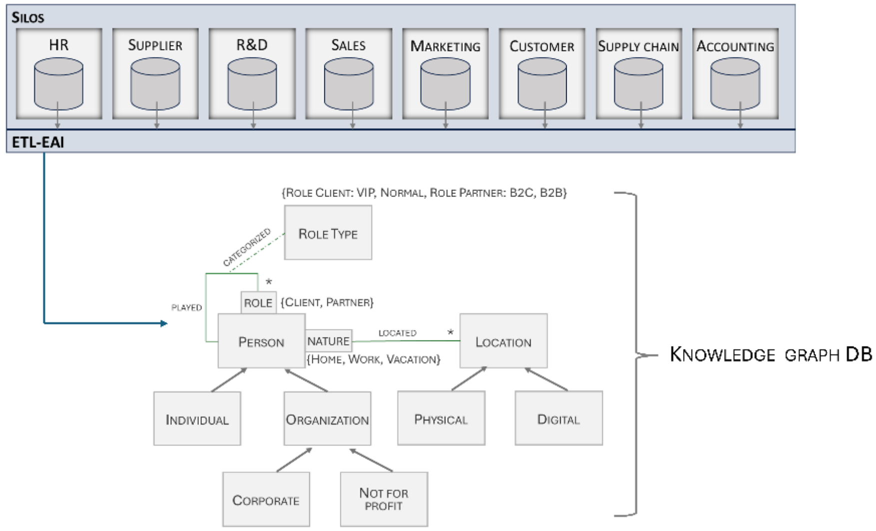
The ontology is an instance of the semantic model of information that includes fundamental integrity rules, such as the life cycle of business concepts. **This ontology is also the catalyst for accumulating personal and collective knowledge held by actors within the organization**. This knowledge is crucial for training AI algorithms. The aim is to facilitate the collection and classification of individual and collective knowledge through writing. Finally, the platform benefits from governance at the enterprise level to ensure its relevance over time.



The semantic platform acts as a digital twin of the information system, masking the details of applications and databases from both AI algorithms and the organization. It allows for correcting quality defects in existing systems. It is essential for deploying AI on a large scale and ensuring its proper governance (responsible AI).

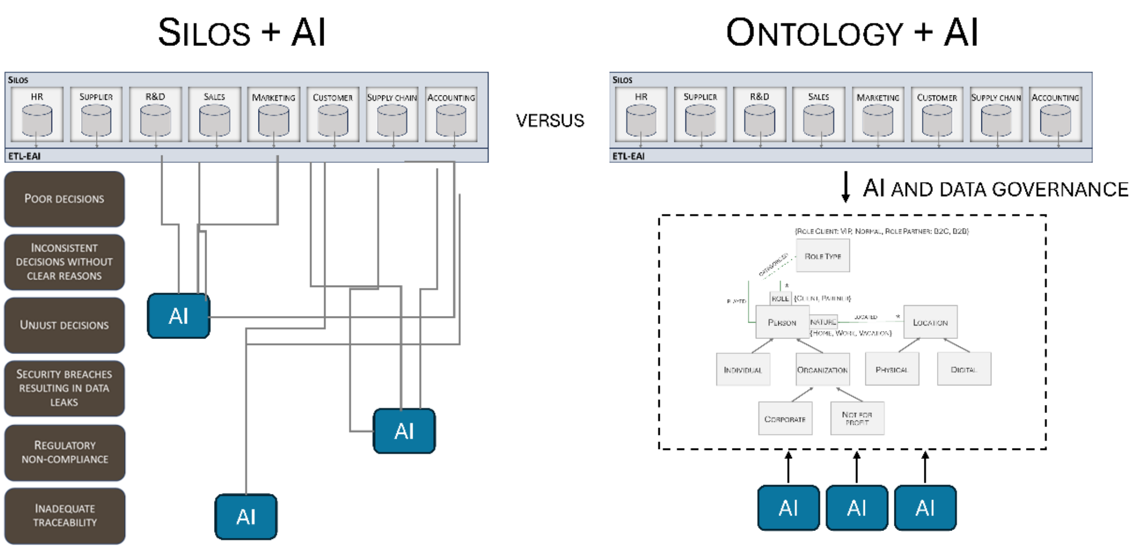
An example of a semantic model is presented in the figure opposite. An instance of the semantic model (ontology) describes the business concept of a 'Person'. You can observe various taxonomies, that is, classifications like the types of person such as Individual, Organizational, etc., or the types of roles including Customer, Partner, and Employee.

The ontology is implemented in a data repository that is transversal to the databases of the silos. **A good practice is to use knowledge graph-oriented database technology** (see last part of this report: “SUPPLEMENTARY MATERIAL: GRAPH DATABASES”).



Synchronizing data between the silos and the ontology requires good mastery of data flow management and technologies such as ETL and EAI. This data synchronization is a common topic in IT. When the ontology is read-only, synchronization is relatively simple to implement, but it becomes more complex when the ontology is also being updated.

In the left part of the figure below, AI algorithms are directly connected to data sources inside the silos. This leads to the propagation of data quality defects into the AI algorithms. Moreover, the lack of centralization prevents the establishment of a global governance for AI applications and operations. The problems are then multiple: the quality of decisions with AI is limited due to adherence to silos; inconsistent decisions between silos emerge, leading to a lack of trust in AI. Furthermore, unauthorized access to certain data within the silos leads to security breaches, poor application of regulations, and ultimately a lack of traceability in AI operations.



In the right part, AI algorithms rely on the use of an ontology that masks the data quality defects in the silos and conceals the complexity of accessing databases within these same silos. The ontology allows for unified governance of AI and associated data. The quality defects are then eliminated.

If you start your AI projects without the semantic platform, you need to be careful not to go too far or you risk creating complexity and fragilities that will be difficult to correct.

AI with silos and data quality defects cannot be considered 'responsible AI.' The semantic platform facilitates the building and deployment of responsible AI across the entire business system. Therefore, it is highly recommended to implement an initial version of the ontology from the first AI algorithms.

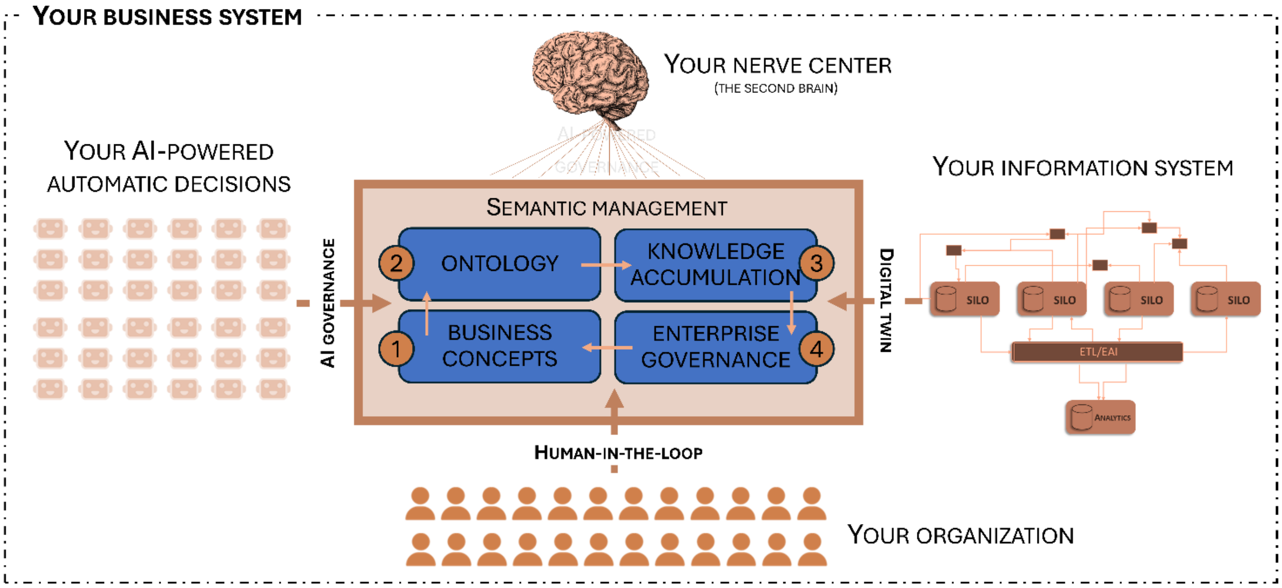
SECOND BRAIN

Our vision integrates a final concept whose role is to supervise the behavior of the semantic platform. By ensuring the quality of its execution, it is also the governance of the entire business system that is put under control.

Indeed, the semantic platform can also be a source of errors, just like the AI decision-making algorithms, not to mention the quality defects and malfunctions coming from the information system. Without the use of AI, it is impossible to code algorithms that are powerful and flexible enough to supervise the behavior of all applications, data, and processes. Today, with AI, this becomes possible. The principle is to provide a AI system with as much documentation as possible about the information system, data, algorithms, organization, and regulations. This can include specifications, user guides, database schemas, UI description and screenshots, vocal descriptions, minutes reports, etc. This AI observes the system's behavior to detect atypical, erroneous, or fraudulent actions. It can also discover improvements that a human would not be capable of detecting. It is called '**AI-powered governance**,' the nerve center of the business system, also referred to as the 'second brain.' **The richer the portfolio of automated decision-making AI algorithms, the more the overall intelligence of the business system increases, necessitating an AI-powered governance system.**

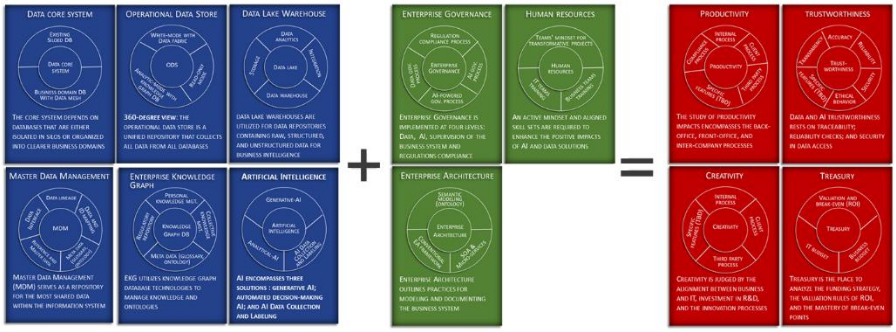
COMPLETE ARCHITECTURE

Here is the complete diagram of our vision with the concepts we have just described: at the center is the semantic platform with connections to the information system, the organization, and the AI algorithms. Above, the AI-powered governance (second brain) ensures the control of entire system. **Without this architecture, the large-scale deployment of AI is neither reliable nor secure. In other words, the semantic platform provides the features to apply "responsible AI" across the entire business system.**



IMPLEMENTATION (TRAIDA FRAMEWORK)

PRINCIPLES

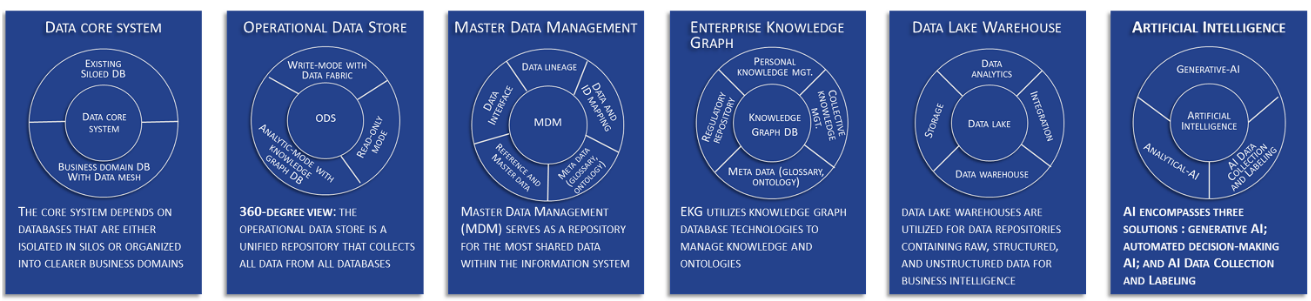


The TRAIDA framework (Transformative AI and Data solutions) consists of a series of cards. Each card addresses a domain that contributes to the large-scale deployment of data and AI. The cards in blue cover the technical perspective, and the cards in green cover the governance perspective. In red, we have the business cards. Each card allows for the accumulation of best practices and describes the enterprise's context. The technical and governance perspectives must be aligned with the business perspective. This idea is expressed with the equation "IT + Governance = Business". TRAIDA comprises 13 cards and a total of 49 topics divided into 21 technical topics, 10 governance topics, and 18 business topics.

All these cards and topics must be taken into consideration for the study and implementation of the semantic platform for the large-scale deployment of AI and its data.

The TRAIDA framework aids in the complete and holistic analysis of your technical context, your governance, and the alignment with the business. You have the option to customize the cards to your situation, to add or remove them, and to adapt the topics. However, we advise you to first try the initial version of TRAIDA to get used to its use.

TECHNICAL CARDS



Data core system:
- ✓ Existing siloed-DB
- ✓ Business domain DB with Data mesh (data as a product)

Operational Data Store:
- ✓ Read-only mode (including virtualization)
- ✓ Write-mode with data fabric (including CDI)
- ✓ Analytics-mode (with knowledge graph DB)

Master Data Management:
- ✓ Data lineage (flows)
- ✓ Data interface (SOA)
- ✓ Data and ID mapping
- ✓ Reference and master data
- ✓ Metadata (glossary, ontology)

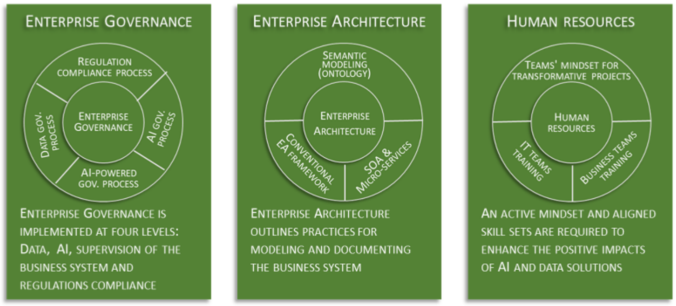
Enterprise Knowledge Graph:
- ✓ Personal knowledge mgt.
- ✓ Collective knowledge mgt.
- ✓ Regulatory repository
- ✓ Metadata (glossary, ontology)

Data Lake Warehouse:
- ✓ Data analytics
- ✓ Integration
- ✓ Data warehouse
- ✓ Storage

Artificial Intelligence:
- ✓ Generative-AI
- ✓ Analytical-AI
- ✓ AI data collection and labeling

The technical perspective of the TRAIDA framework explores how data management and AI technologies are implemented according to the following six cards, from left to right in the figure above: data core system, operational data store, master data management, enterprise knowledge graph, data lake warehouse, and AI. Each card contains a series of topics. For instance, the "data core system" card includes topics related to databases in silos and those developed by business domain with data mesh.

GOVERNANCE CARDS



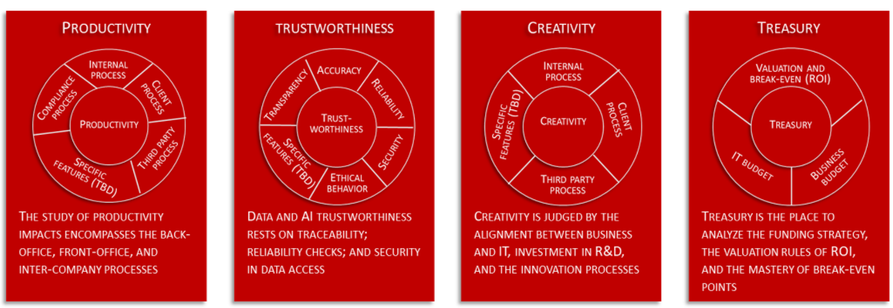
- ✓ REGULATIONS COMPLIANCE PROCESS
- ✓ AI GOVERNANCE PROCESS
- ✓ AI-POWERED GOVERNANCE PROCESS
- ✓ DATA GOVERNANCE PROCESS (QUALITY, CLEANSING…)

- ✓ CONVENTIONAL EA FRAMEWORK
- ✓ SEMANTIC MODELING (ONTOLOGY)
- ✓ SOA & MICRO-SERVICES PRACTICES

- ✓ TEAMS' MINDSET FOR TRANSFORMATIVE PROJECTS
- ✓ IT TEAMS TRAINING
- ✓ BUSINESS TEAMS TRAINING

The figure above shows the cards for governance with their topics: enterprise governance, enterprise architecture, and human resources. The topic of semantic modeling (ontology) is located in the "enterprise architecture" card.

BUSINESS CARDS



- ✓ INTERNAL PROCESS
- ✓ CLIENT PROCESS
- ✓ THIRD PARTY PROCESS
- ✓ COMPLIANCE PROCESS
- ✓ SPECIFIC FEATURES (TBD)

- ✓ ACCURACY
- ✓ RELIABILITY
- ✓ SECURITY
- ✓ ETHICAL BEHAVIOR
- ✓ TRANSPARENCY
- ✓ SPECIFIC FEATURES (TBD)

- ✓ INTERNAL PROCESS
- ✓ CLIENT PROCESS
- ✓ THIRD PARTY PROCESS
- ✓ SPECIFIC FEATURES (TBD)

- ✓ VALUATION AND BREAK-EVEN (ROI)
- ✓ IT BUDGET
- ✓ BUSINESS BUDGET

The figure above displays the four business cards: productivity, reliability, creativity, and treasury. The TRAIDA framework considers these as fundamental generic domains for assessing the alignment of the business with technical and governance perspectives. To accommodate business topics dedicated to your context, the first three cards provide the opportunity to define "specific features."

PROCEDURE

The procedure for using the TRAIDA framework is executed in three steps:

1. First, the framework is used to assess your level of technical and governance maturity necessary for the success of your transformation projects with AI. The cards colored in blue and green are used.
2. Then, the business perspective is added. This involves evaluating the alignment between your business objectives and the levels of technical and governance maturity. The cards colored in red are used.
3. Finally, the last step focuses on your transformation projects around AI and data management. This involves building a strategic and operational vision, based on the analysis of the results of the two previous steps.

In steps 1 and 2 of the procedure, the analysis covers the entire information system. In step 3, the scope is no longer holistic but rather limited to the portfolio of transformation projects.

1. EVALUATION

The TRAIDA scoring system operates on the topics of each card. It is based on these three levels of maturity:

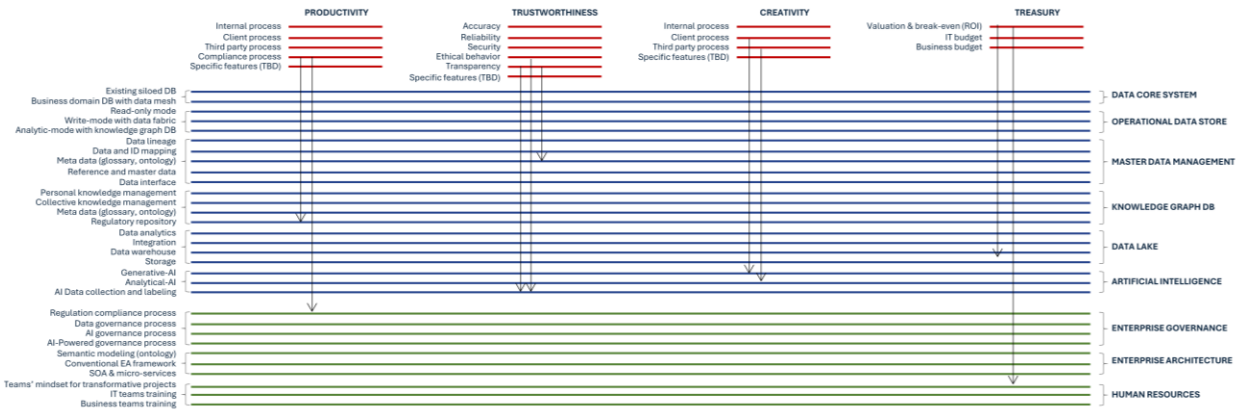
- 'Low' indicates that there is no large-scale experience.
- 'Medium' suggests that the company has begun to explore large-scale deployment.
- 'High' means that the company has established functional practices for large-scale deployment.

The measurements focus on technical and governance perspectives, using two forms highlighting scores for the existing system and those for the target vision. The blue form is for the technical cards, and the green form is for the governance cards.

2. ALIGNMENT WITH BUSINESS

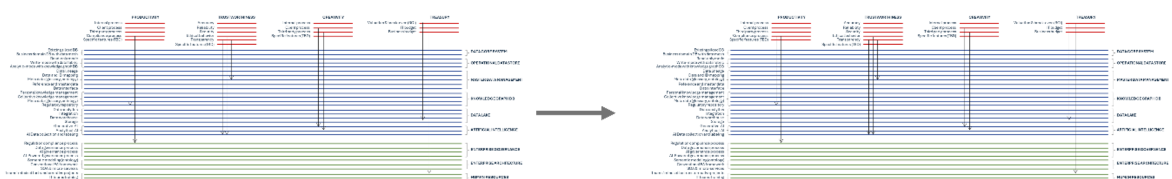
A matrix describes the alignment of the business cards with the technical and governance cards.

The principle involves drawing lines of dependency from business topics to technical and governance topics. This mapping initially concerns the analysis of the existing situation derived from the first step of the procedure. The arrows can adopt the "Low", "Medium", and "High" rating system to provide more information on the maturity of business alignment with technical and governance perspectives.



For practical reasons, it is advisable to select priority business topics to create a first variant of the matrix. Other variants then apply to less priority business topics.

Based on the matrix of the current state of the system, improvement goals are identified for new technical and governance perspectives. The matrix is then updated into various variants to target different expected achievements.

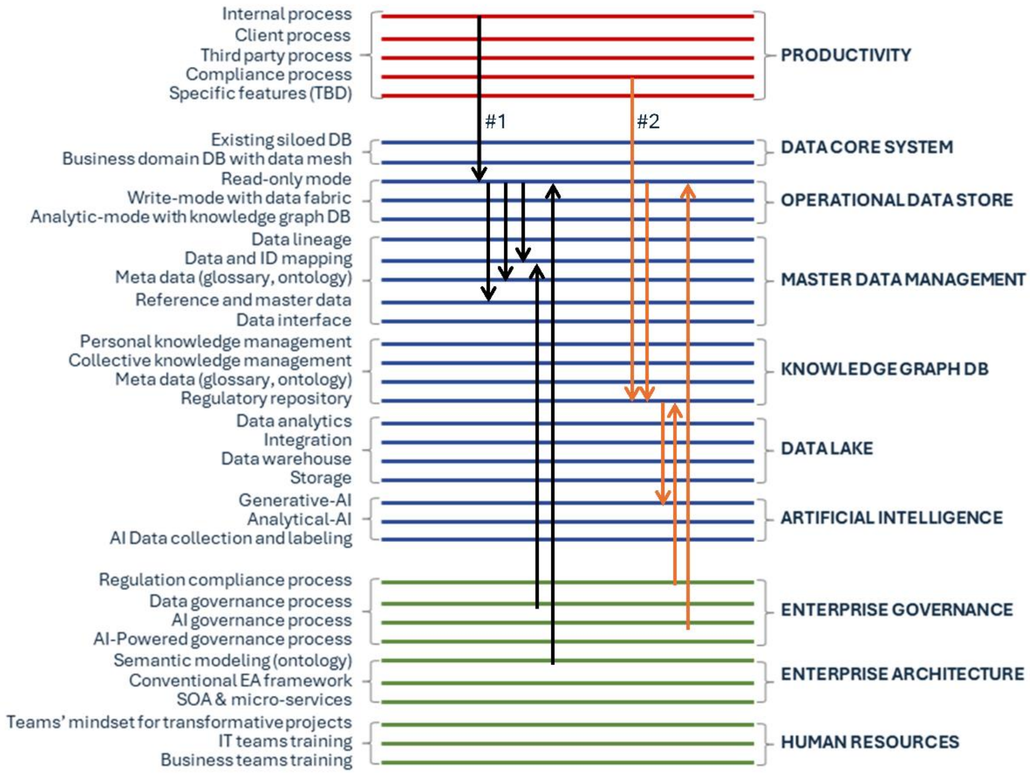


CURRENT SYSTEM EXPECTED ACHIEVEMENTS

Thanks to the TRAIDA framework, the study of business alignment with technology is based on matrices that helps to clarify thinking and decision-making.

3. PORTFOLIO OF TRANSFORMATION PROJECTS

The matrix below highlights an example of a transformation project in the field of 'Customer Data Integration (CDI),' marked with the business objective of 'Productivity".



In this example, the arrows represent the dependency links between business topics and those of technology and governance. Here, the black arrows describe the implementation of the solution, and the orange arrows describe the use of the customer repository to track regulatory compliance:

- #1 – Implementation of Customer Data Integration – CDI (black arrows): Increase in the productivity of the internal customer management process by implementing a read-only operational data store (ODS) that relies on Master Data Management system (MDM) for the management of reference and master data, and deduplication of identifiers. Data governance supports deduplication rules and data lifecycle management. Implementation of semantic modeling to create a unified model with a 360-degree representation of customers (ontology).
- #2 – Use of CDI as a repository for regulatory compliance (orange arrows): copying of the ODS into a regulatory compliance analysis repository using knowledge graph-oriented database technology. Generative artificial intelligence supports the creation of graphs on customers and helps detect atypical cases, fraud, and business opportunities. The entire system relies on enterprise governance for regulatory compliance and on AI governance.

By applying TRAIDA to your transformation projects, you build and manage your roadmap. It then becomes an excellent communication tool for your stakeholders to explain your strategy in AI and data management.

WHAT ARE THE FIRST STEPS TO TAKE?

To grasp the significance of the “semantic platform” advocated in this paper, it is necessary to identify the risks associated with deploying AI systems without such an infrastructure.

Therefore, imagine your business system with multiple AI systems or agents interacting directly with your databases, without central supervision. Also consider that the more a system increases its intelligence, the harder it becomes to keep it under control. If your AI systems produce erroneous results due to poor data quality and security flaws, you will have no choice but to disconnect them and seek alternative solutions.

To avoid this situation of regressing, this report proposes a vision of architecture with the “semantic platform” and the TRAIDA framework to guide the implementation. **However, once the content of this paper is validated in your context, there will still be potential barriers to overcome before beginning your journey towards 'responsible AI’:**

- Do not systematically reject the idea of gradually renovating your silos to correct deep data quality defects and streamline certain strategic business processes;

- Do not reject the complexity of scaling up AI, but use it as a lever for action on the path to improving the business system. Indeed, the TRAIDA framework may seem daunting due to the number of topics it covers, but you control its use and decide on the important topics. In other words, do not let complexity hinder your global initiatives, but take control of it. With the advancement of AI, this mindset becomes a necessity to maintain its control.

- AI is a technology that will replace part of humanity's work in intellectual, financial, medical, creativity activities, as well as in physical activities like transportation, logistics, and commerce. For individuals to maintain a competitive edge over this technology, they must cultivate a mindset that is interested in their profession to continually improve and not be surpassed by AI. In other words, employees can no longer act passively in their work. They must better understand their company and the use of AI in their profession. They must make an effort to formalize their personal knowledge in writing and share it in a collective approach to success. They need to develop an entrepreneurial mindset within their company, which equates to an “**intrapreneurial**” attitude. Without an evolution of your organization and teams towards this mindset, it is likely that AI will lead to a degradation of knowledge in the company. Who can imagine a company operating with only AI systems, without the presence of humans capable of keeping control? **Such an organization might show strong results during its rise to power but would eventually die under an effect of "AI+AI consanguinity”**. The sustainability of organizations will pass through the reinforcement of the teams' mindset for entrepreneurial engagement with their employers, supported by a "Humans + AI" coupling.

CONCLUSION

This report is an introduction to considering a holistic vision of AI deployment in businesses, along with the necessary data for their effective execution. It is also named “responsible AI”. This topic is challenging to address for two main reasons:

1. **Its analysis requires bringing together multidisciplinary expertise in AI, data governance, and support for actors in their relationship to work**. Within the Engage-Meta community, we work towards this union of expertise and provide support tools such as the TRAIDA framework. Our publications are free to use, under an open-source Creative Commons license.

2. Its implementation, even if gradual, can be a concern for decision-makers given the cross-functional nature of the approach. This is an usual barrier for ambitious transformation projects on paper, which often end up leading to vertical projects that are more or less well-synchronized. **Given the massive impact of AI on the future of businesses, a tepid policy of cross-functionality is no longer viable**. This is why this paper emphasizes the importance of the architectural vision, even if it may appear conceptual to less experienced readers. It finds a more operational expression in the TRAIDA framework, also described in this paper. Finally, it is implemented concretely with technical solutions of knowledge graph-oriented databases, and others that we have not listed in this report.

If you wish to delve deeper into the understanding of the architecture described and the TRAIDA framework, do not hesitate to contact the authors.

SUPPLEMENTARY MATERIAL: GRAPH DATABASES

The core technology of the semantic platform is based on knowledge graph-oriented database technology. It allows for the development of three natures of data repositories necessary for the platform:

1. The management of ontologies;
2. The accumulation of individual and collective knowledge necessary for enriching artificial intelligence algorithms.
3. The 'second brain' repository for AI-powered governance, which includes creating knowledge repositories for regulatory purposes. For example, a graph sourced from the text of a regulation (using LLM + graph coupling) is used to monitor compliance with policy rules.

If you are new to the technology of knowledge graph-oriented databases, we will review the main technical criteria to understand its contributions. The table below includes two technologies for comparison with graphs. The most classic is OLTP (Online Transaction Processing) with tools like Oracle or Microsoft SQL Server, located on the right side of the table. In the center is the knowledge graph technology with tools like Neo4J, Stardog or the Palantir AI platform. Finally, on the left, we present generative AI with LLM for Large Language Model with tools like ChatGPT or Claude (Anthropic). Although LLM is not a database in the strict sense, it is nevertheless a repository of knowledge. Other technologies exist, such as vector databases or Data Lake Warehouse. In this paper, we limit ourselves to OLTP and LLM, which are sufficient for comparison with knowledge graph-oriented databases:

	LLM	Knowledge Graph	OLTP
Probabilistic	✓	✗	✗
Deterministic	✗	✓	✓
Transactional - Integrity	✗	✗ ✓	✓
Cardinality management	✗	✗ ✓	✓
Human language readable	✓	✓	✗
Risk of hallucination	✓	✗	✗
Cognitive capability (e.g., inferred relation)	✓	✓	✗
UI on structured data & governance business features	✗	✓	✓
Data update on large volume & real-time	✗	✓	✓
	Generative AI	Semantic management	Transactional data

- Only LLM has probabilistic behavior. If you ask the same question multiple times to a tool like ChatGPT, you will get different answers. **OLTP and knowledge graph technologies are deterministic**.
- For transaction management, i.e., data storage integrity, OLTP technology applies it perfectly thanks to a formal data schema that describes the structures of the data and their constraints. Consequently, OLTP is secure but also rigid when it comes to absorbing evolutions in data structures and inappropriate for information without a predefined structure. On the other hand, LLM processes text freely, without underlying structure or concern for integrity. **Knowledge graphs technology brings interesting flexibility as it can operate with or without transaction and integrity management**. Indeed, the business use cases determine the implementation of integrity rules. Thus, a designer of a knowledge graph database works without a net compared to an OLTP database designer. In other words, the flexibility of the knowledge graph comes with the cost of mastering the use cases and associated modeling rules necessary to enforce integrity.

- The criterion of cardinalities is at the same level as integrity and transaction management. A business constraint like "a client can only be managed by a single salesperson" is solidly implemented in OLTP and cannot be violated. **In a knowledge graph, implementing this rule is also natural, but it is always possible to foresee use cases that go beyond this constraint**, for example, if some clients need to be managed by several salespeople located in different areas. In LLM, this type of constraint does not make sense, even if one could ask ChatGPT to check in a set of documents if each client is attached to a single salesperson. This is a data quality control request only, not the creation of new data.

- The ability for a human to read data is not possible in OLTP, and it requires a technical language like SQL. It has the advantage of enforcing the use of a formal syntax and grammar that avoids ambiguities. However, it is reserved for technically qualified users. In LLM, natural language is the exchange interface, which can induce semantic ambiguities when the text of queries is not well formulated. However, in data analysis, it is now quite common for a natural language query, for example with ChatGPT, to be translated into a formal data access language like SQL. **Here again, the knowledge graph offers interesting flexibility. Indeed, a human can discover and manipulate data through a graph that does not require IT skills**. It is therefore human-readable. Additionally, various technical languages for creating and manipulating data allow formal management in the same vein as SQL.

- LLM incurs the problem of hallucination. This is a consequence of the creative and probabilistic text generation mechanism of LLMs. **OLTP and knowledge graph technologies do not create hallucinations. Their operation is deterministic**.

- The capacity for transitive reasoning exists with LLM and with knowledge graphs. The classic example in the world of artificial intelligence is: "All men are mortal, Socrates is a man, therefore Socrates is mortal." **Utilizing this transitivity of reasoning is interesting and feasible with knowledge graph technology**. It does not exist in OLTP.

- **The creation of business applications is facilitated with OLTP and knowledge graph technologies**. They offer governance frameworks and rapid application development tools. This is not yet the case with LLM.

- Finally, the last criterion concerns the ability of technologies to manage large volumes of data in update and real-time. **Here OLTP stands out, but depending on the use case, the knowledge graph can also be used**. This criterion is not relevant for LLM whose learning process operates on large volumes of data but without real-time multi-user access. Once the learning process is completed, users work in isolation on their own data for the personalization (fine-tuning) phase and then querying with prompts, without mass data updates.
