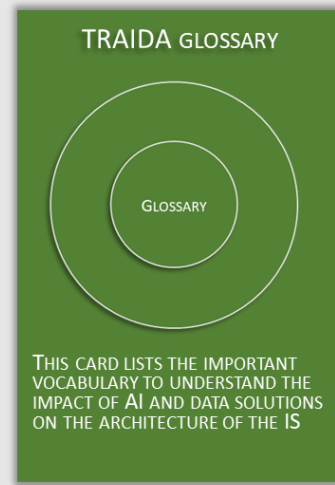


TRAIDA GLOSSARY

To increase your chances of spreading a culture of AI and data management that is understandable by all of your technical and business teams, it is essential to establish and share a glossary of AI and data solutions terms. Although popular, some of these terms do not always have a definition commonly recognized by the market. You will therefore need to decide on your vocabulary choices. This card gives you the starting point for this semantic work, which is fundamental to building and managing your transformation with AI and data management, encompassing many concepts and techniques.



(Under construction - adapted progressively as the TRAIDA cards are written).

D

Data fabric

A data fabric is a comprehensive set of technologies designed to streamline data integration processes, including referencing data sources, data cleaning, and unifying data formats. It supports the configuration and testing of decision-making algorithms (e.g., machine learning, AI training), as well as the deployment and monitoring of processes and data in production environments. While data fabric can assume some roles of a data hub, it primarily enhances data and AI governance at scale. Rather than replacing MDM, ODS, and EKG repositories, it coordinates them. In a data mesh context, data fabric extends to managing data caching, inter-database transactions, workflow management, and long transactions.

Data hub

A data hub primarily functions as a data flow integration bus, incorporating technologies like EAI (Enterprise Application Integration), ETL (Extract - Transform - Load), and ESB (Enterprise Service Bus). Depending on the solution, a data hub can manage metadata (mainly at the flow level), map IDs across silos, visualize unified data, and store certain operational data akin to an ODS (Operational Data Store). Coupled with a data mesh approach, it can also handle data caching and long transaction management. While some vendors market data hubs as universal data management platforms, they often fall short of fully implementing MDM, ODS, and EKG systems. It's typically more effective to use data hubs for integrating data flows and supplement them with dedicated solutions for MDM, ODS, and EKG.

Data mesh

Data Mesh is a data architecture approach that organizes data by business domains or concepts, rather than by functional or organizational silos. It uses semantic modeling and a technical infrastructure to manage transactions between business concepts spread across different databases. Data Mesh

enhances data governance and reduces duplication. It is a set of architectural principles rather than a specific technology. Implementing a Data Mesh requires leveraging data fabric and data hub technologies, tailored to the specific context of each company.

**Data fabric, hub
and mesh
overview**

The term 'data mesh' is relatively straightforward to define, as it refers to a data architecture that organizes data by business concepts to reduce silos. In contrast, 'data hub' and 'data fabric' are more complex to define precisely, as different software vendors encompass various concepts within these terms. At TRAI DA, we prioritize identifying the needs of the three fundamental repositories—Master Data Management (MDM), Operational Data Store (ODS), and Enterprise Knowledge Graph (EKG)—regardless of the chosen data hub and data fabric solutions. No single software solution can universally manage these three repositories on the same platform. Additionally, evolving AI governance requirements are driving the development of new generations of data fabric, focusing on managing AI prompts and related briefs.