Towards a Data-Centric Architecture in the Automotive Industry

Daniel Alvarez-Coello, Daniel Wilms, Adnan Bekan, Jorge Marx Gómez

Summarized by: Noah Lo-Hive Nam

Abstract

In the past two decades, vehicle software architectures have been evolving to support data-driven functionalities. Several enterprises from different domains are actively focusing on the improvement of their data architectures by redefining their data models. The goal is to facilitate robust support for analytics and artificial intelligence.

However, the automotive industry and its newly connected vehicles require a wide variety of data, increasing the need to develop and adopt standards. The paper highlights ongoing research conduct by BMW Research Department, focusing on the design of software architectures and addresses the need for adopting standardized data models towards a data-centric architecture in the automotive industry.

I. Introduction

Recent years have witnessed significant advancements in data and analytics capabilities. This shift led data, rather than intuition, to be the driver of behind digital innovations and business decisions. Nevertheless, a consensus has emerged around harnessing the full potential of data through a modern and scalable data architecture.

This progress has a significant impact on the automotive industry where vehicles incorporate diverse data-driven functionalities, impressive computation resources and hundreds of sensing devices. In this sense, their software architecture evolution is characterized by a strong emphasis on optimizing data utilization to support a myriad of applications and services (e.g., application ecosystems, traffic control)

As future innovation will depend on the speed data can be interpreted and consumed, and the ease data is integrated across domains, standards emerge as crucial factors.

II. Digital Transformation

Digital technologies dramatically reshape the industries, and many companies are pursuing transformation to gain benefits or to keep up with competitors. Digital transformation requires a different mindset and company's overall software architecture mut be redesigned.

An example of digital transformation is digital photography, which has completely transformed activities. Kodak, one of the most famous companies in the photo industry (which invented the digital camera in 1975), decided to persist ignored the disruptive change, which led the company to unimaginable losses.

Applying the lessons from Kodak to the automotive industry implies that recognizing the essential role of data and analytics is paramount. The authors argue that digital transformation should start with a shared understanding of data to redesign existing architectures focusing on four main transformation stages. These stages guide the transition from architectures characterized by isolated data silos to those adopting a cutting-edge AI server with a data-centric approach.

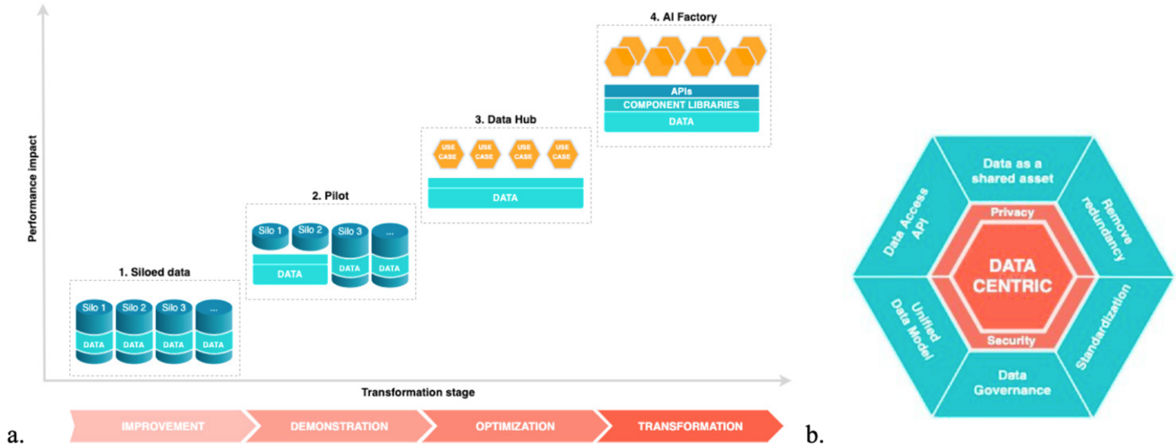

Fig. 1. (a) Stages of transformation to move from an architecture with siloed data towards a state-of-the-art architecture that supports AI and analytics [9]. (b) Some of the aspects involved in a data-centric architecture, where data is the center and most important piece [11].

As shown in Fig. 1, four stages clarify the path from a data and analytics organization to an actual AI factory.

- Stage 1 (Siloed data): access to data sources is limited to a single group within the organization and the data are primarily us-case centric.
-
- Stage 2 (Pilot): a demonstration of the value of analytics can be executed without organizational and cultural changes.
- Stage 3 (Data Hub): the organization has to rearchitect and optimize itself to aggregate and label data from many siloed sources.
- Stage 4 (AI Factory): the organization establishes and adopts a standard operating model for AI, and a major transformation allow the full exploitation of existing data.

III. Modern Data Architecture

A data architecture is composed of "models, policies, rules, or standards that govern which data is collected, and how it is stored, arranged, and put to use in a database system, and/or in an organization". To ensure its effectiveness, designed data architecture should be validated with the unique requirements of end-consumers.

Additionally, Modern Data Architecture (MDA) responds to business demands for speed and agility by enabling organizations to quickly find and unify data across its IT infrastructure. Achieving these needs involves focusing on key principles of MDA:

- **Consider data as a shared asset:** Treating data as a shared asset facilitates reusability. Crucial step in transformation.
- **Address redundancy and eliminate data copies**: Prevalence of data silos and application-centric mindset often leads to duplicate data, increasing the complexity of the data architecture.

- **Describe data with a flexible unified data model**: A common understanding is essential for sharing, interpreting, and processing data throughout the system.

- **Provide the right interfaces for delivering data**: Well-defined and user-friendly interfaces are key for reusability, constituting a crucial technical asset of MDA.

- **Define policies and rules for data governance**: At best, the governance process should be directly linked to the technical implementation.

- **Develop and adopt standards**: The more open the unified data model is, the higher the chance of reusability.

In the context of vehicle software architectures are usually brand-specific and Original Equipment Manufacturers (OEMs) have been developing them focusing on company/brand requirement. To improve data architecture without affecting brand-specific solutions, a common understanding of the data is necessary. Thus, the adoption of standards not only enhances flexibility and scalability but also improves data correctness through collaborative development to address imperfections in data models.

IV. Foundation for a Modern Data Architecture

Enterprises have traditionally adopted an application-centric paradigm in their software architectures, allowing applications to modify data models. However, this approach has been criticized by McComb, who argues it leads to a wasteful allocation of resources in the long run. The author proposes a novel data-centric paradigm based on making data self-describing, with explicit context, and adherence to standards. To address these challenges, one effective approach involves employing the well-established hierarchy of Data, Information, Knowledge, Wisdom (DIKW), as depicted in Figure 2.

This framework serves as a foundation for a good architecture by organizing data in a hierarchical structure, allowing for better management and utilization of information. Taking a bottom-up approach starts by defining a model for what is considered the primary data source in a specific domain. Subsequently, additional components are gradually introduced to enhance the model. In essence, the higher the data is represented in the DIKW hierarchy, the more explicit context is readily available for future use.

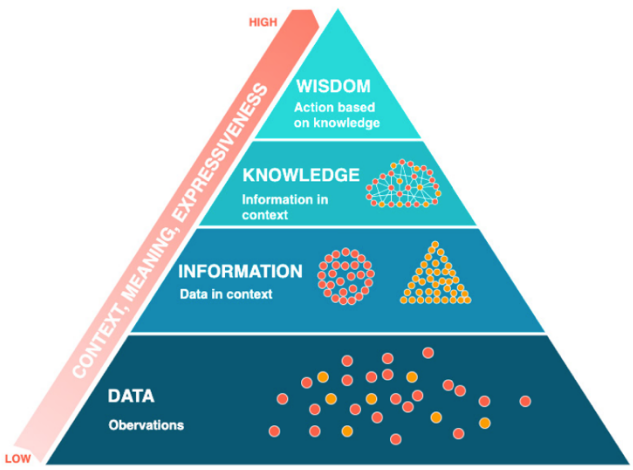

Fig. 2. Data, Information, Knowledge, Wisdom (DIKW) hierarchy [18][19]. Differences between layers are given by the amount of explicit context (i.e., meaning, value, structure, applicability, and expressiveness). Information is defined in terms of domain taxonomies, whereas knowledge covers relationships and interactions between taxonomies from different domains and conceptualize them into a domain ontology.

Using vehicle data as a starting point and the automotive domain in mind:

Toward information: Taking as an example vehicle data. This taxonomy aims to contextualize isolated vehicle observations from sensors using the Vehicle Signal Specification (VSS). This approach establishes a common vocabulary for communicating vehicle data, which has a direct positive impact on performance.

Towards Knowledge: The focus lies on extracting significant value from data integration across diverse sources. The subsequent step involves utilizing the Vehicle Signal Specification (VSS) as a foundational vocabulary for constructing an ontology that encompasses vehicle data. To provide context to vehicle information, a Driving Context Ontology is proposed, specifically designed to make sense of cross-domain driving event data.

Towards Wisdom: Through this layer, there is the capability to formally query the context using the knowledge layer and execute specific actions tailored for particular use cases.

V. Conclusion

Adopting modern data architecture is vital for a digital transformation, while unmanaged complexity poses a major obstacle to achieving an AI factory. The paper proposes that successful implementation of an automotive taxonomy is key to overcome this challenge. The authors think that collaborative approach will be the most successful and that the only way to succeed is to take one step at a time. The first step is to create the right mindset with a framework at hand to understand the reasoning behind it and guidelines on how to proceed. The cornerstones of a modern data architecture for automotive presented in this paper can be applied to facilitate the transformation.