Scientific and Social Research, 2026, Volume 8, Issue 4
http://ojs.bbwpublisher.com/index.php/ssr
Online ISSN: 2981-9946
Print ISSN: 2661-4332

Horizontal Data Middle Ground Design of Traffic Dig Data Asset Management

Ranran Dai*

Suzhou College of Information Technology, Suzhou 215200, Jiangsu, China

Author to whom correspondence should be addressed.

Abstract: How to open up the data center of the task-related industry and build it according to a unified standard to achieve the goal of technology cost reduction, application efficiency improvement, and business empowerment is the problem faced by big data asset management. This paper proposes a horizontal data platform in the background of the traffic big data industry application, trying to solve the problem of data sharing across different industries. Its core includes two aspects: one is the technical ability to apply data, and the other is the ability to provide a data service platform.

Keywords: Traffic big data; Data asset management; Horizontal data platform; Data business; Cross-industry data sharing

Online publication: May 25, 2026

1. Introduction

When it comes to big data, one will definitely think of the Hadoop ecosystem. The Hadoop open source ecosystem has better solved the problem of distributed computing and storage of big data. It can also be said that IaaS for big data is solved. But in addition to IaaS, big data applications need to understand the business and provide PaaS services. It can implement data asset management, such as data collection, cleaning, processing, and governance. Only in this way can the data business be truly realized. At present, there is a big gap between IaaS and PaaS. On the one hand, this gap needs products with large computing power to help improve the efficiency of data processing; on the other hand, it also needs experts who understand business to make use of such big computing tools for the implementation of data business.

How to make up for this gap, Ali first proposed the concept of the data middle ground. Because Ali has a large number of business branches, such as: Taobao, Tmall, etc, each system has its own system and data source, all of which do a lot of services on their own systems, but there is a lack of data sharing between these systems, If one need to dig deep into a certain information, but one cannot query it in the system, one need to use another system to check, that is to say, the data cannot be shared between each system, lead to

low efficiency. Each system also has conflicts between functions and data, between services and applications.

In order to solve these problems, Ali began to integrate mining data to create a data middle ground, from the beginning, only to do data monitoring and statistics, to later data operation and analysis, to search personalized, customized marketing, and then to intelligent. Gradually, the various systems merge together and establish a unified system. Even if the business is expanded, it can be easily incorporated into the middle ground and operate with the same technology and model [1–2].

When there is no data middle ground, no matter whether it is Ali interior or each businessman, everybody has their data center, computer room, small database. However, when the data is accumulated to a certain volume, the cost in this aspect will be very high, as the quality and standard of the data are different, which will lead to inefficiency and other issues. Therefore, through the data middle ground technology, the massive data is collected, calculated, stored, and processed; meanwhile, the standard and caliber are unified [3].

After the data of the Ali system is unified by the data middle ground, standard data will be formed, and then stored to form a big data asset layer, so as to ensure efficient services for all businesses and merchants of the group.

The data middle ground defined by Ali is intrinsically different from the data middle ground discussed in this article. Ali is concerned with the data-sharing problem of vertical services. For the sake of distinction, Ali's data middle ground is a vertical data middle ground. As discussed in this article, the purpose of the data middle ground is to integrate data sharing problems between different industries, so it is called the horizontal data middle ground.

Before constructing the data middle ground in the review, the article first examines the nature of big data.

2. Essence of big data

Big data can tell decision makers some potential rules, and prove or judge decisions with data. In the past, people used data to prove that their decision was right or wrong. Now, data is used to guide people to make the right decision [4–5]. In the era of big data, the sample is the whole, and it can prevent counterfeiting and deviation.

General characteristics of big data have been much discussed, as shown in **Figure 1**, which uses "cross" to describe the sum to characterize big data.

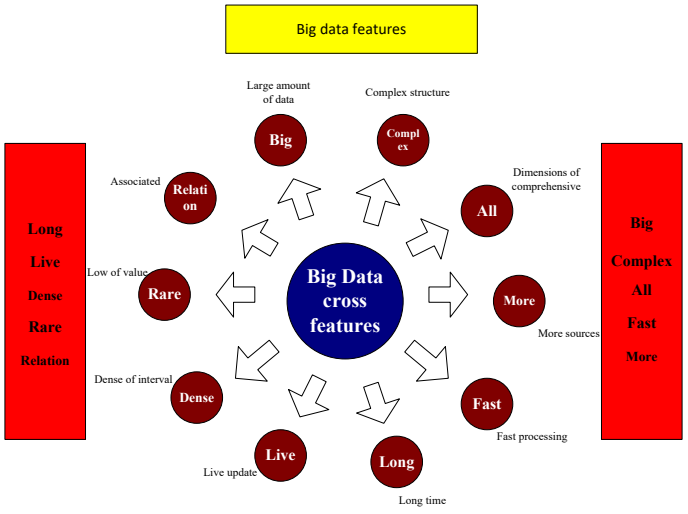

Figure 1. Big data features

This cross feature is easy to understand except for the word "long." "Long" means that the time range of big data is long enough. In general, the longer the time, the more one can find periodic changes in things. For example, the global economic crisis has a cycle of about 10 years, almost every ten years or so. If the time span of the collected data is too short, it is unlikely that such periodicity will be found in the data.

It is the current mainstream to depict big data with 4V, as shown in **Figure 2**.

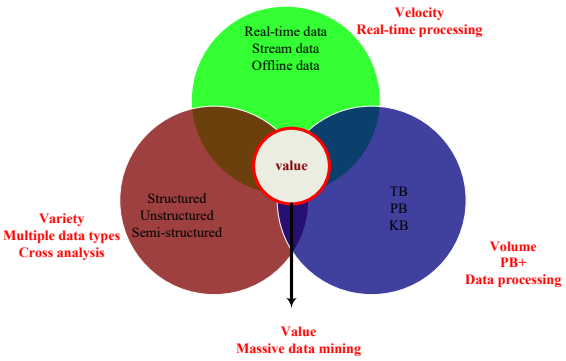

Figure 2. Big data "4V" feature

With so many features, what is the most essential feature? Although the amount of data is large, it only does one kind of thing (getting news, playing games). The data structure is simple, highly repeatable, and just grouped, which has nothing to do with big data. With the increase of data, data fragmentation, logic chaos, and low mining degree make it impossible to form big data, no matter how large the number is and how complex the structure is. Therefore, a high correlation degree of data is one of the essential characteristics of big data.

A classic example often cited in big data is the "pizza case", which covers information in the fields of medicine, banking, transportation, telecommunications, etc. This example shows that true big data is borderless, real-time, cross-industry data sharing [6–7]. Therefore, the essence of big data is the numerous factors, interconnection, and large amounts of real-time data.

Figure 3 shows the causes of traffic congestion at intersections [8].

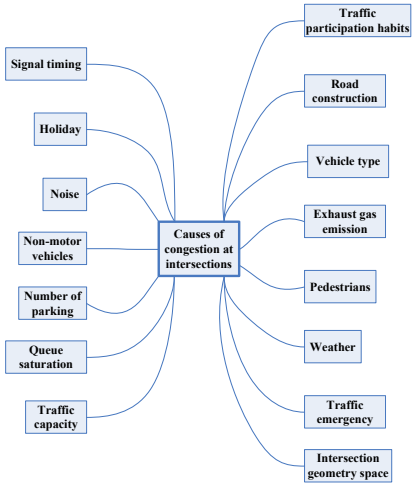

Figure 3. Causes of traffic congestion at intersections

As can be seen from **Figure 3**, traffic congestion events involve weather, municipal, humanities, transportation, environmental protection, insurance, and other multi-domain information. Only by properly managing information shared in various fields can traffic incidents be effectively handled. This is a typical application scenario for horizontal data middle ground.

3. Horizontal data middle ground design of traffic dig data asset management

Data is the foundation of artificial intelligence. Data is an asset in the Internet era. It is the foundation of artificial intelligence. Data assets can form new data through management, which gives us unlimited creative imagination.

Horizontal data middle ground is a data asset management platform that undertakes technology, leads business, builds specifications, and can be connected and extracted globally. The goal is to efficiently meet the demand of real-time processing of foreground traffic events. Its structure can be divided into three layers: data model layer, data management layer, and data service layer, as shown in **Figure 4** [9].

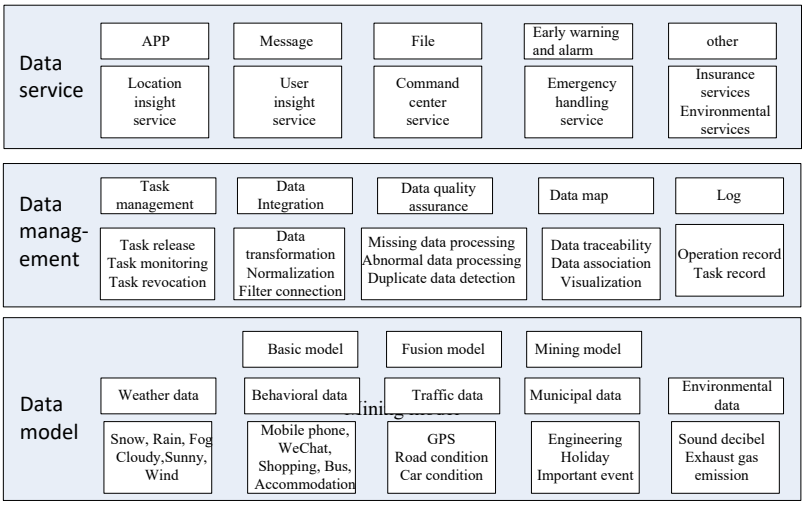

Figure 4. Horizontal data middle ground architecture for traffic big data asset management

The goal of the construction of a horizontal data middle ground is to make it a part of the traffic management business, and make machine intelligence a decision-making link of traffic event handling, so as to make traffic management intelligent.

Business disposition is the destination of horizontal middle ground data convergence, and is also the basis for horizontal middle ground service capability.

If cross-industry data sharing cannot be realized, whether it is a topic, a report or extract the data, the current basic chimney data production mode or project construction mode will inevitably lead to the lack of precipitation and unsustainable development of data knowledge, as a result, models cannot truly become reusable components and cannot support the rapid response and innovation of data analysis [10].

(1) Data model layer. The task of the data model layer is to collect cross-domain data synchronously based on tasks. The data model adopts the data warehouse model, including three models: the basic model is generally relational modeling, which mainly realizes the standardization of data; the fusion model is

generally dimensional modeling, which mainly realizes data concentration and synchronization and data association; the mining model is partial to application.

The unified data model makes a good aggregation of data in related business fields and solves the demand for data intercommunication, which is of great significance.

(2) Data management layer. The data management layer is the core of the entire architecture, the purpose is to achieve data business. If the vast amount of traffic data is compared to the soil of data, then the role of the data management layer is not simply to pile up the sand into mountains, and then provide a set of tools to "gold rush" when needed. The most important task of the data management layer is to establish a reliable "data tree" for the transportation department. Nutrients are extracted from the soil of the original data, and the branches of the data are reorganized according to the basic objects of the business, thereby providing clear and reliable business entities for the upper decision-making, and quickly, accurately, and easily support the subject data analysis and new business capabilities.

The essence of data integration is to organize data scattered in multiple fields separately according to the association relation of data and the inherent basic business object logic.

(3) Data service layer. With the deepening of traffic big data operation, various big data applications are emerging in an endless stream, and there is an urgent demand for data services. If big data is not serviced, it cannot be scaled.

4. Horizontal data middle ground application

Figure 5. shows the field of data dependence in the data from the perspective of data sharing.

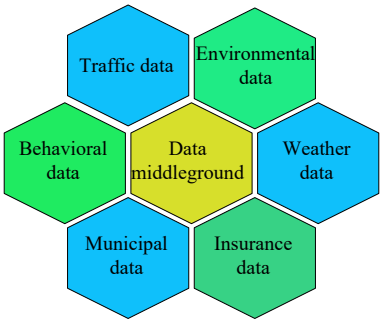

Figure 5. Data middle ground from a data sharing perspective

Figure 6 shows the services provided in the data middle ground from the data service.

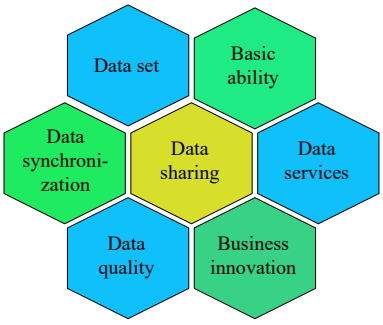

Figure 6. View the data center from the data service

(1) Data set. Refers to the collection of business data.
(2) Data synchronization. With business data from all walks of life, you can organize the required data synchronously according to different tasks.
(3) Data quality. Because the data of different industries may be inconsistent, data quality assurance is necessary to standardize the synchronized data.
(4) Basic ability. Be able to query, track, and apply the processed data.
(5) Data service. An application of data.
(6) Business Innovation. Develop new applications for existing data.

5. Conclusion

The challenge brought by the data era is not only the explosive growth of data volume, but more importantly, how to manage, govern, and make good use of this data. Obviously, the traditional big data construction methodology cannot meet the demand.

If the construction of big data is compared to the construction of tall buildings, then the big data platform development is a pile driver, an excavator, a bulldozer, a tower crane, and so on. The data middle ground is the floor plan, the house design, the house construction, etc., and decoration is a data business.

In the future, based on the era of data, there may be many company platforms. Now several major Internet giants are doing their own platforms, but more for serving the enterprise itself. However, due to their small size, startups lack both customers and the ability to aggregate data, which requires the enabling of a horizontal data middle ground. How to do a good job in the construction of horizontal and vertical data middle ground, and better empower the small and medium-sized enterprises in the entire industrial chain. It is a work that will be explored in the next few years, providing a one-stop PaaS.

Funding

2021 Education Reform Project of Suzhou College of Information Technology, “A Practical Study on Improving Classroom Teaching Effectiveness: A Case Study of the ‘Computer Network Security Technology’ Course” (Project No.: 2021YJ02)

Disclosure statement

The author declares no conflict of interest.

References

[1] Gu J, 2023, Design of Intelligent Traffic Visualization Platform Based on Big Data Architecture. Advances in Computer and Communication, 4(3): 172–176. https://doi.org/10.26855/acc.2023.06.013

[2] Yang DY, Reflection on the Construction of Urban Traffic Big Data Analysis Platform. Communication & Shipping, 2019(1): 58.

[3] Wang YH, Ban, XG, Cui ZY, 2019, An Artificial Intelligence Platform for Network-wide Congestion Detection and Prediction using Multi-Source Data. University Transportation Centers US Transportation Collection.

[4] Respatia S, Bhaskarb A, Chungc E, 2018, Traffic Data Characterisation: Review and Challenges. Transportation Research Procedia, 2018(34): 131–138. https://doi.org/10.1016/j.trpro.2018.11.024

[5] Lee C, Lee S, 2022, Exploring the Contributions by Transportation Features to Urban Economy: An Experiment of a Scalable Tree-Boosting Algorithm with Big Data. Land, 2022(11): 577. https://doi.org/10.3390/land11040577

[6] Yuan Q, Liu Z, Li J, et al. 2014, A Traffic Congestion Detection and Information Dissemination Scheme for Urban Expressways using Vehicular Networks. Transportation Research Part C, 47(Oct.): 114–127. https://doi.org/10.1016/j.trc.2014.08.001

[7] Chen LW, Chang CC, 2017, Cooperative Traffic Control with Green Wave Coordination for Multiple Intersections Based on the Internet of Vehicles. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47(7): 1321–1335. https://doi.org/10.1109/TSMC.2016.2586500

[8] Chow A, 2023, Nature and Causes of Urban Traffic Congestion — A Data Analytic Approach, thesis, City University of Hong Kong. https://www.hkuits.hku.hk/2018-conference/wp-content/uploads/presentations/Andy%20CHOW.pdf

[9] Zhang C, Zhu B, Wang J, et al., 2023, Development and Application of Shanghai Big Data Platform for Transportation Planning. Shanghai Urban Planning & Design Research Institute, published online. https://www.chinautc.com/upload/accessorychinautc/20241/202418932181049885.pdf

[10] Zhu Y, Liu X, Wang Y, 2013, Pervasive Urban Sensing with Large-Scale Mobile Probe Vehicles. International Journal of Distributed Sensor Networks, 2013(5): 177–182. https://doi.org/10.1155/2013/762503