**Data Governance –**

**For your selected company describe the data governance mechanism that you**

**implement it? Explain your rationale.**

Today, most retail companies have greatly improved the way they collect data. As retail enterprises pay more attention to data, a large amount of enterprise data will accumulate. Data governance includes not only how to effectively revitalize and use data to meet the specific business scenario needs of each business department, but it also tremendously impacts management of information assets and drives initiatives towards the unified enterprise view of information. Retail companies like grocery chains have as few as three to five systems and as many as dozens of systems. However, each system is often divided into multiple suppliers, and the internal data of the company is also isolated. Therefore, the big question that needs to be answered is “How do we effectively use data to meet specific business scenario needs for each business department?” At the same time, it is imperative to avoid repeated development, and wasting funds has become a problem faced by the information and data departments. The key points of our data governance system design are that data stays relevant, up-to-date, and consistent across the multitude of sales and marketing platforms.

In general, the first step is to **collect and standardize enterprise data**. The importance of data governance is self-evident. For most retail companies, due to the complex internal data structure, each specific application needs to be cleaned by the service provider, which is time-consuming and labor-intensive. This means the efficiency is very low at the first step. However, standardized governance of the enterprise’s internal data is the first step in effective utilization. There is a saying in the field of informatization — "garbage in, garbage out" — which means that in the instance that we use messy data, the research results produced become worthless as well. Data pollution may occur in various links such as data generation, collection, transmission, circulation, processing, storage, extraction, and exchange. Therefore, to ensure the realization of data governance goals, it is necessary to control the entire process of data. Data quality, data file exchange, metadata, data life cycle, primary and secondary data sources, data security, data responsibility, etc. form a unified data governance specification. Although our group split up the work to create the different OLTP systems, we came together to ensure that everything was consistent and no unnecessary or duplicate data was added, and we discussed the ways that the systems could be connected on different tables and fields so data could flow easily through the systems.

The second step is **data model management and label combing**. The core of data governance is data model management. At present, there are many fields and tables in the original enterprise database without annotations, and the meanings become ambiguous. Fields and tables with the same name can have different meanings, while others with different names may have the similar or identical connotations. This phenomenon of inconsistency with redundant fields and enumerated values ​​is common, and these problems will directly affect the system's recognition of data. For example, in both the **sales** and **inventory** OLTPs, there was as store table though one was called **store** and the other called **location.** These two tables are essentially the same thing, a table that can be used to link the OLTP systems but causes confusion due to lack of communication and mislabeling. Data modeling makes the data structure richer and clearer, facilitating the unification of data caliber. After an enterprise has accumulated many data models, it is necessary to perform retention or deletion management in a timely manner, instead of data model management and control. Then, these historical problems will cause a lot of trouble for the new generation of system transformation.

The third step is the **application of enterprise algorithms and artificial intelligence**. After the collection and standardization of enterprise data, and the control and labeling of the data model, the data can be managed, supplemented by corresponding algorithms and artificial intelligence, and applied in specific business scenarios. Take data model management as an example: artificial intelligence can help companies achieve the perfect integration of experience models and computer models and build a knowledge map of products and members.