**Literature Review**

**Introduction**

AliExpress is an online e-commerce wholesale website established by Alibaba which is the largest online retailer in China. It provides international online shopping services around the world and builds a cross-border platform to connect sellers and consumers without face-to-face transactions. As the number of users on the website has soared dramatically in recent years, the big data technologies that support the transaction and security of AliExpress are also evolving to meet companies' demands. In this project, we decided to investigate and introduce the big data technology used in AliExpress or Alibaba group to generate insights from its technological development.

**The evolution of big data technology**

AliExpres belongs to the Alibaba group, which also owns another two famous online sites Taobao and Tmall in China. The transaction volume shown in the graph (Figure 1) illustrated that Alibaba’s business growth was facing a data explosion problem.

A screenshot of a cell phone

Description automatically generated

**Figure 1**: Transaction volume increases at Alibaba.

Some researchers in Alibaba enterprises stated that the company optimized its big data processing system to improve the performance of processing large amounts of data. The company’s data platform was changed from the Oracle application cluster in 2009 to ODPS in 2015, and the data processing time was shortened to near real-time. The new ODPS (MaxCompute) system constructs a data warehouse platform to manage massive and various data in a distributed way. Furthermore, the new data processing system helps the company solve the obstacles regarding data volume and data streaming, which are difficult to handle by the previous Oracle data platform.

A close up of a sign

Description automatically generated

**Figure 2**: Big data computing process.

Figure 2 illustrated that the ODPS data system assisted the company team in solving several issues. The first problem was the data storage capability, which increased from 100TB in 2009 to 100EB in 2014. With the solution to various data sources, data processing time had also decreased dramatically. We can found that the new generation processing system boosts the data storage and data calculation performance of massive data(Chen et al,2015).

**The recommender system in Aliexpress using big data technology**

Because of the extraordinary performance improvements on big data technology, The subsidiary company of Alibaba, AliExpress, is capable of using machine learning running on the big data processing system. The company collects large amounts of personal data from registered customers to build a recommendation system and it also sells online products based on these users' preferences. The two authors Lukucheva and Semenovich (2019) conducted a survey of the big data technology used in the AliExpress recommendation system to determine their advantages in enhancing the loyalty of Russian customers buying goods from online retailers. The survey shows that in order to build a recommendation system, customers on AliExpress are required to provide their data and collect these data during the usage of AliExpress. The company utilizes big data tools to analyze the user's personal information, such as likes, comments, or browsing information, etc and then executes the recommended algorithm to predict the user's hobby or preferences on similar goods. In this case, big data tools can also help customers personally find what they like, and provide them with better services when using online shopping sites. Furthermore, after using this big data analyzing technology, the revenue of Aliexpress in the Russian e-commerce market had also increased since 2010(Zhou et al, 2017).

**References**

Lukicheva, T., & Semenovich, N. (2019, November). Big Data as a success factor of AliExpress in the Russian market: advantages and opportunities as seen by the eyes of consumers. In *Third International Economic Symposium (IES 2018)*. Atlantis Press.

Chen, J., Tao, Y., Wang, H., & Chen, T. (2015). Big data based fraud risk management at Alibaba. *The Journal of Finance and Data Science*, *1*(1), 1-10.

Zhou, J., Li, X., Zhao, P., Chen, C., Li, L., Yang, X., ... & Qi, Y. A. (2017, August). Kunpeng: Parameter server based distributed learning systems and its applications in alibaba and ant financial. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1693-1702).