



## Designing Courier Service (*Jastip*) Application by using Spark-based Big Data technology

Ananda Dessi<sup>1</sup>, Fauzan Nasafi<sup>2</sup>, Gresshinta<sup>3</sup>, Gunawan Wang<sup>4</sup>

<sup>1,2,3,4</sup> Information Systems Management Department, BINUS Graduate Program-Master of Information Systems Management, Bina Nusantara University, Jakarta, Indonesia 11480.

<sup>1</sup>ananda.budiansa@binus.ac.id, <sup>2</sup>fauzan.nasafi@binus.ac.id, <sup>3</sup>gresshinta@binus.ac.id,

<sup>4</sup>gwang@binus.edu

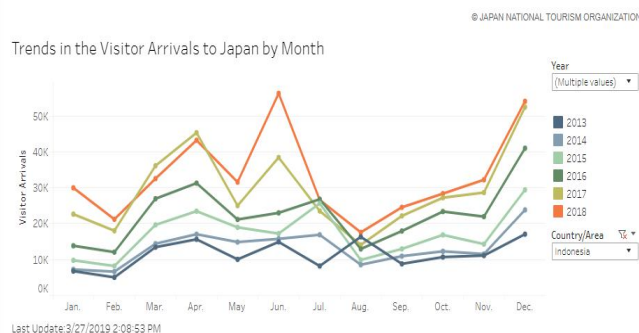
### ABSTRACT

Jastip is an informal service that offers assistance to people in need or want to buy something but cannot go to the desired place to buy themselves for various reasons. Those who offer help will give you a service value for each item offered for them to buy. This phenomenon is a new business model that has the potential to be developed. Based on this phenomenon, it is necessary to design an application that can accommodate the phenomenon of Jastip. This application will design as a mobile application using Spark based Big Data technology. StealDeal is an online service application through a mobile application that provides discount product recommendations on websites that exist in Japan. In this application object design service only issues a test object design using a large data processing method in the system architecture design. Focus point of this design is the architectural design.

**Key words:** Jastip, Big Data, Mobile Application, Spark.

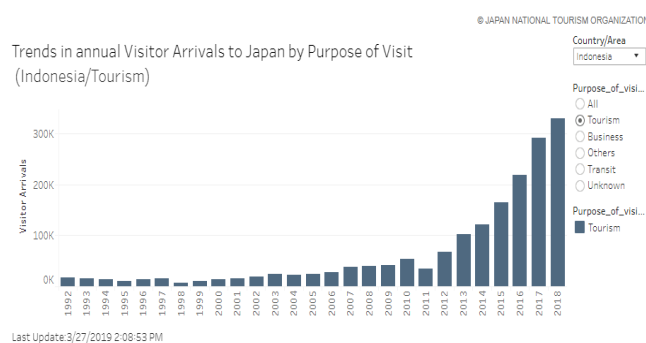
### 1. INTRODUCTION

Jastip is an informal service that offers assistance to people in need or want to buy something but cannot go to where they want to buy for themselves for various reasons [1]. Jastip phenomenon is a new business model that has the potential to be developed [1]. With the physical phenomenon it is possible to achieve the locations that are often visited by the people of Indonesia. Every year there is an increase in visits from Indonesian to Japan [6].



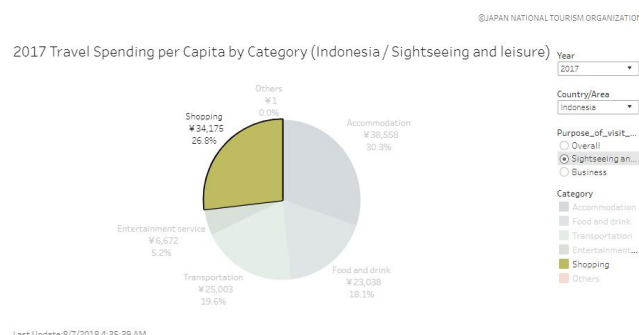
**Figure 1:** The annually visited trend of Indonesian annually.

The above data shows that an average increase in visits occurs every month, especially in June and December. From the visit of the Indonesian, it is also known that every year is an increase visits with the purpose of travel [6].



**Figure 2:** The annually visited trend of Indonesian based on the purpose of visiting.

On the visit, it is known that 26.8% of Indonesian public expenditure is made to make purchases [6].



**Figure 3:** Indonesia's tourist spending trend in Japan in relation to spending

The trend of traveler spending related to the spend, only through the spending rate by the cost of accommodation incurred by Indonesian tourists in Japan.

Therefore, it is necessary to do the design of the application that can accommodate the phenomenon of the Jastip, which is supported by big data technology. Especially related to the utilization of big data in terms of information search and recommendation system. There are four positive factors in

implementing of Big Data Analytics (BDA) that can be beneficial for the company, including: (1) Information search. (2) Recommendation System. (3) Dynamic Pricing. (4) Customer Services [2]. The utilization of big data in terms of recommendation system has been implemented by Netflix, eBay, Amazon, monsters and other retail stores which have all been recommended [2]. Other than that, in terms of information search, the speed and convenience of collecting information is one of the benefits for customers when they do online shopping [2].

## 2. LITERATUR REVIEW

### 2.1 Jastip

Jastip is an informal service that offers assistance to people in need or want to buy something but cannot go to the desired place to buy for yourself for various reasons [1].

### 2.2 Big Data

Big Data is a large and complex data set, with various kinds of data. Big Data can be divided into structured and unstructured data and have five V as different characteristics (Volume, Velocity, Value, Variety and Veracity) [3].

### 2.3 Positive Factor in Big Data Implementation

The positive factors of application of Big Data analysis applications included offering Information retrieval, recommendation systems, dynamic pricing, and customer service to interact with member communities [2].

- Information Search: Emotionally driven consumers easily drive their wishes and purchase demand with information networks. The speed and convenience of collecting information is one of the benefits for customers when they do online shopping. With analytics tool, the website can filter and search various kinds of data and present it to customer. And with text miner technology, the website can search and record the relevance history from the libraries, catalogs. It is all about relevancy and offering the right product or service to the right person at the right price through the right channel at the right time.
- Recommendation System: For the recommendation system, involves relationship between customers and website. Customers provides their personal preferences information, such as hobbies, needs, while website offer the recommendations based on customer informations. With recommendation system, it can help customer to choose product or service can meet their needs and expectation [4].
- Dynamic Pricing: With Big Data, the website can give personal price to the customer based on the location, product or time data of customer.

- Customer Service: for customer service, Big Data implementation can help the company to optimize the customer service result in order to make customer happier. It can help company to get the information about customer's level of happiness not only through company official channels but from another media such as social media, forum. Beside that, it can help company to detect and handle the customer complaint faster. At the in the end, it can improve the customer service performance and make customer satisfaction level higher.

### 2.4 Characteristics of Big Data

There are some characteristics of big data [7]:

- Comprehensive
- Enterprise Ready
- Minimal Maintenance
- Integrated
- Extensibility
- Scalability
- Permits ad-hoc queries
- Low latency
- Robust and reliability

### 2.5 The Benefit of Spark

There are some benefit of Spark implementation [8]:

- Comprehensive framework
- 100 times faster in memory, 10 times faster on the disk
- Supports SQL queries and data flow and offers machine learning and graph-oriented processing functions.

## 3. PROPOSED ARCHITECTURE

In designing the service application, using the following Spark-based big data technology is the proposed architecture:

### 3.1 Big Data Analysis Architecture

In designing Big Data Analysis Architecture can be divided into 3 components, namely the preparation of Big Data Pipeline, System preparation that will display data analysis and Data Store results [5].

The following are the architectures compiled related to the Big Data Analysis for the StealDeal application:

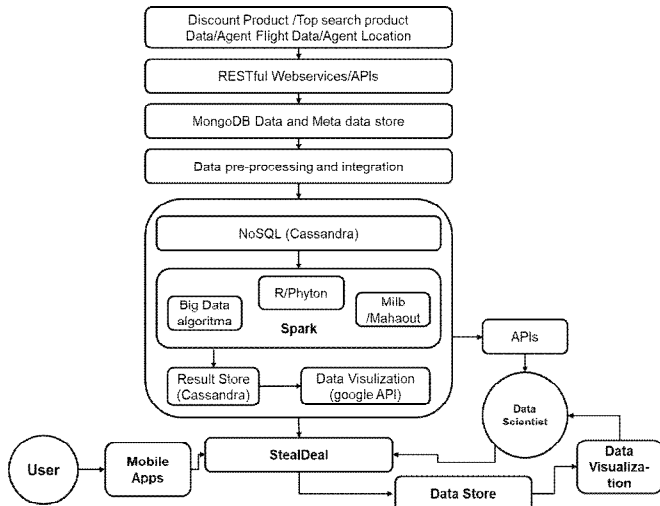


Figure 4: StealDeal BDA Architecture

### 3.2 Big Data Pipeline Arrangement

The main role of the System developer was to build the Big Data Pipeline. The Big Data Pipeline in the proposed architecture is mainly divided into five phases, 1) data extraction, 2) raw data storage, 3) data preprocessing and integration, 4) analytical engine, and 5) data visualization [5].

The following is a workflow diagram of the StealDeal Big Data Pipeline:

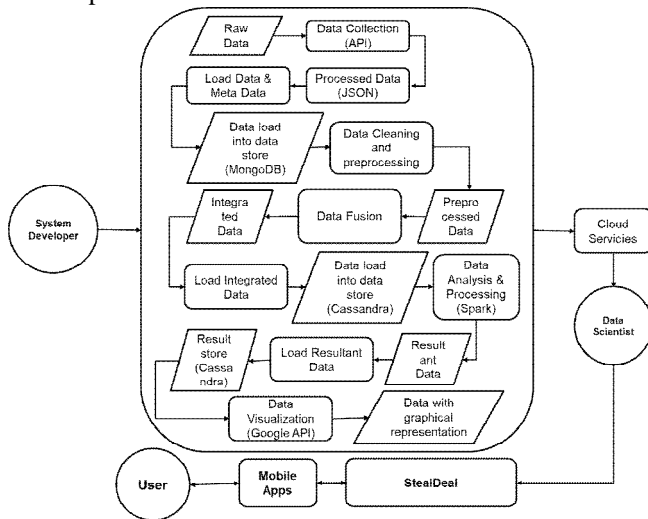


Figure 5: StealDeal Big Data Pipeline

- Data Extraction

Data collected in the stage of data extraction is the product data that is discounted in the webstore that exists in Japan, any product data that is often sought by users through its browser handles, flight schedule data from the agent and the location of the residence Agent in Japan. Data from these sources are retrieved and processed using the API and stored in the Raw Data Store.

- Raw Data Store

To store the semi-structured Data, the best option is Using MongoDB. MongoDB is an opensource database that provides high availability, high performance and automatic scaling in the provision of databases [5].

- Data Processing and Integration

Fusion Data is the process of integrating some Data and knowledge representing the same realworld objects Consistent, accurate and rewarding representation. Raw data stored in MongoDB is integrated into an integrated format in such a way that it can be searched, asked, and analyzed using analytical machines [5].

- Analytical Engine

In the stage of analytical engine is done data processing and analysis of data for information obtained. The Platform used at this stage is using Spark. Spark is the most suitable platform for implementing repetitive machine learning jobs and Interactive analysis. Spark can surface as much as 10x Hadoop in disk analysis and 100x memory in memory. In fact, Spark does not have its own storage system, it can access data from a variety of data sources, including HDFS, S3, Cassandra, HBase and any Hadoop data sources. The Spark app can be implemented using Scala, Python, Java and R. Machine learning algorithms implemented using MLIB and Mahout Libraries [5].

- Data Visualization

At this stage the visualization of information that has been produced in the Analytical Engine stage. These visualizations can be displayed using tools like Google API.

### 3.3 StealDeal

StealDeal is an online service application through mobile apps that provides discount product recommendations on websites that exist in Japan. In addition, StealDeal can also provide product recommendations according to the product that is often sought by user information through the browser that is on his handset. In StealDeal Also, user can choose the agent for the appropriate agent availability agents at the time and place expected by the user.

## 4. CONCLUSION

In the design of this object service application is only a test object of the design issue of the use of big data processing methods in the design of a system architecture. The focal point of this design is the architectural design process. Selection of service applications as a design object due to the needs expressed in the background.

## REFERENCES

1. I. S. Muslich and Irwansyah, **Instagram dan Fenomena Jastip di Indonesia**, pp. 1–5, 2000.

2. T. M. Le and S. Y. Liaw, **Effects of pros and cons of applying big data analytics to consumers' responses in an e-commerce context**, *Sustain.*, vol. 9, no. 5, 2017. <https://doi.org/10.3390/su9050798>
3. Ishwarappa and J. Anuradha, **A brief introduction on big data 5Vs characteristics and hadoop technology**, *Procedia Comput. Sci.*, vol. 48, no. C, pp. 319–324, 2015. <https://doi.org/10.1016/j.procs.2015.04.188>
4. L. Hongyan and L. Zhenyu, **E-Commerce Consumer Behavior Information Big Data Mining**, *Int. J. Database Theory Appl.*, vol. 9, no. 7, pp. 135–146, 2016. <https://doi.org/10.14257/ijdta.2016.9.7.12>
5. P. Shah, D. Hiremath, and S. Chaudhary, **Big data analytics architecture for agro advisory system**, *Proc. - 23rd IEEE Int. Conf. High Perform. Comput. Work. HiPCW 2016*, pp. 43–49, 2017. <https://doi.org/10.1109/HiPCW.2016.015>
6. <https://statistics.jnto.go.jp/en/>
7. B. Manoj, K. V. K. Sasikanth, M. V. Subbarao and V. Jyothi Prakash, **Analysis of Data Science with the use of Big Data**, *International Journal of Advanced Trends in Computer Science and Engineering*, 2018
8. M. Banane and A. Belangour, **Querying massive RDF data using Spark**, *International Journal of Advanced Trends in Computer Science and Engineering*, 2019. <https://doi.org/10.30534/ijatcse/2019/68842019>