International Journal of Advanced Smart Convergence Vol.8 No.2 102-108 (2019)
http://dx.doi.org/10.7236/IJASC.2019.8.2.102

IJASC 19-2-12

A Study on Open API of Securities and Investment Companies in Korea for Activating Big Data

Gui Yeol Ryu

Dept. of Computer Science, Seokyeong University, Korea
e-mail: gyryu@skuniv.ac.kr

Abstract

Big data was associated with three key concepts, volume, variety, and velocity. Securities and investment services produce and store a large data of text/numbers. They have also the most data per company on the average in the US. Gartner found that the demand for big data in finance was 25%, which was the highest. Therefore securities and investment companies produce the largest data such as text/numbers, and have the highest demand. And insurance companies and credit card companies are using big data more actively than banking companies in Korea. Researches on the use of big data in securities and investment companies have been found to be insignificant. We surveyed 22 major securities and investment companies in Korea for activating big data. We can see they actively use AI for investment recommend. As for big data of securities and investment companies, we studied open API. Of the major 22 securities and investment companies, only six securities and investment companies are offering open APIs. The user OS is 100% Windows, and the language used is mainly VB, C#, MFC, and Excel provided by Windows. There is a difficulty in real-time analysis and decision making since developers cannot receive data directly using Hadoop, the big data platform. Development manuals are mainly provided on the Web, and only three companies provide as files. The development documentation for the file format is more convenient than web type. In order to activate big data in the securities and investment fields, we found that they should support Linux, and Java, Python, easy-to-view development manuals, videos such as YouTube.

Keywords: *Big Data, Hadoop, HTS, Open API*

1. Introduction

Korea has 22 major securities and investment companies such as SK Securities Co., Hanwha Investment & Securities Co., Shinyoung Securities Co., Bookook Securities Co., Hanyang Securities Co. Samsung Securities, Kyobo Securities Co., KB Securities Co., Hyundai Motor Securities, Daishin Securities Co., Hi Investment & Securities Co., NH Investment & Securities Co., Kiwoom Securities Co., eBest Investment & Securities Co., Mirae Asset Daewoo Co., Yuanta Securities Co., Eugene Investment Co., Korea Investment

Manuscript Received: Apr. 16, 2019 / Revised: Apr. 21, 2019 / Accepted: Apr. 28, 2019
Corresponding Author: gyryu@skuniv.ac.kr
Tel: +82-2-940-7752, Fax: +82-2-919-0345
Dept. of Computer Science, Seokyeong University, Korea

& Securities Co., KTB Investment & Securities Co., Hana Financial Investment Co., IBK Securities Co., and Cape Investment & Securities Co.. These companies are engaged in the business of brokering services of stocks, direct investment using assets held, investment bank (IB) such as M&A, initial public offering (IPO), new issue of stock, and capital increase by issuing new stocks.

Securities and investment companies offer Home Trading System (HTS) which allows individuals to trade stocks anytime, anywhere. Users analyze and sell stocks using the HTS system developed by securities and investment companies. They develop HTS and recommend investment using artificial intelligence engine. These systems report the stock status using the screen developed by them. This approach can be changed in the big data era.

Large data before big data were extracted and used in database. The software that manages the database is a database management system (DBMS). DBMS is suitable for data retrieval, and is costly because it is commercial software. In the database age, it was the right way that stock investors use DBMS. The limitations of this system are that investors fail to accumulate data directly, and they would follow recommendations of securities and investment companies, or make personal subjective judgments. Even if data is accumulated, it takes much time to analyze it. So it is difficult to utilize it. These limitations can be overcome in the big data era. Since big data is basically saved as text, we can develop the programs that can make the calculations that investors want in almost real time. In addition, software related to big data are open source and can be used free of charge. You only need to reserve storage space. To build big data systems, Hadoop is the storage management software. Hadoop distributes large files and fetches scattered data at developer's request. This system is highly scalable and has no difficulty in storing and reading large amounts of data. You need a program to read and analyze stored files, which is a MapReduce program developed by Java. The MapReduce program is divided into the Map programming, where data is composed of key and value, and the Reduce program which implements the program for analysis purpose. These programs allow you to view the results in almost real time. Because it is implemented in Java, it can be implemented free of charge. R can be used to implement more diverse analytical models and visualizations. R is also free. The storage problem can be solved with a cloud service at a given cost. Therefore, big data systems can implement powerful programs almost free of charge. Stock data changes in real time and produces a large amount of data. Investors make decisions based on screens, analysis results, or investment recommendation provided by securities and investment companies. To use their own investment analysis or method, you must use Excel or other application software. It must be determined after analyzing. So there is a time difference. If there is a time difference in stock investment, it becomes a stumbling block to activation. Therefore, real-time analysis and investment decisions are important factors in stock investment. Big data provides a tool to solve these problems. In this paper, we analyze the data format of open API provided by securities and investment companies in Korea and study whether they are suitable for big data.

2. Analysis of big data and open API of securities and investment companies

2.1 Characteristics of big data

Various organizations are defining big data. Gartner (2012) defines Big Data in the following. "Big data" warrants innovative processing solutions for a variety of new and existing data to provide real business benefits [1]. But processing large volumes or wide varieties of data remains merely a technological solution unless it is tied to business goals and objectives. McKinsey Global Institute (2011) defines Big Data as datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze [5]. Wikipedia (2019) defines Big Data as follows [8]. "Big data" is a field that treats ways to analyze, systematically extract information from, or otherwise deal with data sets that are too large or

complex to be dealt with by traditional data-processing application software. Data with many cases (rows) offer greater statistical power, while data with higher complexity (more attributes or columns) may lead to a higher false discovery rate. Big data challenges include capturing data, data storage, data analysis, search, sharing, transfer visualization, query in, updating, information privacy and data source. Big data was originally associated with three key concepts: volume, variety, and velocity. Other concepts later attributed with big data are veracity (i.e., how much noise is in the data) and value. Big data is applied to various fields as in the paper of Kim (2018), and Lee and Lee (2013) [2, 7]. Cho (2013) analyzed the various issues in Big Data [10].

Volume should handle more than several petabytes in size, which is difficult to manage with existing database management system. Variety refers to various types of data include text / numbers, images, audio, and video. Data structures include structured data used in R-DB, semi structured data including XML and HTML data, and unstructured data not stored in fixed fields. Velocity means to analyze and process data in almost real time. Big data has various application areas such as Chang and Ko (2019) [9]. The McKinsey Global Institute (2011) categorizes data produced and stored by industry sector as shown in Table 1 [5]. Securities and investment services produce and store a large data of text / numbers.

Table 1. The type of data generated and stored varies by sector

Sectors	Video	Image	Audio	Text/numbers
Banking	Medium	Medium	Medium	High
Insurance	Low	Low	Low	High
Securities and investment services	Low	Low	Low	High
Discrete manufacturing	Medium	Medium	Low	High
Process manufacturing	Medium	Medium	Low	High
Retail	Medium	Low	Low	High
Wholesale	Low	Low	Low	High
Professional services	Medium	Medium	Medium	High
Customer and recreational services	Medium	Low	Medium	Medium
Health care	Low	High	Low	High
Transportation	Medium	Medium	Low	High
Communications and media	High	Medium	High	High
Utilities	Medium	Medium	Low	High
Construction	Low	Medium	Low	Medium
Resource industry	Medium	Medium	Low	High
Government	High	Medium	High	High
Education	High	Medium	High	Medium

The KB management Institute (2012) reported average data holdings per company per industry in the US based on IDC and McKinsey & Company data, as shown in Table 2 [3]. Securities and investment services have the most data per company on the average.

Table 2. Average data per company per industry in the US

Industry	Data volume (unit: terabyte)
Securities and investment services	3,866
Banking	1,931
Communications	1,792
Devices	1,507
Manufacturing	967
Insurance	870
Raw materials	825
Transportation	801
Distribution	697

2.2 Applications of Big data for Securities and investment services

According to McKinsey Global Institute (2011), banking services and securities and investment services produce and store a large amount of data in the form of text/numbers [5]. According to Kim and Hur, Kim (2013), Gartner revealed that the demand for big data was 25% for financial services, 15% for services business, 15% for manufacturing, 2% for government, 11% for education and 7% for healthcare. Results are shown in Table 3 [4]. Seo (2013) reported that big data is used in banks, insurance companies and credit card companies in Korea [6]. Insurance companies and credit card companies are using more actively than banking companies. However, researches on the use of big data in securities and investment companies have been found to be insignificant. In the securities and investment services, a large amount of text data for stocks is produced daily. The securities and investment companies display data on its HTS, provide their analysis results, and use artificial intelligence (AI) to advise investors on investment.

Table 3. Demand for Big Data

Industry	Ratio of Demand
Financial services	25%
Services business	15%
Manufacturing	15%
Government	2%
Education	11%
Healthcare	7%

Big data can be used as a base technology for system trading which can make users invest by their own trading rules. System trading is a method of trading that raises the return on investment by eliminating arbitrary judgments and prejudices, and using consistent trading rules. System trading can be implemented based on individual knowledge and creativity. Big data is needed to implement system trading. Big Data, therefore, serves as an infrastructure to support investors' creativity in the stock market. In order for investors to build big data systems, they must receive and store real-time data for stock from securities and investment companies. Investors can obtain almost real-time stock data through open APIs provided by them. Therefore, open API provided by them is an important issue. From now on, we will analyze the open API supported by 22 major securities and investment companies in Korea.

2.3 Analysis of open API supported by major securities and investment companies

In the securities and investment sector, investors must receive and store data through the open API in order to utilize big data. We will analyze the open API supported by major securities and investment

companies to support big data. There are 22 major securities and investment companies in Korea. HTS is 100% supported. But there are only 6 companies that support open API, only 27%. Table 4 shows the HTS and open API support status of 22 securities and investment companies in Korea.

Table 4. open API supported by 22 major securities and investment companies in Korea

Securities and investment services	HTS	Open API
SK Securities Co.	Yes	No
Hanwha Investment & Securities Co.	Yes	No
Shinyoung Securities Co.	Yes	No
Bookook Securities Co.	Yes	No
Hanyang Securities Co.	Yes	No
Samsung Securities	Yes	No
Kyobo Securities Co.	Yes	No
KB Securities Co.	Yes	No
Hyundai Motor Securities	Yes	No
Daishin Securities Co.	Yes	Yes
Hi Investment & Securities Co.	Yes	No
NH Investment & Securities Co.	Yes	No
Kiwoom Securities Co.	Yes	Yes
eBest Investment & Securities Co.	Yes	Yes
Mirae Asset Daewoo Co.	Yes	No
Yuanta Securities Co.	Yes	Yes
Eugene Investment Co.	Yes	No
Korea Investment & Securities Co.	Yes	No
KTB Investment & Securities Co.	Yes	No
Hana Financial Investment Co.	Yes	Yes
IBK Securities Co.	Yes	Yes
Cape Investment & Securities Co.	Yes	No
Total	22	6

From now on, we will analyze the open APIs of the six securities and investment companies. The operating system (OS) of the client is only Windows. They do not support Linux which is used in big data. Supported data models are Component Object Model (COM), OLE Control Extension (OCX), and Dynamic Link Library (DLL), which are all defined by Microsoft. COM is a binary-interface standard for software components introduced by Microsoft in 1993. It is used to enable inter-process communication object creation in a large range of programming languages. OCX is a proprietary technology developed by Microsoft that allows embedding and linking to documents and other objects. OLE (Object Linking & Embedding) is a proprietary technology developed by Microsoft that allows embedding and linking to documents and other objects. DLL is Microsoft's implementation of the shared library concept in the Microsoft Windows and OS/2 operating systems. These libraries usually have the file extension DLL.

The programming languages of open API are mainly VB, Excel, VC ++, C #, VB.net, and MFC, which are provided by Microsoft. In addition, Daishin Securities Co. supports Python. Python is a good language for receiving and analyzing web information. So it is useful for analyzing and merging with web information. eBest Investment & Securities Co. supports Java. Java is a programming language that implements the big data platform, which is advantageous for linking with big data.

In order to use the open API, developers should follow manual. 6 securities and investment companies provide manuals through a web browser. Kiwoom Securities Co., eBest Investment & Securities Co., and Yuanta Securities Co. provide the files for manual. The file type manual is easy to develop software. In particular, Kiwoom Securities Co.'s manual is very helpful. Manuals of Kiwoom Securities Co., and eBest Investment & Securities Co. are easy to read. Manuals of Yuanta Securities Co. and Daishin Securities Co. are readable. And manuals of IBK Securities Co., and Hana Financial Investment Co. are difficult to read. Therefore, to make the manuals easier to read, it is desirable to be provided as files. Video tutorials using YouTube are also needed. The results are summarized in Table 5.

Table 5. Characteristics of open API of securities and investment companies

Securities and investment services	OS	Data Controller	Supported Language	Manuals Method	Degree of explanation
Daishin Securities Co.	Windows	COM	VB, Excel, VC++, Python, etc	Web	Middle
Kiwoom Securities Co.	Windows	OCX	VB, Excel, MFC, etc	Web, File	High
eBest Investment & Securities Co.	Windows	COM, DLL	Delphi, Java, Excel, C#, VB, VB.net, C/C++, etc	Web, File	High
IBK Securities Co.	Windows	COM	VB, VC++, Delphi, Access, etc	Web	Low
Hana Financial Investment Co.	Windows	OCX	MFC, C#, VB, Delphi, Excel, etc	Web	Low
Yuanta Securities Co.	Windows	COM, DLL	VB, C#, Delphi, C/C++, etc	Web, File	Middle

Open APIs provided by securities and investment companies require that developers receive data by Windows. This is a limitation that data cannot be directly received by a big data system. Real-time analysis and decision making are important for system trading. And big data System is useful for analyzing stored data in almost real time. Big data is therefore an important infrastructure for system trading. However, in order to apply big data system to system trading, it is necessary to receive data from Windows and upload data to the Hadoop platform of Linux. Five securities and investment companies, excluding Daishin Securities Co., do not support Java. Java is a programming language for analyzing big data. These problems are limits to real-time analysis and decision-making. Therefore securities and investment companies provide open APIs which support Linux and Java for activating big data and system trading.

3. Conclusion

Big data was originally associated with three key concepts, volume, variety, and velocity. Other concepts such as veracity are later attributed. The McKinsey Global Institute (2011) categorizes data produced and stored by industry sector, shows that securities and investment services produce and store a large data of text/numbers. The KB management Institute (2012) reported average data holdings per company per industry in the US. Securities and investment services have the most data. Gartner found that the demand for big data in finance was 25%, which was the highest. We can see securities and investment services produce the largest data such as text/numbers, and have the highest demand. Seo (2013) reported that insurance companies and credit card companies are using big data more actively than banking companies in Korea. However, researches on the use of big data in securities and investment companies have been found to be

insignificant. We surveyed 22 major securities and investment companies in Korea. We can see they actively use AI for recommend. As for big data of securities and investment companies, we studied open API. Of the 22 securities and investment companies, only six are offering open APIs. The characteristics of open API are as follows. First, the user OS is 100% Windows. Second, the languages used are mainly VB, C#, MFC, and Excel provided by Windows. Third, development manuals are mainly provided on the Web, and only three companies are provided as files. Because the development documentation of the file format is more convenient than web type, they need to provide documentations of the file format. There is a difficulty in real-time analysis and decision making since we cannot receive data directly using Hadoop, the big data platform, because Linux is not supported. In order to activate big data in the securities and investment fields, they should support Linux. Second, Java and Python also be supported. Third, they should provide easy-to-view development manuals and videos such as YouTube.

Acknowledgement

This research was supported by Seokyeong University in 2019.

References

[1] Gartner, The Importance of 'Big Data': A Definition, Retrieved from https://www.gartner.com/en/documents/2057415, 2012.
[2] Jungkee Kim, Web Server Log Visualization, International Journal of Advanced Smart Convergence, Vol.7, No.4, pp101-107, 2018.
DOI: http://dx.doi.org/10.7236/IJASC.2018.7.4.101
[3] KB management Institute, Understanding big data and implications for financial services, 2012.
[4] Kim Ji-Woon and Hur Jun, Kim Jang-Il, Use of Big Data by financial institutions, The Journal of Korea Institute of Electronic Engineers, Vol.40, pp.795-800, 2013.
[5] McKinsey Global Institute, Big Data: The next frontier for innovation, competition, and productivity, 2011.
[6] Min-Kyo Seo, A Case Study on the Use of Big Data in the Korean Financial Institutions, E-Trade Review, Vol.11, pp.115-134, 2013.
DOI: http://dx.doi.org/10.17255/etr.11.4.201311.115
[7] Seong-Hoon Lee and Dong-Woo Lee, Big Data Processing and Utilization, The Journal of digital policy & management, Vol.11, No.4, pp267-271, 2013.
[8] Wikipedia, Big Data, Retrieved from https://en.wikipedia.org/wiki/Big_data, 2019.
[9] Young-Hyun Chang and Chang-Bae Ko, A Study on Efficient Building Energy Management System Based on Big Data, International Journal of Advanced Smart Convergence, Vol.8, No.1, pp82-86, 2019.
DOI: http://dx.doi.org/10.7236/IJASC.2019.8.1.82
[10] Young In Cho, Understanding Big Data and Its Main Issues, JOURNAL OF KOREAN ASSOCIATION FOR REGIONAL INFORMATION SOCIETY, Vol.16, No.3, pp31052, 2013.