

TRENDS IN DATA WAREHOUSING TECHNIQUES

Naveen Kunnathuvalappil Hariharan

Sr. Hyperion SME & Department of Information Technology, United States

ABSTRACT

Financial data volumes are increasing, and this appears to be a long-term trend, implying that data management development will be crucial over the next few decades. Because financial data is sometimes real-time data, it is constantly generated, resulting in a massive amount of financial data produced in a short period of time. The volume, diversity, and velocity of Big Financial Data are highlighting the significant limitations of traditional Data Warehouses (DWs). Their rigid relational model, high scalability costs, and sometimes inefficient performance pave the way for new methods and technologies. The majority of the technologies used in background processing and storage research were previously the subject of research in their early stages. The Apache Foundation and Google are the two most important initiatives. For dealing with large financial data, three techniques outperform relational databases and traditional ETL processing: NoSQL and NewSQL storage, and MapReduce processing.

1. INTRODUCTION

Technological innovations such as advanced analytics, machine learning, AI, big data, and the cloud have penetrated and transformed how financial institutions compete in the market thanks to digitization. Large corporations are using these technologies in order to carry out digital transformations, meet consumer demand, and improve profit and loss. While most businesses are storing fresh and important data, they aren't always sure how to make the most of it because the data is unstructured or not captured within the organization ([Schwertner 2018](#)).

Companies must adapt to these changes in a planned and thorough manner as the financial industry rapidly advances toward data-driven optimization. Financial businesses will be able to fully use the capabilities of unstructured and high volume data, identify competitive advantages, and develop new market opportunities with efficient technology solutions that satisfy the advanced analytical demands of digital transformation ([Liebowitz 2013](#)).

Big financial data has emerged as a useful idea for businesses seeking high commercial value, and it is often described by its intangible properties (volume, variety, velocity, value and veracity) ([Garmaki et al. 2016](#)).

The point at which financial data becomes “big” is subjective, therefore it remains an abstract word, generally characterized as data that is too big, too rapid, or too difficult to process using present techniques. Massively parallel processing; mixed and complex analytical workloads (e.g., ad hoc querying, data mining, text mining, exploratory analysis, and materialized views); flexible storage to support data from multiple sources; real-time operations (stream processing): the concept has gained significant notoriety in many business areas, such as healthcare, retail, manufacturing, or modern city changes.

Because of the high demand for input/output efficiency, scalability, and elasticity, which can be accomplished through parallel storage and processing, Big Data Warehousing represents a fundamental shift for organizations facing several difficulties in their traditional Data Warehousing systems, namely bottlenecks throughout the collection, storage, processing, and analysis of data. Another thing to consider is traditional DWs' stringent modeling approach, which only emphasizes the importance of BDWs. The state-of-the-art in BDW, however, indicates the concept's youth, as well as ambiguity and a lack of standard techniques to building BDWs based on their qualities.

2. BIG DATA

Big data consists of data sets that are so massive and complicated that they are impossible to process using typical data processing software or on-hand database management solutions (Big data). It's a jargon or catchphrase for a big volume of both organized and unstructured data that's too large to be processed using standard database and software procedures. Petabytes (1,024 terabytes) or exabytes (1,024 petabytes) of data containing billions to trillions of records of millions of people are examples of large data (Sri and Anusha 2016) (Shukla et al. 2015).

Social networks, web-logs and web pages, smart mobile devices and sensors (especially those belonging to the IoT paradigm) are the greatest generators of big data, but statistical open data, multimedia (videos, images), and telecom data are also important big data sources (Agrawal et al. 2014; Thai et al. 2016).

It's vital to deal with big data on time and in (real-) time, but it's also tough due to the large amounts of data that come in at a rapid rate. Dealing with a chunk of data can't always be put off because the data will be lost if it isn't. The (un)structuredness of data can range from unstructured to semi-structured, which makes creating a formal data model difficult. The necessity for exploratory analysis of such data is frequently mentioned, and data visualization is viewed as a tool for solving this challenge (Khan et al. 2018).

Big data is large, quick, unstructured, and comes in a variety of forms. There are various distinguishing characteristics:

1. Variety: Data comes in a variety of forms, including structured and unstructured data. Numeric and text fields make up structured data. Images, video, audio, and a variety of other sorts of unstructured data are examples. There are several data sources as well. Data from ERP systems and other operational systems are among the conventional sources of structured data. Data that is unstructured can be found in a variety of places, including social media, the Web, RFID, machine data, and more. Unstructured data exists in a variety of sizes and resolutions, and it can be analyzed in a variety of ways (Jeble et al. 2018). Video files, for example, can be tagged with labels and played, but video data is rarely computed, and the same is true for audio data. Network distances can be calculated using graphic data. Although feelings in Facebook texts and tweets can be assessed, they cannot be directly compared (Vasarhelyi et al. 2015).

2. Velocity: The Internet substantially enhances the speed with which data may be moved, from emails to social media to video files. Cloud computing allows for rapid and easy sharing from any location. People can rapidly share their info with each other through social media programs. Data generation and access are also sped up with mobile access to these applications.

3. Volume: Websites have evolved into excellent data stores. The clickstreams of users are recorded and saved for further use. Users can now become data prosumers thanks to social media programs like Facebook, Twitter, Pinterest, and others (producers and consumers) (Bello-Orgaz et al. 2016) (Leung and Zhang 2016). The number of data shares has increased, as has the size of each data element. High-definition videos have the potential to enhance the total amount of data shared. Video, music, text, data, and other autonomous data streams are generated by social networking sites, websites, RFID applications, and other sources (Bendre and Thool 2016).

3. BIG FINANCIAL DATA

Data capture, data storage, search, sharing, transmission, analysis, and visualization are all demanding tasks when dealing with a huge volume of financial data. Mobile devices that collect data, aerial sensory technologies, software logs, cameras, microphones, radio frequency identification scanners, and wireless sensor networks are all becoming more common. Big data is currently and will continue to be a trend. In

today's big data environment, the volume of data is growing at breakneck speed, and the rates of data can be continuously processed: storing, analyzing, and reporting in real time. These databases include billions and trillions of numerical entries, making them indispensable not only to the financial industry but also to government, telecommunications, and academic research.

In the financial sector, big data includes a vast amount of stock data, quotation data, and pricing data (Fang and Zhang 2016). According to the Financial Information Forum, U.S. exchanges generated fewer than 65 billion transaction and quote signals each year less than a decade ago (FIF - Financial Information Forum). The data volume on the exchanges recently surpassed 1.5 trillion, and it shows no signs of slowing down. A fundamental capital markets difficulty is acquiring, classifying, and evaluating this data without incurring latency. Because financial data is sometimes real-time data, it is continuously generated, resulting in a tremendous amount of financial data being produced in a short period of time (Alexander et al. 2017). For example, the NYSE has 1500 firms that hold 1570 different types of stock, and these companies will conduct billions of business transactions per day, resulting in a massive volume of financial data. Furthermore, the world includes hundreds of stock exchanges and financial organizations, and big data has established itself as a trend in this new financial climate. Using typical relational database management systems to deal with massive data is simply too complex. The volume of market data is increasing, and this appears to be a long-term trend, implying that data management development will be critical in the coming decades (Yang et al. 2014).

4. BACKGROUND PROCESSING AND STORAGE TECHNOLOGIES

The majority of the technologies utilized in the research on massive financial data warehousing were formerly the topic of research in their early stages (Shen et al. 2015). The Apache Foundation and Google's initiatives are the most important. The following text explains the most important systems created in these projects.

Apache Hadoop is a distributed programming system that is open-source. It is built around a scalable distributed file system called HDFS for data storage and a MapReduce programming language for data processing. This most popular MapReduce implementation, which combines parallel storage and data processing, was inspired by Google File System, which can be credited with kickstarting the current MapReduce revolution (Rana et al. 2018). Hadoop supports a variety of NoSQL databases, and numerous software packages, such as HBase, Hive, Pig, Spark, Sqoop, and others, run on top of or alongside Hadoop, such as Cloudera Impala (Sidhu et al. 2015) (Ishwarappa and Anuradha 2015).

Google File System (GFS) is one of the most important platforms for research in this field, but its NoSQL database Google BigTable is equally noteworthy (Kalid et al. 2017). Unstructured data is stored in BigTable as uninterpreted bytes in the form of triplets that provide a row key, column key, and timestamp. Data versioning is enabled through timestamps, while column keys are used to group data of the same type into sets called column families (Yang 2015).

HBase is a column-oriented database that operates on top of Hadoop and is modeled after Google BigTable. Bloom filters and data compression on the column level are supported in HBase, and tables can be utilized as input and output datasets for MR processing (Das and Kumar 2013) (Shahriari and Baraani-Dastjerdi).

Apache Hive is a technology whose introduction accelerated research in the field of big data warehousing (Du 2015). Hive is a Hadoop-based infrastructure that supports a variety of databases and file systems. It was built with the goal of integrating relational database concepts (declarative query language, tables, columns, and so on) to Hadoop's unstructured environment (Hive 2013). Hive was also classified as a data warehouse since it incorporated a key warehousing concept: a system catalogue metastore storing table metadata.

Table partitions are stored as subdirectories in the table's directory in Hive. Table data is a set of queries written in HiveQL, a SQL-like query language, that are converted into MR jobs and run on Hadoop (Thusoo et al. 2009) (Thusoo et al. 2010). Data structuredness was improved by using a schema-on-read technique, which speeds up data loading by deferring data validation to the query run-time phase rather than the data loading phase. Hive, on the other hand, still lacks a robust implementation of one of relational databases' most crucial features: the optimizer. Its optimizer is simple and rule-based, and it has been improved through study (Bouarar et al. 2014).

Hive is frequently said to be good for examining datasets, however the Apache Pig platform is better for processing data flows. Pig Latin, which can also be used in other frameworks like Spark, is a programming language that operates on top of Hadoop and can be used to create programs and MR tasks (Olston et al. 2008; Swarna and Ansari 2017).

5. BIG FINANCIAL DATA STORAGE AND PROCESSING CONCEPTS

There are three techniques that are superior to relational databases and typical ETL processing for dealing with the previously identified problematic qualities of huge financial data: NoSQL and NewSQL storage, and MapReduce processing.

NoSQL database

NoSQL (Not Only SQL) is a trend that is frequently misunderstood as an antithesis to relational databases, however it does not imply that "SQL" is completely obsolete (i.e. relational). Rather, NoSQL refers to the extension or change of relational storage (Kaur and Sachdeva 2017).

NoSQL databases are largely open-source projects that are currently being actively developed, and they are typically characterized by horizontal scalability, distributed computing, and improved fault tolerance at the cost of more flexible data consistency. They do, however, lack a higher-level declarative language similar to SQL, as well as an uniform data model, which could be due to the many database families' major variances. There are four types of NoSQL databases (Mason 2015): a) column-oriented databases b) document-oriented databases c) key-value store d) graph databases e) object-oriented databases ff) XML databases g) other databases: multimodel, multidimensional, and multivalued databases (Gupta et al. 2017).

Map Reduce

The most popular framework for processing massive amounts of data, MapReduce (MR) (Lee et al. 2012), is also suited to big financial data. It is based on the shared-nothing architectural concept, which refers to a system made up of distributed independent computer nodes that might be virtual or non-virtual. Each node has its own local disk and main memory, and is linked to other nodes via a network. Since there is no shared memory between the nodes, the fundamental feature of this design is that it does not have any contention points (thus the shared-nothing phrase) (Karloff et al. 2010). According to these characteristics, the MapReduce framework is made up of processing nodes that are specifically built for scalability, fault tolerance, data and processing dissemination, and load balancing in cloud environments, while keeping the implementation details hidden from the programmer (Jiang et al. 2010).

NewSQL

NewSQL systems have recently been proposed. The goal of these systems is to attain the same level of scalability and availability as NoSQL databases while maintaining the ACID transaction properties and extensive functionality of relational databases (Pavlo and Aslett 2016). This technique may be appropriate in situations where traditional RDBMS-based systems have been in use for a long time, but there is a recent need

for increased scalability and performance. The following are the main issues in the design and deployment of such systems (Kaur and Sachdeva 2017) (Grolinger et al. 2013):

a) The data model; b) the consistency model; and c) the architecture. The NewSQL concept is still in its early stages of development, with only experimental systems like RubatoDB available.

Storage of Holographic Data

(HDS) Holographic data storage (HDS) addresses ODS's issue by storing data throughout the media's volume rather than only on its surface(Coufal et al. 2000) . Multiplexing is used by HDS to record several holograms in the same amount of data. A laser source's light is split into two beams. The first is a signal beam that is encoded with binary data by passing it via a Spatial Light Modulator (Curtis et al. 2010). The second is a reference beam that is recombined with the modulated signal beam before being recorded on the media. A hologram is created by the pattern created by the overlapping beams. The hologram's distinguishing feature is the unique volume address provided by the reference beam (Ruan 2014). The data is retrieved using the same reference beam that was used to record the hologram. The signal beam, on the other way, exists independently of the other holograms.

6. CONCLUSION

The volume, diversity, and velocity of Big Data are increasingly highlighting the serious constraints of traditional Data Warehouses (DWs). Their rigid relational model, high scalability costs, and occasionally inefficient performance pave the way for new methodologies and technologies. The notion of Big Data Warehousing has recently gained popularity, with the goal of studying and proposing innovative solutions to cope with Big Data issues in Data Warehousing contexts. The Big Data Warehouse (BDW) is a scalable, high-performance system that uses Big Data techniques and technologies to support mixed and complex analytical workloads (e.g., streaming analysis, ad hoc querying, data visualization, data mining, simulations) in a variety of emerging contexts, such as Smart Cities and Industries 4.0. However, because this field is still in its infancy, the vagueness of the notions and the absence of consistent methodologies persist.

Integration of traditional data warehouses with big data is a difficult scenario since, in the past, there was just structured data to deal with, but today there is also other (big) data to deal with, which might improve analytics by adding context. The use of MapReduce and NoSQL databases has become a key aspect of big data warehousing research. This article examines how new concepts have been employed to bridge the gap between traditional data warehouses and big data analytics demands. This paper focuses on novel technologies and approaches for incorporating NoSQL databases and MapReduce systems with one another or other systems to enhance or speed up data processing and storage in data warehouses, or to implement new algorithms for computing the multidimensional data cube, which is the key data structure of a data warehouse.

REFERENCES

- 1) Agrawal, D., C. Budak, A. El Abbadi, and T. Georgiou. 2014. "Big Data in Online Social Networks: User Interaction Analysis to Model User Behavior in Social Networks." International Workshop on Databases and Expert Systems Applications: Proceedings. https://link.springer.com/chapter/10.1007/978-3-319-05693-7_1.
- 2) Alexander, Lewis, Sanjiv R. Das, Zachary Ives, H. V. Jagadish, and Claire Monteleoni. 2017. "Research Challenges in Financial Data Modeling and Analysis." Big Data 5 (3): 177–88.
- 3) Bello-Orgaz, Gema, Jason J. Jung, and David Camacho. 2016. "Social Big Data: Recent Achievements and New Challenges." An International Journal on Information Fusion 28 (March): 45–59.

- 4) Bendre, Mininath R., and Vijaya R. Thool. 2016. "Analytics, Challenges and Applications in Big Data Environment: A Survey." *Journal of Management Analytics* 3 (3): 206–39.
- 5) Bouarar, Selma, Ladjel Bellatreche, Stéphane Jean, and Mickaël Baron. 2014. "Do Rule-Based Approaches Still Make Sense in Logical Data Warehouse Design?" In *Advances in Databases and Information Systems*, 83–96. Springer International Publishing.
- 6) Coufal, Hans J., Demetri Psaltis, Glenn T. Sincerbox, and Others. 2000. *Holographic Data Storage*. Vol. 8. Springer New York.
- 7) Curtis, Kevin, Lisa Dhar, Adrian Hill, William Wilson, and Mark Ayres. 2010. *Holographic Data Storage: From Theory to Practical Systems*. John Wiley and Sons.
- 8) Das, T. K., and P. M. Kumar. 2013. "Big Data Analytics: A Framework for Unstructured Data Analysis." *International Journal of Engineering Science*. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.411.6697&rep=rep1&type=pdf>.
- 9) Du, Dayong. 2015. *Apache Hive Essentials*. Birmingham, England: Packt Publishing. <https://play.google.com/store/books/details?id=4S7WBgAAQBAJ>.
- 10) Fang, Bin, and Peng Zhang. 2016. "Big Data in Finance." In *Big Data Concepts, Theories, and Applications*, edited by Shui Yu and Song Guo, 391–412. Cham: Springer International Publishing.
- 11) "FIF - Financial Information Forum." n.d. Accessed August 24, 2021. <https://fif.com/>.
- 12) Garmaki, M., I. Boughzala, and S. F. Wamba. 2016. "The Effect of Big Data Analytics Capability on Firm Performance." *PACIS*. <https://core.ac.uk/download/pdf/301369542.pdf>.
- 13) Grolinger, Katarina, Wilson A. Higashino, Abhinav Tiwari, and Miriam A. M. Capretz. 2013. "Data Management in Cloud Environments: NoSQL and NewSQL Data Stores." *Journal of Cloud Computing: Advances, Systems and Applications* 2 (1): 22.
- 14) Gupta, Adity, Swati Tyagi, Nupur Panwar, Shelly Sachdeva, and Upaang Saxena. 2017. "NoSQL Databases: Critical Analysis and Comparison." In *2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN)*, 293–99. ieeexplore.ieee.org.
- 15) Hive, Apache. 2013. "Apache Hive." [intel.ch. https://www.intel.ch/content/dam/www/public/us/en/documents/case-studies/big-data-apache-hadoop-framework-spotlight.pdf](https://www.intel.ch/content/dam/www/public/us/en/documents/case-studies/big-data-apache-hadoop-framework-spotlight.pdf).
- 16) Ishwarappa, and J. Anuradha. 2015. "A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology." *Procedia Computer Science* 48 (January): 319–24.
- 17) Jeble, Shirish, Sneha Kumari, and Yogesh Patil. 2018. "Role of Big Data in Decision Making." *Operations and Supply Chain Management An International Journal*, January, 36–44.
- 18) Jiang, Dawei, Beng Chin Ooi, Lei Shi, and Sai Wu. 2010. "The Performance of MapReduce: An in-Depth Study." *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases* 3 (1-2): 472–83.
- 19) Kalid, Sultana, Ali Syed, Azeem Mohammad, and Malka N. Halgamuge. 2017. "Big-Data NoSQL Databases: A Comparison and Analysis of 'Big-Table', 'DynamoDB', and 'Cassandra.'" In *2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, 89–93. ieeexplore.ieee.org.
- 20) Karloff, Howard, Siddharth Suri, and Sergei Vassilvitskii. 2010. "A Model of Computation for MapReduce." In *Proceedings of the 2010 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 938–48. *Proceedings. Society for Industrial and Applied Mathematics*.
- 21) Kaur, Karambir, and Monika Sachdeva. 2017. "Performance Evaluation of NewSQL Databases." In *2017 International Conference on Inventive Systems and Control (ICISC)*, 1–5. ieeexplore.ieee.org.
- 22) Khan, Nawsher, Mohammed Alsaqer, Habib Shah, Gran Badsha, Aftab Ahmad Abbasi, and Soulmaz Salehian. 2018. "The 10 vs, Issues and Challenges of Big Data." In *Proceedings of the 2018 International*

- Conference on Big Data and Education. New York, NY, USA: ACM. <https://doi.org/10.1145/3206157.3206166>.
- 23) Lee, Kyong-Ha, Yoon-Joon Lee, Hyunsik Choi, Yon Dohn Chung, and Bongki Moon. 2012. "Parallel Data Processing with MapReduce: A Survey." *SIGMOD Rec.* 40 (4): 11–20.
- 24) Leung, Carson K., and Hao Zhang. 2016. "Management of Distributed Big Data for Social Networks." In 2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), 639–48. ieeexplore.ieee.org.
- 25) Liebowitz, J. 2013. "Big Data and Business Analytics."
- 26) Mason, Robert T. 2015. "NoSQL Databases and Data Modeling Techniques for a Document-Oriented NoSQL Database." In *Proceedings of Informing Science & IT Education Conference (InSITE)*, 3:259–68. proceedings.informingscience.org.
- 27) Olston, Christopher, Benjamin Reed, Utkarsh Srivastava, Ravi Kumar, and Andrew Tomkins. 2008. "Pig Latin: A Not-so-Foreign Language for Data Processing." In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, 1099–1110. SIGMOD '08. New York, NY, USA: Association for Computing Machinery.
- 28) Pavlo, Andrew, and Matthew Aslett. 2016. "What's Really New with NewSQL?" *SIGMOD Rec.* 45 (2): 45–55.
- 29) Rana, Poonam, Vineet Sharma, and P. K. Gupta. 2018. "Exploration of Apache Hadoop Techniques: Mapreduce and Hive for Big Data." In *Advances in Computing and Data Sciences*, 543–52. Springer Singapore.
- 30) Ruan, Hao. 2014. "Recent Advances in Holographic Data Storage." *Frontiers of Optoelectronics* 7 (4): 450–66.
- 31) Schwertner, K. 2018. "The Impact of Digital Transformation in Finance Sector." *Социальные и экономические системы*. <https://elibrary.ru/item.asp?id=36367682>.
- 32) Shahriari, Farrokh, and Ahmad Baraani-Dastjerdi. n.d. "A Method of Query over Encrypted Data in HBase Database by Using Bloom Filter Algorithm." *Academia.edu*. https://www.academia.edu/download/33003671/Maghaleh_-Bloomfilter_Encryption_for_HBase-Eng.pdf.
- 33) Shen, Y. C., C. T. Yang, S. T. Chen, and W. H. Cheng. 2015. "Implementation of Software-Defined Storage Service with Heterogeneous Object Storage Technologies." *Proceedings of the ASE*. <https://dl.acm.org/doi/abs/10.1145/2818869.2818891>.
- 34) Shukla, S., V. Kukade, and S. Mujawar. 2015. "Big Data: Concept, Handling and Challenges: An Overview." *International Journal of Computer*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.695.5850&rep=rep1&type=pdf>.
- 35) Sidhu, Rohan, Deanna Chea, Rachita Dhakal Annie Hur, and Mark Zhang. 2015. "Implementation of Hadoop and Sqoop for Big Data Organization." *aeop.uark.edu*. 2015. <https://aeop.uark.edu/files/2017/07/TeamB2.pdf>.
- 36) Sri, Psga, and M. Anusha. 2016. "Big Data-Survey." *Indonesian Journal of Electrical*. <http://section.iaesonline.com/index.php/IJEEI/article/view/195>.
- 37) Swarna, C., and Zahid Ansari. 2017. "Apache Pig-a Data Flow Framework Based on Hadoop Map Reduce." *International Journal of Engineering Trends and Technology (IJETT)* 50 (5): 271–75.
- 38) Thai, M. T., W. Wu, and H. Xiong. 2016. "Big Data in Complex and Social Networks."
- 39) Thusoo, Ashish, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Suresh Anthony, Hao Liu, Pete Wyckoff, and Raghobham Murthy. 2009. "Hive: A Warehousing Solution over a Map-Reduce

- Framework.” Proceedings of the VLDB Endowment International Conference on Very Large Data Bases 2 (2): 1626–29.
- 40) Thusoo, Ashish, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Ning Zhang, Suresh Antony, Hao Liu, and Raghotham Murthy. 2010. “Hive - a Petabyte Scale Data Warehouse Using Hadoop.” In 2010 IEEE 26th International Conference on Data Engineering (ICDE 2010), 996–1005. ieeexplore.ieee.org.
- 41) Vasarhelyi, M. A., A. Kogan, and B. M. Tuttle. 2015. “Big Data in Accounting: An Overview.” Accounting Horizons. <https://meridian.allenpress.com/accounting-horizons/article-abstract/29/2/381/99282>.
- 42) Yang, Jade. 2015. “From Google File System to Omega: A Decade of Advancement in Big Data Management at Google.” In 2015 IEEE First International Conference on Big Data Computing Service and Applications, 249–55. ieeexplore.ieee.org.
- 43) Yang, Shuang, Kun Guo, Jianping Li, Yingjiong Zhong, Rui Liu, and Zili Feng. 2014. “Framework Formation of Financial Data Classification Standard in the Era of the Big Data.” *Procedia Computer Science* 30 (January): 88–96.