5C, A NEW MODEL OF DEFINING BIG DATA

Liang-Jie Zhang1,3, Jing Zeng1,2
1 Kingdee International Software Group Co., Ltd., China
2 Research Institute of Information Technology, Tsinghua University, China
3 National Engineering Research Center for Supporting Software for Enterprise Internet Services
zhanglj@ieee.org

Abstract

Big data as an emerging paradigm has revolutionized the IT, which is embodied with features of 4Vs (volume, velocity, variety, veracity). However, with the rising of the digital economy 2.0, 4Vs features merely give the non-functional criteria and they cannot precisely depict the essence of big data and fail to address how to apply the big data in actual scenarios. Therefore, to advance the down-to-earth application of big data in the digital economy 2.0, we propose a brand new model – 5C (creator, channel, center, context, consumer), which will fundamentally redefine the requirements of big data and give a novel methodology for big data. In the proposed 5C model, creator elucidates the body of creation of big data. Channel refers to the transmitting issues of big data. Center denotes how to make data capitalized and further form valuable assets. Context is the application scenario of big data based on different platforms. Consumer involves users of big data under specific contexts. To demonstrate the proposed 5C model, we use big data architecture of Amazon and business opportunity map of Kingdee as the cases to illustrate what we propose. We also summarize the successful rules for people to embrace big data under the digital economy 2.0. Finally, we conclude that the presented 5C model can be fully used as basic principles and guideline for leading the construction of enterprise big data.

Keywords: Digital Economy 2.0; Big Data; 5C;

1. INTRODUCTION

With the rapid development of internet of things (IoT), currently it is coming into an era of internet of everything. It is a vital impetus for the applications of big data. Thereby big data has already become the strategic assets of enterprises, individuals and government. Along with the refined scenarios of enterprise management and mobile office, the essence of big data also has also been updated.

As a buzzword, big data has gained popularity in the current society and industry for years. Nonetheless, many existing research groups are more inclined to emphasize the significance of concept in big data rather than investigating the essence of big data, which means that mining the connotations and to maximize the value of big data.

In the actual application contexts, enterprises may confront various tricky problems in every day. For example, companies expand the marketing institution but their performance still keeps decreasing; partner abandons them but they cannot find the reasons; company substantially invests in new product line but receives little effect; Research and development (R&D) receives tremendous investments, but it is still difficult for company to dominant the market. There are many reasons to explain them; however, the fundamental reasons are that the enterprises lack data thinking. They can be summarized as follows: 1) Precise marketing is insufficient to meet the requirements of market; 2) Although enterprises own the data of partner and customer, but data mining and analysis are totally ignored; 3) It lacks effective associations between equipment data and business data; 4) The investment of traditional R&D neglects the data-driven innovations of product and business model. In future many problems can finally be considered as data problems, such as Google AlphaGo (Silver, 2016) win human player in Go game. To some extent big data can significantly change many aspects of enterprises.

Typically, big data is exhibited with 4Vs (volume, velocity, variety, veracity) features defined by IBM scientists (IBM, 2014). The volume represents the
scale of data, the volume of the data produced by recent two years has achieved the 90 percent of all previous produced data. The velocity demonstrates the speed of data production, such as the data speedily produced by 4G devices, social network, etc. The variety denotes the type of the data, which not only encompasses the structured data (database table), but also includes semi-structured (XML, Html, etc) and un-structured data (video, voices, social media, etc). The last is about veracity which tends to illustrate the truth of the big data.

By virtue of the deep development of big data concept, some researchers also show their understanding about big data. Doug Laney, which is the researcher fellow at the vice president level of Gartner, argued that he had proposed 3Vs (volume, velocity, variety) in his research report in 2001 (Laney, 2001). Another Scientist, John R. Mashey, (Mashey, 1997) coming from SGI considered big data volume as a pressure which infrastructure confronted. These also reflect the diverse understanding of different persons and stages for big data.

Therefore, with the rising of digital economy 2.0, which is an evolution of digital economy, we should refine the requirements of big data era. The definition of 4Vs features merely gives the non-functional criteria and they are unable to exhibit the intrinsic quality of big data and how it can create value. According to our practice of many years, as long as it has complete data set for specific industries and domains, it can be called big data. We focus on getting the insight of data and creating value by delivering the data among different organizations and different applications, furthermore, forming value chain. Typically, we use 5C (creator, channel, center, context, consumer) model to refine the big data in digital economy 2.0 era. The first core element creator is the main body of data production. The second core element channel refers to the data communications. The third core element center represents the assets of data. The fourth core element context depicts the scenarios of created value by application data. The last core element consumer is the user of created value by digital economy. They pave the way to promote the applications of big data in digital economy 2.0.

The reminder of this article is organized as follows: Section 2 briefly discusses the concept of digital economy 2.0 and introduces the 5C model for big data in detail. Section 3 gives two cases to validate the model in use. Conclusions are summarized in Section 4.

2. BIG DATA IN DIGITAL ECONOMY 2.0
2.1 Digital Economy 2.0

Digital economy was first mentioned in the book Digital Economy (Tapscott, 1996) which was written by an America writer in 1996. In this book, digital economy sometimes can be described as internet economy, new economy or web economy, which indicates that some economy will be generated through cardinal number specially calculated by numbers. Actually, only the word of “new economy” was applied in many countries among the above three words. In contrast, the implication of the internet economy and web economy are not in accord with the connotation of digital economy according to the 20 years’ development process.

Thus, many researchers have already discussed the issue that which kind of economy can be called the digital economy. Is it the internet economy or the web economy? Although they were the latest concept in 20 years ago, they may seem inaccurate due to the appearance of internet technologies. However, the growth of the mobile internet with the wearable devices and the IoT devices has changed the situation, that is, these economic forms do not merely rely on the internet any more. The definition of Internet is changing due that internet itself just is a channel.

In 2001, an economist, Thomas Mesenbourg, had said digital economy should contain three key elements (Mesenbourg, 2001). First is the infrastructure mainly including hardware, software, and telecoms. Second is the e-business which means the enterprise digitization. In other words, the business process and the whole value chain should be transformed on the basis of digitalization. Finally, the third element is e-commerce including logistics and online sales etc. It indicates the transaction and transport of goods can be fulfilled via digit.

For the first element, the rapid development of hardware, software and communications, supporting infrastructure has significantly changed compared to past. In the underlying hardware, the circuit is highly integrated, and has much lower energy consumption and cost. From the perspective of microcosms, the chip occupies much smaller area and owns much lower price. In the real scenarios, the chip can reach the 0.75 of square-millimeter (Nakase et al., 2001), the size of computer is greatly developed from small to tiny, even to invisible. All these emerging embedded computers sharply drive the development of sensors, such as MEMS, NEMS, etc. They also to some extent construct the basis for IoT. From the
perspective of macrocosms, it refers to the high performance computing (HPC), including GPU/CPU, cluster, supercomputer. The advanced hardware technologies such as Google HPC cluster (Google, 2017), Sunway Taihulight (Fu et al., 2016) of China national supercomputer which can achieve 93 petaflops, and can offer very powerful guarantee for developing data processing business in computing resources. The software is also greatly improved with the hardware development, especially artificial intelligence (AI) and data processing technologies. AI is refashioning due to deep learning (LeCun et al., 2015) which allows machines to automatically discover the knowledge via a large number of labeled data and neural network. Distributed data processing starts to dominate the market of data analytics because of the powerful distributed data processing software which can run on a large scale of PCs. It strongly reduces the cost of computers. These software technologies pave the way for data analytics business. Moreover, telecoms or network as an infrastructure has been greatly facilitated by various novel businesses, for instance, IoT, VR-AR, smart city, video or voice services. The emerging network largely enhances the sharing of data between human and human, human and things, things and things. These infrastructures possibly bring potentially new value for digital economy.

Regarding to the e-business, what the most concern for companies is how to make good use of computers to proceed their businesses. The value chain is reorganized due that data is playing a key role for business. By analyzing the data of company business, companies can learn which business needs to be adjusted if necessary. In fact, data can be used to sense the requirements of market. On the one hand, there are a plenty of available data collected many years. On the other hand, it has affordable computing resources to conduct data analytics for average companies.

The e-commerce is another important element of digital economy. Traditionally, logistics and online sales can be digitalized by using internet technologies and mobile internet technologies. Actually, either internet or mobile internet can be identified as a platform where multiple applications of e-commerce can be deployed and run to realize their businesses. Also, the large company can build platform to provide services. For example, Taobao of Alibaba Group in China offers a platform for millions of small businesses and individual entrepreneurs to open online shop. Here, platform can gain the profit by providing high-quality services. That is, services model as an important way can significantly improve the profit by seamlessly jointing customers and businesses.

From the history of digital economy, it is observed that due to the development of the base of digitalization or information, the internet has become mobile internet and then the mobile internet has become the IoT and finally the IoT turns to virtual reality (VR-AR). In other words, the platforms transform from internet to VR-AR. Therefore, what we should process is to push the digital economy to march toward a new stair. That is the digital economy 2.0.

2.2 5C Model for Big Data

As a new economy mode, the digital economy 2.0 cultivates renewed spirit for big data. Along with the development of cloud computing, IoT and big data storage, process, analytics and visualization, big data as a value-added resource greatly promotes the finance, transportation, hotel, aviation, education, etc. As the Figure 1 demonstrates, the big data in the digital economy 2.0 possesses five core elements under the impetus of the mentioned technologies. It consists of creator, channel, center, context and consumer, which will start a new era of big data from 4V to 5C.

![Figure 1. Core Elements of Big Data in Digital Economy 2.0](image)

**Creator.** The first core element is creator, who aims to figure out the principal part that creates data in this new era. Data is the only resource which can be blended on different separate platforms. Thus, what kinds of principal part for creating data should
be the most important element to the digital economy 2.0. It can be divided into three types of sources: IoT, existing intelligent systems, human computer interaction.

As Figure 2 illustrates, the IoT and VR has appeared under the impetus of the technology so that the creator of data has also changed. And now the main part of data creator becomes IoT which connects everything, such as surveillance cameras, refrigerators, mobile phones and air conditions. As for smart city, the equipment in manufacturing industry and the system in city or company will create data. And all this data will create more data through integration.

The second critical attribute of creator is heterogeneity, which is also due that the various sources can result in multiple types of data, such as some data created by camera, or wearable devices which determine that different creators can create diverse data. The heterogeneity of creator determines the variety of data value when the big data applications are deployed. This is due that the different types of data require different data processing technologies, thus further affecting the cost in development and deployment.

The third attribute of creator is domain-specificity. The data can come from multiple domains, such as industry, agriculture, finance, transportation. The creator of corresponding domains has different needs for the big data applications. In order to the feasibility of specific domains, it is very necessary to exactly identify the features in corresponding application domain and build the pilot applications to validate the creator.

The last attribute of creator is real-time. In IoT, many sensors can produce data in real-time. Due to the application scenario becoming prevalent in recent years, the real-time of creator brings much more challenges in data processing infrastructure. Therefore, when we intend to carry out big data, the real-time of creator should be thoroughly considered to support the decision-making of big data solutions in actual applications.

**Channel.** The second core element is the channels to the data propagation. And internet is one of them. The channel of data propaganda has radically altered compared to the situation in 10 or 20 years ago. As Figure 3 demonstrates, the channel of data propaganda is the IoT which. It is also called ICT internet, corresponding to the traditional internet or the basic physical internet. In contrast, with the appearance of audio and video, the data created by internet end users will be shared through social medium such as smartphones and wearable devices. And finally, all these data will be deposited in the society.
The data channel includes the sources, destination and the transmission medium. In fact, the essential part is the medium or network of data transmission. The emerging IoT and social network have extended the data propagation, which is significantly affected by data scale and communication technologies. The data scale to some extent determines the size of channel. Millions of sensors are deployed into the space, the overall scale is quite large but the data produced by each sensor is very small, so the channel can use various wireless transmission technologies. With the emergence of wireless technologies, such as Zigbee (ZigBee Alliance, 2006), NB-IoT (Ratasuk, 2016), etc, the data created by IoT can be transmitted precisely and speedily. This is also a fundamental change for the requirements of data communications.

In the digital economy 2.0, the channel of big data has revolutionized the past transmission manner of data. When we choose the channel for big data, various factors should be considered. First, the data communication solutions should be determined by specific application requirements. Second is that we should make a tradeoff between the performance and cost. That is due that there are many available channels that can be chosen. Lastly, the service of quality (QoS) for the channel should be acquired. The choice of data transmission solutions should meet the QoS demands of big data. For different data, the channel may be diverse. For example, the pipeline may be suitable for stream data to guarantee the data transmission.

**Center.** After we have the data creator and data propaganda channel, but how can we price and capitalize these data? Therefore, the third core element center is a place for the production and propagation of data. The data created from different industries can be classified, thus becoming valuable and reusable assets in this place. It just like the commodity can be chosen and purchased in Wal-Mart. Each data can be precisely priced and leveraged to realize its value in actual scenarios. It is the key part and basic target to apply big data in business process. The value of big data lies in that it can be used as a service. By the servitization of big data, it can be further identified as an asset to be used by other individuals or organizations. That is why we need to build a data center for the incorporations.

These novel wireless communications drastically reduce the cost and energy consumption for data transmission. For example, zigbee protocol, which works in 2.4G frequency and can reach 250Kb/s speed, can run supported by two 1.5V batteries for at least six months (ZigBee Alliance, 2006). The low-power consumption can ensure the stability of data transmission of IoT. Besides, each chipset of zigbee is less than two dollars which largely promote the applications of IoT.

Another channel is about the propaganda of personal data, which leverages the social network to push or pull the data into individual cellphone. The social network fully embodies the spirit of sharing. People can actively share their interesting photos, short videos and valuable comments via social networking services (SNS). Every person can be the channel of data propagation in our socially connected society. The value of the channel heavily depends on the interaction of human and human.
calculated data has been shared. That is the real reason why the artificial intelligence (AI) can defeat human in a short period. As Figure 4 shows, capitalized data center includes the data center via combination of data and cloud computing, data transaction platform for e-business, open API for invoking and automation and intelligent based on data.

Data center usually refers to cloud computing, because cloud provides unprecedented merits for elastic data storage via virtualization technologies. Especially, the storage virtualization can support multiple object storage and block storage in distributed computing environment. It is quite appropriate for big data storage. The cloud offers services to users. Typically, it totally includes three levels (Mell & Grance, 2011): infrastructure as a service (IaaS), Platform as a service (PaaS), Software as a service (SaaS). IaaS shields the underlying computing resources (CPU, memory, network, etc) and provides the bottom level interface to upper. Furthermore, The PaaS can support a series of system resources (operation systems, database, etc). The SaaS is the high-level abstraction for underlying software and hardware resources. Leveraging the mentioned cloud technologies, big data can be provided as a service to users.

The second is the data transaction platform E-business is the best case. That is also the key part for value of big data. How to make the data center valuable? What we should do is preprocess the data, classify the data and open the assess interfaces. For the reusable data, we should fix a price and capitalize it. Regarding the preprocessing of the data, we should conduct the data cleaning so as to remove the incomplete, inconsistent and inaccurate data. Then, we make a category for the data which can help us distinguish the potential customers. Furthermore, the assess interfaces should be provided to users who can develop new applications by themselves.

The third is to open API to share data, which also takes the data as a service. The sharing can create value in propagating the data process. On the one hand, it maximizes the use of data that improves the data value. On the other hand, it can improve the assess volume of corresponding adhesiveness of users. The open API can be a user or the existing systems used in secondary development.

Finally, the big data can embody its value in augmenting the intelligence and automation. By intrinsic law of hidden in data, the artificial intelligence (AI) is greatly enhanced by big data. Specifically, the neural network can be trained by a large amount of data which has a much better accuracy in issues of classification compared to other approaches. A typical example was AlphaGo (Silver, 2016) which defeated the champion of human in Go game by learning a large amount of playing data created by human players based on deep learning.
technologies. Data has presented an unbelievable intelligence in digital economy 2.0, which is unimaginable in early years.

Thus, if we want to develop the value of data, it needs to establish an exchanged standard due that data center is just a carrier. The exchanged standard means a definite price can be utilized in the exchange between the data and data or people. We all hope data can embed in different industries. Consequently, different companies can share all the data through external interfaces so that the data has been endowed with intelligence and elaborative faculty. In addition, the formation of the data planning ability, inferential capability or learning ability is attributed to the basic ability which we should outfit in the capitalized data center.

Context. The fourth core element is that how enterprise solves their problems by utilizing the data assets created from different industries and areas. In this way, the data should be suitable for creating business occasions. For example, the business of drops a taxi and Uber were derived by the data. If they do not exactly process the information about the drivers or the cars, they are impossible to maintain the business successfully. All the data, such as the place and time you take the taxi and the destination has been recorded. And then the company would exchange the data with the cooperative hotel. Finally, the company can attract customers accurately by utilizing these data. Hence, a passage published in an America magazine even has praised the Uber as a company which harnessed big data.

In this point, it is very necessary to find suitable applicable scenarios in specific industry so as to explore the value of the digital economy 2.0, let the data create value for end users. In addition, various business scenarios for big data to create value are of great importance. The context involves various aspects, including communications among people, e-commerce, manufacture, design, marketing, after-sales service, as demonstrated in Figure 5.

People communicate with each other via the channels of communication, such as WeChat and Weibo. The value of the data communication is in an ascendant in this scene. In 25th September, 2016, Google had released a new generation of IM tools (Google Allo, 2017), integrating gadget window and incorporating personal schedule. In such small window, users can subscribe services by the figures. The services also extremely rely on big data.

The second is e-commerce context. In this process, a substantial of data is accumulated in the transactional process. The data can exactly specify the user purchase habit and their preferences. Many e-commerce websites collect the data of user click to recommend suitable commodities to potential customers. Specifically, it is enabled by big data analytics and processing technologies.

Certainly, various capabilities of data analysis are the same in the electric supplier industry and a traditional manufacturing industry. Big data can enhance the efficiency of each component in value chain about 5%-15%. Customer service in the product after integrating a large amount of data, can be also got help and it improves customer satisfaction to achieve the design satisfied the marketing requirements, precise marketing and after-sales service. Some researchers in Forbes (Gordon, 2013) use big data analytics to identify valuable opportunities. By mining a large amount of data, the sale leader can predict the new sale chances instead of making a decision only by historical sale performance at a specific field.

A variety of application scenarios will bring us new sources of income which is able to be directly achieved. The method is as follows: firstly, we make the data capitalized, and then utilizes it in specific application scenarios. The value is created from the interactions between data in service form and end-users. The end-users can truly benefit from these application scenarios because they are the origin of big data value in specific applications. Therefore, they are play important roles in data capitalization

Furthermore, to design big data solutions, we should find the matched scenarios for applications. Three types of issues should be examined. First, the application scenarios should provide enough data volume once you decide to implement the big data. Many application scenarios, for example, some simple text records in local supermarket management system, are inappropriate for big data scenarios. Then, multiple domains can share the data to achieve win-win results of multi-industry. For instance, the big data of e-commerce and the big data of logistics can be integrated to improve the efficiency of commodity delivery. Third, the new application scenarios should be explored. With the advent of big data, novel business model may bring the new economic growth point in corresponding industry areas.

Consumer. The last core element is consumer. All the products or serves need to be consumed by
consumers. However, the problem is that what kinds of forms or results about this situation should be carefully considered in this new economic era.

In the age of the digital economy 2.0, for the most important consumers – users of data services, what critical roles do they play? These consumers can drive the development of the digital economy requirements and habits, lifestyles will also change the idea of people on product design. Therefore, in such an era, as Figure 6 shows, the digital consumer comprises three major categories: natural persons who get information and interact with the world; the second, everything interconnected, for example, each car has hundreds of sensors, how to use the data will became a key problem; application scenarios, various software, hardware and systems, together with data prompt to add value in scenarios. Moreover, invisible users are embedded into software and hardware.

![Figure 6 Consumer of Big Data in Digital Economy 2.0](image)

Different from machines, the human consumers are more inclined to the human-centric and sensory representation about the data services. Hence, they enjoy the visualization of data and user-friendly interfaces when using the big data services. Developing visualization technologies of big data can help ordinary consumers’ understanding.

The second consumer comes from IoT that senses the things existed in the world by sensors and actuators. The volume of consumer is so huge that the traditional data processing technologies do not suffice. Usually, a number of clusters and distributed computing platform, such as Spark (Zaharia et al., 2010) and Hadoop Mapreduce (Hadoop, 2009) can be used to address the volume problem by a huge number of consumers.

The third consumer is the existing system including multiple software systems (Saas), which get the data from data center and further process it to offer valuable services for end-users. They are the miners of data value and discoverer of useful knowledge from big data. Although these consumers are not traditional within your understanding, they will be the propeller of digital economy 2.0.

How the connections among these three types of consumers are handled? The design of business scenarios is usually the combination of all three types of consumers, and their relationships between data consumers and data creator can be transformed reciprocally. The five elements contained in the digital economy 2.0 are the data generator, data channel, data center, data business application scenarios to the data consumer. The data is intangible but is creating a completely new economy, thus it cannot be measured by traditional GDP. It also indicates that people and materials can create equivalent products. It is a combination of tangible materials that determine the fluctuation of GDP.

In order to improve the GDP, we try to enhance the productivity, boost sales and expand market. However, this is far away from the traditional definition of GDP. The new digital economy, like a famous American analyst wrote (Worstall, 2016): what do digits and production means? What people consume is beyond the GDP records.

Evidently, data plays a vital role in the digital economy 2.0, but it is not yet recognized by public and government and even is not properly described in the digital economy 2.0. Therefore, it is predictable that new GDP definition will appear in recent years. In fact, various businesses, as engaged in the software and enterprise internet industry, are closely related to the digital economy, since everyone may be the data creators, data integration, and application scenarios. Some people focus on the end consumer, once we consider the digital economy 2.0 as a whole, the five core elements can be well understood. Whatever industry we are engaged in, it is essential to find the pain point of user. That is, we need to understand end users’ requirements, and create a novel way that traditional GDP cannot measure or implement.

3. CASE STUDY

Faced with the opportunities and challenges brought by big data in digital economy 2.0, how can
the managers of enterprises survive in new digital economy era and lead the trend? Here, we demonstrate two cases to show some successful experiences of dominated enterprises within industry.

Usually, we use the lifecycle of service innovation to describe the specific application scenarios, which can be divided into six phases (Zhang et al., 2007). As Figure 7 demonstrated, they are summarized as follows:

**Consulting and Strategic Planning:** The key activity in this phase is to invite third-party consulting companies to work out strategic planning. Customers entrust customer support plan (CSP) to analyze its potential position in the market and the IT professional plan for the next five years. At the end of the consulting phase, CSP helps customers finish a request for proposal (RFP).

**Services Engagement:** Customers distribute RFP to multiple service providers to bid for the project; then customers make agreement with one service provider, that is IT provider (ITP).

**Services Delivery:** The third phase includes the formation of service delivery teams and governing project management, solution creation, work breakdown, macro design, development and implementation, testing and deployment.

**Services Billing:** In this phase, the service provider gathers payment the contracts from the customer.

**Services Management:** In this phase, since the ITP is in charge of the solution creation and IT service management tasks, ITP could predict the traffic knowing that customers want to add a new service, and recommends that customers to increase IT resources for the new service.

Given a case in real big data application scenario of big data, we take Amazon web services (AWS) big data architecture (Raghupathy, 2015) as an example, as illustrated in Figure 8. The architecture is one of the best interpretations for the proposed 5C model of big data.

In the first phase, service consulting and strategic planning focus on the data collect services, since it is observed that data collection requirements are exactly defined. The creator plays a critical role in the data collection services. As previously discussed, it is mentioned that the creator refers to multiple data sources specified in the AWS big data architecture. They originate from three types of sources. The first are applications which refer to web apps, mobile apps and logstash (Anicas, 2015). The second sources are from logging records, such as flume (Hoffman, 2013), log4j (Gülcü, 2003), fluentd (Fluentd, 2017), etc. The last are the emerging IoT applications which produce a substantial of data in recent years. These three types of sources are transactional data, search data, file data and stream data, respectively.
In the second phase, various services of storage components are needed to help to store the data from the data sources. Therefore, these data services can refer to multiple IT providers, which are oriented to database systems, NoSQL cache, search SQL, file storage and stream storage. In this process, the channel can affect the selection of IT providers due that the transmission of these data may leverage various communication manners. The performance of data transmission is also sharply affected. Meanwhile, the context influences the storage providers. Diverse application contexts are mentioned in the architecture, such as NoSQL storage and search, file data storage and stream data storage.

The third phase and fourth phase are about the development of data analysis services and corresponding data service operation. By the ETL process, which is realized by the pipeline or channel between data storage and analytics, the data are cleaned and conversed into available data set. Furthermore, storage services are employed by data analytics services. Multiple analytics services are described in the architecture, including machine learning (ML) services, interactive services, batch services and stream processing services. The architecture is exhibited by the proposed center concept, which refers to capitalized data center. Amazon can mine the big data value while leveraging multiple analysis services. It uses Amazon ML to accomplish the machine learning. Amazon redshift and presto are used to provide interaction for big data. Amazon elastic MapReduce platform aims at solving batch process of big data. It also provides stream processing solution by integrating multiple data

Figure 8 Overview of AWS Big Data Architecture (Raghupathy, 2015)
analytics platform, including spark stream, storm, AWS kinesis, and AWS lambda. All these big data analytics services are to capitalize the data and make it more valuable to be further consumed within various applications. Besides, it presents many contexts of application when we conduct the big data analytics.

In the phase of services billing and management, the consume element in the Amazon architecture which mainly involves how to use and manage the data services. It refers to the consumer in the proposed 5C model. The object which requires analysis services consumption coming from five parts. They are predictions, analysis and visualization, notebooks, IDE as well as applications and API. Actually, they involve two kinds of users. One is business users who need the analysis and visualization services. They focus on the presentation of data so as to find the value of the data. Another is data scientists and developers who pursuit both analysis and visualization, IDE and notebook. Furthermore, billing strategies vary from user to user. Business users have much higher payment ability than data scientists. Therefore, the bill can be personalized according to diverse demands of users. In addition, the storage of AWS is well maintained, and hence it can be directly consumed by users, which exhibit that data center can be capitalized and the use of data can be purchased by various users.

Another case about the 5C model has been applied in practice is business opportunity map developed by Kingdee (Kingdee, 2017). To process the big data of enterprise, in the service consulting and planning phase, Kingdee has built the big data center of national enterprise (Kingdee, 2014). As demonstrated in Figure 9, Kingdee offers 23 million information of companies. For the service engagement, it confirms that service demands originate from providing credit information services for management of companies. The requirements of business opportunity map lie in making users much easier to discover how many companies in the
buildings and their corresponding products. The searching process is shown in Figure 10, which can help users quickly gain the channel of business opportunity.

Moreover, the best practice of services delivered by business opportunity map of Kingdee is how to use 5C model in the real scenario. Firstly, the creator of the big data originates from the companies information, including the state of the companies including registered capital, scope of business and major managers. Other than the traditional product sale model, the compensation of business opportunity map depends on use fee for the access of required company information.

The second is about the channel of the data. Multiple technologies are employed to transmit the data, such as internet, social network, smart cellphone and even the VR. It is an emerging technology to demonstrate data and applications. We also construct the VR platform on the basis of Kingdee VR and enable it to become a part of business opportunity map solution, which can facilitate the fast realization of business scenarios in 3D and seamlessly embed the real-time data into 3D virtual environment. Furthermore, the intelligent transformation of various business scenarios can be realized in VR scenarios.

The third is the application scenario of business opportunity map, we converse the enterprise data into available data services which can be stored in the capitalized data center. We open the API of the developed services in business opportunity map so as to share the data assets that will greatly promote the development of big data in the era of digital economy 2.0. The sharing aims to start the trial in new economy era and find the insight of big data. Once the value chain of big data is formed, it will bring the earth-shaking change in the application scenarios of big data.

In the service operation phase, for the context of business opportunity map, it provides visualization for business users to find the company information in a digital map. It refers to the technologies of web services used in the fourth phase of service lifecycle. In general, plenty of location information about the company is well marked in the map via interaction between visualized presentation services and data services. This context of business can produce the value relying on taking data as a service, thus prompting company find the information they truly want in a flash.

The last in services billing and management phase refers to consumers of the map data. They are the business manager and marketing personnel, who are eager to find the company information to pursuit cooperation with other companies or to purchase the required products. The digital value is realized through addressing the panic point of companies and finding business opportunities. In this phase, we focus on the service billing and management so as to provide personalized and stable services for end users. For the services bill of big data, we have also opened the ecosystem of big data center, the API of which has been assessed by 3.7 billion times.

The 5C model of big data provides a guideline to ensure how to implement big data service in various phases of business lifecycle. In the practice of Kingdee, the 5C model can address the application implementation of big data. When we conduct the practice of big data, it is essential to consider the five core elements defined in the proposed 5C model.

Finally, based on the study, the successful rules for people to embrace big data under the digital economy 2.0 can be summarized as follows:

1) We should adequately cultivate the abilities of data thinking.
2) We should improve the construction abilities of services and products in big data era via 5C model.
3) The new technologies, such as VR and AI, should be converted into available services that can be shared or an open API, thus prompting their propagation to create value.

4. CONCLUSIONS

In this article, we introduce a new insight about big data in the digital economy 2.0, which indeed brings us a new prospect and entrepreneurial opportunity for the application of big data. We apply 5C model to guide the implementation of big data and conduct the down-to-earth practice by the proposed methodology.

The 5C are concisely and exactly summarized for recognizing the key characteristics of big data in real applications. It gives the insightful descriptions from data creation, data channel, data asset, context of data application and data consumer. Based on the model, they can provide a reference for researchers and practitioners to comprehensively understand and grasp the core of big data.
We hope that with the inspiration of the 5C model, we can greatly facilitate cooperation in this field, jointly create a completely new generation of digital economy and make our own contributions to big data in the digital economy 2.0.

5. ACKNOWLEDGMENT

This work is supported by the central grant funded Cloud Computing demonstration project of China, R&D and industrialization of the SME management cloud (Project No. 20112448), hosted by Kingdee Software (China) Company Ltd., under the direction of National Development and Reform Committee of China; the construction fund of National Engineering Research Center for Supporting Software of Enterprise Internet Services (Project No. 2012FU125Q09); the Guangdong province project, under grant No. 2015B010131008; the fund for leading talents of Guangdong Province [2012]342; the fund for National Key Research and Development Plan (Project No. 2016YFB1000803).

REFERENCES


Authors

Liang-Jie (LJ) Zhang is the Chief Technology Officer (CTO) and Senior Vice President of Kingdee International Software Group Company Limited. He has served as the President of Shenzhen Big Data Alliance since 2013. He also served as a director of The Open Group from 2011 to 2012. Prior to joining Kingdee, he was a Research Staff Member and Program Manager of Application Architectures and Realization at IBM Thomas J. Watson Research Center as well as the Chief Architect of Industrial Standards at IBM Software Group. Dr. Zhang has published more than 160 technical papers in journals, book chapters, and conference proceedings 50 granted patents.

Jing Zeng is the postdoctor of Kingdee International Software Group Co., Ltd., China. His research interest includes IoT and big data.