Ann. Data. Sci. (2017) 4(4):433–440
DOI 10.1007/s40745-017-0129-9

How China Deals with Big Data

Yong Shi[1,2] · **Zhiguang Shan**[3] · **Jianping Li**[4] · **Yufei Fang**[5,6]

Received: 7 June 2017 / Accepted: 28 September 2017 / Published online: 14 November 2017

Abstract On September 5, 2015, the State Council of Chinese Government, China's cabinet formally announced its Action Framework for Promoting Big Data (www.gov.cn, 2015). This is the milestone for China to catch up the global wave of big data. Since 2012 big data became a hot issue for scientific communities as well as the governments of many countries (Lazer et al. in Science 343:1203–1205, 2014; Einav et al. in Science 345:715, 2014; Cate in Science 346:818, 2014; Khoury and Ioannidis in Science 346:1054–1055, 2014). At the 2013 G8 Summit, the leaders of Canada, France, Germany, Italy, Japan, Russia, U.S.A. and United Kingdom agreed on an "open government plan" (www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex, 2013). China's framework, however, mainly emphasizes the integration of all trans-departmental data and establishes a number of government-driven national big data platforms so as to provide big data services to research, public and enterprises. The framework not only demonstrates a strong commitment of the Chinese government on big data, but also covers a wide range of governmental branches, enterprises and institutions far more than that of other countries. In

✉ Yong Shi
yshi@ucas.ac.cn

1 School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China

2 Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing, China

3 Informatization Research Department, State Information Center, Beijing 100045, China

4 Institute of Policy and Management, Chinese Academy of Sciences, Beijing 100190, China

5 School of Computer and Communication Engineering, University of Science and Technology, Beijing 100083, China

6 Informatization Research Department, State Information Center, Beijing 100045, China

addition, the framework shows an interpretation of big data that differs from other countries. If its objective is achieved, China would become a strong "big data country".

Keywords Big data · Chinese government · Big data framework · Data sharing · Data opening · Data service

1 The Contents of the Action Framework

The framework can be interpreted as a Top Level Design, which consists of three national platforms: National Data Opening Platform, Trans-Departmental Data Sharing Platform and Internet based National Data Service Platform (see Fig. 1). By 2020, the Chinese government commits to complete 10 key big data projects for three platforms so as to provide big data applications in a number of public areas, including credit, transportation, healthcare, employment, social security, geography, culture, education, science and technology, agriculture, environment, safety and security, product quality, statistics, meteorology and ocean service. These projects will finally assemble a numbers of big data systems across a wide range of governmental departments, industries, academic and education institutes. After the announcement of the Action Framework, the Chinese government has published a planning program to further nail down key tasks for related departments to carry out its respective potions of Fig. 1 in terms of the responsibilities, road maps and target dates into 2020. This program offers a strong support for the completion of the Action Framework.

2 The Impact of the Framework on Chinese Culture

Importantly, the Chinese government tries to use the implementation of the framework to influence the Chinese culture and social value towards data and builds social data awareness. In May 2013, a group of the Chinese and international scholars, including the authors, brain-stormed an academic definition of big data as "a collection of data with complexity, diversity, heterogeneity and high potential value, which are difficult to process and analyze in reasonable time". It is comparable to well known 4V (Volume, Variety, Velocity and Value) model of big data. To encourage the Chinese leaders to speed up the policy guideline, they further stated that "big data is new type of Strategic Resource in digital era and the key factor to drive Innovation, which is changing the way of human being's current production and living" [7,8]. A Chinese version of big data can be regarded as the large-scale data set produced and being utilized from China's modern informatization process, the totality of data source in the current information society and the whole set of data, not only internet data, but also governmental and commercial data [9]. The framework calls for the entire Chinese society to have big data thinking and also drives social data awareness. The traditional Chinese decision making relies on qualitative thinking, not quantitative thinking. Such a cultural behavior has burdened nation's science and technology development. Thus, the Action Framework aims to

Fig. 1 China's big data top–down design

change the current Chinese culture by enhancing data awareness and promoting data spirit.

3 The Strengths of the Framework

Comparing with other countries, the strengths of China's framework are *strong government-drive, high R&D investment, and booming internet business*. Although many other countries struggle with large scale data handling, China is more likely to succeed in two aspects. First, over 80% of the most valuable data and information in China are gathered and controlled by Chinese government and its departments. Therefore, this framework focuses on trans-departmental data sharing, which is more realizable and effective. Moreover, with the largest amount of data, Chinese government yet has the worst data utilization capability. Thus, focusing on improving this weakness will lead to a significant step towards the Big Data China. Second, the operating mechanism of Chinese government is very special. The large scale data

Table 1 China's provincial government plans on big data

2010	Chongqing
2011	Inner Mongolia, Xinjiang
2012	Guangdong, Shaanxi
2013	Beijing, Shanghai, Tianjin, Jiangsu, Zhejiang
2014	Guizhou, Anhui, Fujian, Hainan, Henan, Hubei, Hunan, Liaoning, Shandong,Shanxi, Sichuan, Yunnan, Guangxi, Ningxia, Tibet
2015	Gansu, Hebei, Heilongjiang, Jiangxi, Jilin, Qinghai

handling of American and other countries is motivated by the free market; however, the big data action of China is motivated strongly by the Chinese government, which is much more effective since the government has a decisive influence on the national society and economic activities. For last 6 years, all China's 31 provincial governments (Provinces, Autonomous regions, and Municipalities) also made their policy and planning to support big data projects (see Table 1). Among them, Beijing municipal government, especially, Zhongguancun High-Tech Zone, is the flag of big data movements in China. China's two special administrative regions: Hong Kong and Macau are advanced in big data applications. For example, big data stock has been hot at Hong Kong exchange market while Macau government considers building a credit data management for its Casino VIP halls.

Since 2014, China's big data research has been supported by four major funding agents: NSFC (National Natural Science Foundation of China), MOST (Ministry of Science and Technology), CAS (Chinese Academy of Sciences) and NDRC (National Development and Reform Commission) [10]. Each of big four has its own focus to support various projects ranged from the fundamental research to the industrialization of big data. NSFC is the key player to fund scientific research projects on big data. For instance, in 2014 it provided more than 13 million US dollars (82 million RMB) grants to 19 big data projects, where each has more than 300 thousand US dollars (2 million RMB) (see Fig. 2). In September 2015, NSFC announced its first big data major research plan is called "Big Data driven Management and Decision" with 24 million US dollars (150 million RMB) for 8 years. This project is joined managed by 4 NSFC departments: Management Science, Information Science, Mathematics and Physics and Medical Science. MOST funded two big data projects in cyber space big data computing theory and technology with the total amount of 64 million US dollars (400 million RMB). CAS has established its big data key laboratory, called Big Data Mining and Knowledge Management Lab in 2014. NDRC has drafted this Action Framework and is implementing the Top Level Design (Fig. 1), especially 10 key big data projects, on behalf of the Chinese government. All of the money was not spent casually. Instead, it was spent in a very deliberate way so as to effectively push the big data action forward.

The typical big data application of China that has been well-known around the world is finance, banking and internet business [11,12]. Especially, online commerce has contributed 7% of China's GDP growth in 2014. In September 2014, Alibaba, one

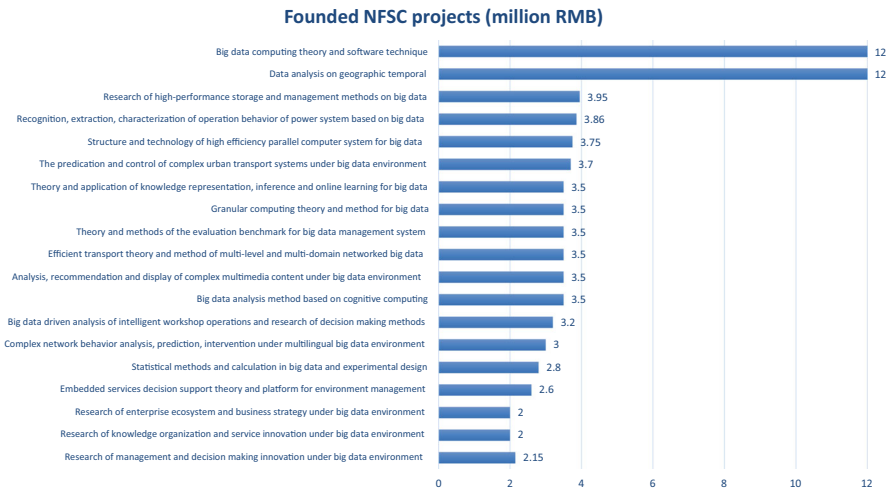

Fig. 2 2014 China's big data projects founded by NFSC (million RMB)

of China's big data companies, went public trading on the New York Stock Exchange. Its value was more than 200 billion US dollars as the largest IPO of the year in New York Exchange. On November 11, 2015, called "Double 11 Day", Alibaba created a record one-day sells of 1.45 billion US dollars (9.12 billion RMB) over about 200 countries and regions. Recently, hundreds and thousands Chinese youth have registered start-up firms with all kinds of innovative ideas. One of examples is Mr. Liu Chengcheng, the 28-year-old master graduate from the University of Chinese Academy of Sciences in 2014, founded 36Kr High-Tech Media. Less than 5 years, the firm becomes China's leading service provider for Internet entrepreneurial enterprise with market value of about 1 billion US dollars (60 billion RMB) and Liu himself was elected as one of 2013 Forbe's entrepreneurs under 30 in China.

4 Challenges and Suggestions to the Framework

There are *several obstacles* which challenge the Chinese government to fulfill the Action Framework.

First, it may lack of leadership to control the progress of carrying out the framework. The major job to manage all respective big data systems in Fig. 1 is to coordinate the actions among different government departments, which can be a tedious, careful and time-consuming work. It could be better to set up a strong National Big Data Committee including scientists and experts under the State Council to supervise the progress of the platforms in Fig. 1.

Second, the framework could be too big to complete by the given deadline. The Chinese government needs further to prioritize the proposed systems and tasks, then concentrate on key targets, and finally expand to others. For example, the big data application system has to handle various credit issues regarding governmental units, production, business, finance, tax, engineering projects, transportation, e-commerce,

food safety, healthcare, intellectual properties, environmental protection, and services. These involve the collaborations among 20 more governmental departments [13]. This is obviously a difficult project to finish. In this case, China can first build a fair credit law, and then use its newly established National Personal Credit System, known as "China Score" from the Peoples' Bank of China as the basis for all other credit applications [14].

Third, although Chinese technology companies are doing quite successfully globally and domestically due to the largest population market in China, the critical problem is that some fast growing companies adopted the business models of other countries, mainly from USA. This is the obstacle for them to gain competitive ability at the global markets for a long run. The Chinese government should encourage them increase the R&D budget and create their own intellectual properties to master the cutting-edge technologies. The Chinese government should take advantages of Chinese scientific community to push industrial innovation. The following actions should be done: (i) relying on scientists in evaluating all governmental big data projects/systems; (ii) requiring big data industries to invite scientists to join their big data application whenever the company accepts the government financial support; (iii) setting up more National Big Data Engineering Centers with key roles of scientists in selected big data companies; and (iv) providing more research funding via not only governmental agents, like big four, but also governmental departments to generate more opportunities for scientists so as to accelerate the technology transfer.

Fourth, like many other countries, China lacks data scientists to carry out this framework. Based on IDC's latest report on China, China's business analytics services market was 1.398 billion US dollars in 2014, up 16.4% from 2013 and will reach 3.027 billion US dollars in 2019 [15]. This shows a big demand of data scientists in China for next 5 years. To respond the market, the leading universities and research institutes in China, such as Tinghua University, University of Chinese Academy of Sciences and Fudan University, have been offering various types of Data Science degrees and Data Scientists trainings. However, the Department of Education should formally recognize data science as foundational major and develop a large-scale education programs to train data scientist for the big demand.

Under the Action Framework, China's big data movement could be speeding up and expanding very quickly, which may enhance the research capability of the international scientific research community and also boost both national and global economy.

5 Recent Progress of the Framework Implementation

Since the release of the Action Framework, the importance of big data has been highly recognized by leaders. An integrated national big data center was proposed as the country seeks to enhance governance capability. In the Belt and Road summit, President Xi Jinping emphasized that big data plays an important role of connecting the twenty first century Digital Silk Road. Not only has the national level paid attention to big data development, local governments also attach importance to big data increasingly. Up to June 2017, more than 40 provinces and cities issued nearly 100 big data development policies and big data industrial plans.

Table 2 Strategic orientation of national big data comprehensive test areas

Strategic orientation	Province or city
National level experimental area	Guizhou
Urban agglomeration area	Beijing–Tianjin–Hebei, Pearl River Delta
Regional pilot area	Shanghai, Henan, Chongqing, Shenyang
Big data infrastructure integration and utilization Area	Inner Mongolia

China has taken some significant steps to fulfill the Framework. In February 2016, under approval of the State Council, 43 departments formed "the National Inter-Ministerial Joint Meeting on Big Data Development" led by NDRC. The Joint Meeting is responsible for the implementation of the Framework, the research on big data development issues, the promotion of government data opening and sharing, and the enhancement of overall management of national data resources.

Government data resources sharing is one of the Joint Meeting's priority works. In September 2016, the State Council issued the Interim Measures for the Administration of Sharing of Government Information Resources, explicitly requested the rules of management, coordination, evaluation and supervision on government data. Since then, rapid progress was made in government data sharing across the country. By the end of 2016, more than 20 data open platforms had been opened and operated throughout the country.

In 2016, the Joint Meeting approved 8 National Big Data Comprehensive Test Area", each of which has its own strategic orientation (Table 2). This measure aimed at solving bottleneck problems including administrative system innovation, public data opening and sharing, big data application, big data industry, big data trading and circulating, data center integration and utilization, international communication and cooperation.

To promote technology research and industrial innovation, the Joint Meeting selected 38 Big Data Development Major Projects from 1700 candidate projects nationwide. More than 600 million RMB financial support were provided to these projects. 11 National Big Data Engineering Laboratories were established, some of them focus on specific application area, and others focus on cutting-edge technologies.

Recently, the Joint Meeting unveiled a package of measures to accelerate the combination of big data policy, industry, research and application. In May 2017, National Big Data Expert Advisory Committee and the National Big Data Innovation Alliance were launched on China International Big Data Industry Expo in Guiyang.

The expert committee is composed of academicians, scholars from scientific research institutes, and representatives of industrial circles. The innovation alliance is made up of more than 70 related entities of the 14 Big Data National Engineering Laboratory. Through these two mechanisms, big data scientists and companies would be gathered together contributing to policy-making, technology consulting and technology transformation. In addition, a series of open data competitions will be hold in

8 National Big Data Comprehensive Test Areas, to promote pubic data opening and encourage innovative applications.

Acknowledgements This work has been partially supported by grants from National Natural Science Foundation of China (Nos. 91546201, 71331005).

References

1. Action Framework for Promoting Big Data (2015). www.gov.cn. (**in Chinese**)
2. Lazer D, Kennedy R, King G, Vespignani A (2014) The parable of Google flu: traps in big data analysis. Science 343:1203–1205
3. Einav L, Levin J (2014) Economics in the age of big data. Science. https://doi.org/10.1126/science.124308
4. Cate FH (2014) The big data debate. Science 346(6211):818
5. Khoury MJ, Ioannidis JPA (2014) Big data meets public health. Science 346:1054–1055
6. Cabinet Office (2013) Policy paper: G8 open data charter and technical annex. London. www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex
7. XSSC (2013) Report on the 462nd session: data science and big data. In: Xingshan Science Conference, Chinese Academy of Sciences, Beijing, May 29–31
8. Shi Y (2014) The bridge. U Natl Acad Eng 44:6–11
9. Shan ZG (2015) Interpretation on action framework for promoting big data. http://news.xinhuanet.com/info/2015-09/17/c_134632375.htm. (**in Chinese**)
10. Sun YT, Cao C (2014) Demystifying central government R&D spending in China. Science 345:1006–1008
11. (2014) Report on China's big data technology and industrial development. China Machine Press, Beijing. (**in Chinese**)
12. Li JP, Zhang YJ, Wu DS, Zhang W (2014) The bridge. US Natl Acad Eng 44:20–26
13. Task Arrangement for Social Credit System Planning Outline (2014–2020) (2014). http://bgt.ndrc.gov.cn/zcfb/201501/t20150105_659410.html. (*in Chinese*)
14. Credit Public Sentiment, Credit reference center of people's bank of China 64(7) (2010). (**in Chinese**)
15. IDC report, China business analytics services market forecast and analysis, 2015–2019. www.idc.com/getdoc.jsp?containerId=prCN25609415