Rethink Data Sharing under the Epidemic of Novel Coronavirus Pneumonia (NCP)

Xie Bin and Han Tianyu

At present, the situation of epidemic prevention and control in the country is changing for the better, but at the same time the situation in the city of Wuhan and the province of Hubei remains tough and complex. During the epidemic, the Ministry of Industry and Information Technology (MIIT) organized the construction of the national support and dispatch platform for key medical supplies to improve the efficiency in supplying emergency response (ER) resources by sharing the data of human, financial, physical and other strategic resources. The telecommunication big data are used to analyze the movement of people and provide important support for epidemic monitoring and accurate policy-making. The sudden outbreak of the epidemic has become a tough test of data sharing capabilities in the modernization of national governance systems and capabilities.

I. Data sharing problems were exposed during the NCP epidemic.

Low efficiency in data collection. Vertical data reporting requires level-by-level reporting, as a result, the communication process takes a long time, data sources are often repeated, and the data consolidation process is complex. General Secretary Xi Jinping pointed out at a meeting of the Standing Committee of the Political Bureau of the CPC Central Committee on February 3, "health bureaus, emergency response bureaus, government offices, CPC county committee offices, women's federations, education bureaus, agricultural and rural affairs bureaus in some cities and counties, they each produced one or several tables and forms, requiring grass-root cadres to fill in and report quickly while the content required in these tables and forms is virtually the same". In terms of horizontal data release, there is no specific platform to access, there are so many communication channels, the information takes so many different forms, making it difficult to consolidate them. With 8 hospitals in Wuhan and hospitals from all over the country asking help from the public, countless "donor communities" emerged across the country. For example, the theme post "crowdfunding donations - joining hands in the fight against the epidemic" on Wuhan Station of Baidu Tieba (literally, it means "let's post", it is an online community platform) and the epidemic

information release platform group on Douban attracted more than 13,000 members in a short time.

Difficulties in data circulation and sharing. Data is not shared among central government agencies, resulting in limited sources of public service data and inaccurate data analysis. Currently, the widely-used red, yellow, and green health codes lack the support of public transportation data, hospital diagnosis and treatment data, behavior and trajectory data, and other data. Consequently, a lot of doubts and complaints surfaced. "Green code was issued to the people returning from Wenzhou to Hangzhou, but why I was given a red code only because I'm from Zhuji", "The three members in my family have the same situation and why my code was changed to red after update". Such complaints kept popping up. Cross-provincial data is unavailable, resulting in insufficient joint regional defense and unbalanced labor flows after the resumption of production and work. On February 26, a case from Hebei was missed, which led to concentrated infections in a workplace in Beijing. The Wuhan University conducted a survey on the flow of migrant workers, which shows that in more than 100 villages with zero infection across the country, the cross-regional labor mobility in most of these villages is less than 20%. Enterprises in the eastern region want desperately to resume production, but they simply can't due to the lack of disease-free and available labor forces, while migrant workers in the central and western regions are idle at home and want to work. But the labor force is frozen across the nation and the economic and social recovery is slow.

The data search results are not credible. Some charity organizations have not fully disclosed the data of received donations and there is a lack of reliable data management mechanisms to ensure the authenticity of the disclosed information, leading to the credibility of the data in the sharing and interaction process being questioned. As of February 3, Wuhan Red Cross Society accepted social donations in an amount of 920 million yuan in, among which only 151 million yuan was disclosed for use, accounting for 16.4% of the total donations. According to public information released on the official Weibo account of Wuhan Red Cross Society on January 31 concerning the distribution of donated materials from January 22 to 28, 36,000 N95 face masks flowed to Wuhan Ren'ai Hospital and 16,000 to Wuhan Tianyou Hospital. However, the information was subsequently changed to "18,000 face masks to Wuhan Ren'ai Hospital and 18,000 to Wuhan Tianyou Hospital", while Wuhan Xiehe Hospital, which was on the front line of the epidemic prevention, only received 3,000 medical surgical masks and the name of this

hospital was even not on the allocation list at the beginning.

II. The NCP epidemic prevention and control requires data sharing capabilities.

The key to consolidating data from multiple sources is to establish a unified data model. To achieve economy of scale and maximize the value of data, it is necessary to aggregate data from government departments and industries that have data sharing needs. A unified data model is the key to solving problems such as inconsistent data standards and the confusions caused by the disclosure of information on multiple platforms. We need to promote the mutual recognition and collection of data across regions, departments, and levels by establishing a unified semantic and grammatical data model for data sharing and standardizing the formats of data storage and exchange. For example, in the Shanghai, Zhejiang, Jiangsu and Anhui of the Yangtze River Delta, a joint prevention and control cooperation mechanism can be established, and a unified notification form can be implemented through mutual recognition of health observation status discharge notices.

The key to the interaction of data stored on multiple platforms is to establish a data query mechanism. To give full play to the value of government data to public institutions and for the public interest, it is necessary to ensure that data flows freely across central government agencies and across the country. The establishment of cross-platform data query is the key to connect multiple platforms and ensure that data is used by multiple subjects without conflict. By establishing traceability of cross-platform data and clarifying the scale of data interactions, issues such as accessibility and detectability of data between platforms can be addressed. At present, the health commissions of Beijing, Tianjin and Hebei, the development and reform commissions of Sichuan and Chongqing have gradually established a joint cross-regional prevention and control mechanism. 28 provinces, autonomous regions, and municipalities across the country have launched a "digital epidemic prevention system" with Alibaba. On February 28, MIIT and the Ministry of Education jointly launched initiatives to leverage broadband networks for teaching and learning purposes. By extracting data from teacher management information systems and student funding information systems, the subsidized data usage packages can be accurately sent to the teachers and students in need.

The key to the verification of multi-agency data is to establish a chain of security trust. Data needs to flow freely in various fields for general good. The

establishment of a fair and transparent data governance mechanism and rules for open data flow consistent with socialist values are the foundation of data sharing. Technologies such as blockchain provide solutions in terms of fairness, feasibility, and transparency in data acquisition and utilization. The key information published by agencies and institutions on the blockchain platform needs to meet the specifications and be recorded on an open and transparent distributed ledger to solve the problem of information asymmetry and reduce data risks. For example, Fudan University provides a blockchain-based application platform in the field of drug traceability. The platform can enable pharmacies to scan drug codes, analyze data with pharmaceutical companies and provide value-added benefits. Based on the structure of the blockchain, the data is secure and cannot be tampered with.

III. Suggestions on further promoting data sharing

To develop cloud data services for small and medium-sized enterprises (SMEs) and accelerate their digitalization process. To fully mobilize the resources of industry, academia, and research institutes, build cloud manufacturing and service platforms for SMEs, create products, solutions and corresponding toolkits that meet the information needs of SMEs, and lower the digital threshold for SMEs. To gradually promote the cloud transformation of SMEs in infrastructure, business systems and other aspects, establish the production management, supply chain management and other applications in the cloud, and improve the IT application capabilities and levels of SMEs.

To build the smart city brain and improve digital governance capabilities. To comprehensively integrate and consolidate government, corporate and social data, connect data scattered in all corners of the city, and integrate AI, cloud computing and other technologies. To conduct real-time analysis, dispatching and management on all domains of a city through the analysis and integration of large amount of data, establish connected, online, intelligent and open city brains, and use the development of smart cities to improve the digital governance capabilities of governments.

To promote the credibility and transparency of public services and establish digital security guarantees. In terms of promoting the credibility of data, to focus on the research in the field of trusted identities, improve performance and reduce costs, so that the identities of a wider range of users can be quickly and accurately authenticated. In terms of data sharing security, to build up a blockchain-based ecosystem, develop scientific and reasonable application standards, and develop infrastructure adapted to new network. To

attract more enterprises to participate in this process and create a safe and reliable new environment.