

## Executive Summary

---

The value of data is rapidly increasing with the 4th industrial revolution, and the spread of Covid-19 is accelerating the digitization worldwide. Accordingly, data analysis are becoming more important, and data utilization capabilities are emerging as core competitiveness of individuals and institutions. Reflecting this trend, KOICA has established a digital transformation strategy improving processes and systems, and modernizing infrastructure, for better collection, management and utilization of data. Current data management system in KOICA is somewhat insufficient compared to the diversity of data which has been collected and produced by KOICA. Data governance is essential: policies and regulations are much necessary for high-quality data management, especially for analyzing the current status and performance of development cooperation projects through data thus utilizing them in establishing business directions and strategies. Therefore, this study examines the definition and key elements of data governance, and review the current data status. Based on the analysis, it would suggest future directions so that it can be used as a basis for establishing data governance for KOICA development cooperation projects.

Due to the massive increase in an amount of data and its value, many organizations and institutions emphasize the need for data governance, yet there seems to be no established consensus on the definition of data governance. Reviewing at a number of previous studies, data governance is defined differently, but some similarities are found that it aims to collect and manage high-quality data which can be continuously used to create various values within the organization. In this perspective, this study intends to define data governance as a framework consisting strategies, policies, standards, roles, and responsibilities that are required in the entire management process from data creation to utilization.

Data as an asset is rapidly increasing in value, yet it is believed that data is still

undervalued in organizations or there is a significant difference of opinion within the organization over making an investment for systematic management. The data governance is necessary for better communication between participants in the entire data supply chain from data creation to utilization, to make good use of data while reducing risks and costs. Data is generated and collected through various channels and there are numerous constraints until it makes a good usage. Also, while data is everywhere, it is often difficult to find the exact data that is needed, and many times it requires a considerable time to process. This can be attributed to the lack of connection between the supply and demand of data. After all, data governance is a framework that serves as a link between each role in the data supply chain. Data governance manages data as a whole, thus it must be structured to conform with strategies and goals according to the nature of the organizations. And it is necessary to examine at what data is generated, who has authority, and how it is used within the organization. There is no fixed order in building data governance, but it is suggested to define suitable data governance under the given circumstances of the organization, from a small part, then gradually expanded to the whole.

The key element of data governance include standardization, quality management, and data life cycle. Data standardization is a process of organizing diversely existing data according to specific criteria, and setting common language and rules between participants, for better data utilization. Establishing standards is an essential procedure for consistent data quality management, and in the process of reorganizing existing data as well. Also, as the scope of data sharing expands, it is important to manage data properly to meet the quality and prevent misuse of them. Data quality management is a process of reviewing its accurate input according to set standards and managing to convey the meaning of the input data value to the end users. Moreover, detailed quality standards need to be established and a quality dedicated team should be organized, to handle management of considerations in data storage, quality diagnosis and improvement, and performance evaluation. Meanwhile, it is almost impossible to collect and store the generated data indefinitely. And sensitive data such as personal information needs to be discarded after a certain period of time,

and managed to prevent long-term storage. Therefore, organization's data should be managed and stored in accordance with the regulations, and discarded when its utilization is complete or transferred to a separate archive for a long-term storage. For this, it is necessary to examine data's life cycle and establish standards for each stage and a rating system according to its nature.

The data mainly generated by KOICA can be seen as various information generated throughout the implementing process of development cooperation projects, such as recordings earned from the projects including performance and evaluation results. Currently, among these data, unstructured data accounts for a large portion due to KOICA's document-oriented work process, and above all, digitization of these data is still insufficient. For ideal data management, it is required to establish an integrated work and data management system by analyzing the life cycle and step-by-step work process of each data. However, as with most institutions, management systems for each business field are individually established at different times, consequently data are distributed and managed in multiple systems.

Currently, a large part of KOICA's work area is managed through the system, and each work is carried out in connection with a number of systems. Yet, the linkage between the systems are limited, and digitization and standardization of data are insufficient, resulting in restrictions on data utilization. As a beginning point for establishing data governance, it is highly required to discuss standardization focusing on development cooperation project data, which is the KOICA's core data. Examining various documents and data at each project stage within the Integrated Project Management(IPM) system, the quantity was considerable yet revealing the need for efficient management in terms of quality control. In particular, clear guidelines on which data to be collected and managed are insufficient, generating numerous data management standards of each person in charge. When utilizing the data, this leaves only little valuable data. Thus, it is necessary to recognize the importance of data management across the organization, strive to digitize data, and build a digital environment, from establishing rules for the definition and input method of data at

each step and task, to establishing quality standards.

Especially in developing countries, overcoming language barriers and limitations of system infrastructures and managing them to quality data will require much more efforts. As the development cooperation data obtained through KOICA's projects includes unique values, systematic management of the data would enhance its great significance. KOICA's data can be used internally to establish business directions and strategies, and externally to provide an opportunity to expand accessibility to developing countries and share information in the field of international development cooperation. It is close to impossible to establish institutional data governance in a short period of time. Establishing data governance is not a one-time project, but a framework that needs to be continuously managed and improved with flexibility. In addition, sufficient financial resources and manpower must be put in place for the institution-wide. Therefore, in short term, it is necessary to establish standards for KOICA's projects, then internalize the quality and life cycle management. And in the long run, KOICA should consider even the establishment of digital infrastructure in developing countries.