Data as a Public Good: Creating Open Data Platforms in Seoul, 2011–16

Introduction

Many of the smartphone applications that became popular in the Republic of Korea in the late 2000s relied on public information access, raising the demand for third-party usage and dissemination of government-produced information. While the demand was particularly strong for data related to widely used public services and to public transportation and other infrastructure, growing requests for public information disclosure showed a heightened demand for government transparency more generally. Demand was especially strong among the nearly 10 million inhabitants of Seoul, Korea’s capital city. However, Korean laws and government officials had long restricted access to public information and data because of security concerns.

In 2011, Park Won Soon was elected as mayor of Seoul and promised greater openness and transparency. Under Park, the Seoul Metropolitan Government (SMG) aimed to make city data and information “public goods” easily accessible through online platforms that would encourage citizen participation in the decision-making processes of the city government (SMG 2014).

To develop and operate such open data platforms, the SMG had to address several challenges. First, the city government had to organize special training sessions to ensure coordination among and engagement of staff members. Second, the SMG took a two-track approach of short-term and long-term strategies to standardize internal and external data systems for more efficient operation while minimizing the risk for service interruption. Finally, the SMG developed a multilayered system of information disclosure to prevent accidental disclosure of sensitive information.

Introduced in 2012, the city’s open data platforms continue to operate today. They have increased the use of open data and have promoted transparency with much higher rates of information disclosure. Open government data projects worldwide can draw lessons from Seoul’s innovative effort to make public data accessible to citizens.

Development Challenge

Until the early 2010s, the SMG offices responsible for managing public information prioritized security over transparency, which led to restricted public access. Increasing requests for disclosure and free use of public information and data, however, prompted the city government to take a new approach to data as a public good. The demand was particularly strong in areas closely related to citizens’ daily lives, such as user-friendly smartphone applications for city navigation that depended upon government geospatial data. Making public data accessible could encourage business growth and promote more effective and transparent uses of public information and data.

Intervention

In 2012, the SMG introduced the Seoul Information Communication Plaza (SICP) and Seoul Open Data Plaza (SODP) to serve as platforms for users looking for relevant, up-to-date information about local government services, projects,
and policies. The SICP focuses on unstructured data such as administrative information and policy reports, while the SODP provides access to public datasets such as those covering public transit ridership trends as well as pollution indexes.

**Addressing the Delivery Challenges**

**Ensuring Coordination and Engagement**

Initially, the open data initiative caused great discomfort among SMG staff members who were concerned about the potential implications of releasing administrative information and data. They were particularly worried that the misinterpretation of public information could cause confusion and undermine confidence in the reliability of public information (SMG 2014). As a result, staff members responsible for implementing the initiative did not readily accept the new procedures and adapt their work accordingly.

To ensure coordination among and engagement of the SMG staff, the Information Disclosure Policy Division organized semiannual training sessions. Created in 2012 with the specific mission of information management and disclosure, this new division actively promoted training based on the firm belief that the safe disclosure of public information required not only technological solutions but also expert human review. The training sessions typically lasted about four days and covered the importance of information sharing, information disclosure processes, previous experiences of monitoring disclosure of sensitive information, and ways to avoid accidental disclosure.

Strong commitment from the leadership also helped resolve staff concerns during the early stages of implementation. The newly elected mayor and his team prioritized the establishment of open government platforms, and their commitment helped change initial staff perceptions.

**Sequencing Tasks for Data Standardization and Quality Management**

Preexisting internal systems used for routine administrative work produced tremendous amounts of data, which complicated the sequencing of tasks for data standardization and quality management. For the SODP to be useful, its data had to be uniform, consistent, and accurate.

SMG staff members faced the question of whether to simultaneously standardize the internal operational systems within government offices and the external information system used for the SODP—or whether to first prioritize the latter. On the one hand, simultaneous standardization could produce more consistent data, because data input into the SODP would be more uniform. On the other hand, a governmentwide project of data standardization could be complicated, time consuming, and costly. Such a major change could have unforeseen implications. Sequencing the standardization tasks would decrease risks but also reduce data coherence and consistency.

To resolve this dilemma, the SMG took a two-track approach. In the short term, the city government pursued the timely standardization of data received from individual departments for the external system. Over the longer term, the SMG initiated a more cautious and comprehensive process to standardize internal systems. This methodical approach enabled a timely release of data to the public without prolonging the project timeline or raising costs significantly.

Ultimately, this choice proved wise. Data standardization turned out to be a slow and difficult process that required continuous efforts to monitor data collection and transmission, to correct errors, and to standardize formats and codes according to internal guidelines.

**Establishing a Multilayered System of Information Disclosure**

A key challenge for developing the SICP platform was the trade-off between maximizing information disclosure and avoiding accidental leakage of sensitive information. Because administrative documents contain a wide range of confidential and personal information that is subject to nondisclosure laws, including national identification numbers and passport numbers, accidental information disclosure could pose significant risks.

To address this challenge, the SMG built a multilayered system of information management that relied on both technological and human resources. First, an internal administrative information system automatically collected
administrative documents. Next, a digital document disclosure system filtered and masked personal and confidential information. In addition to the automated filtering system, the Information Disclosure Policy Division separately managed internal data about the topics most vulnerable to accidental disclosure. The office manually monitored more than 1,000 high-risk cases among the 10,000 administrative documents produced daily. Finally, the processed information was released to the public on the SICP.

The SMG established additional mechanisms that helped strengthen the information disclosure system. In collaboration with the software company Hancom Inc., the city government developed a special masking function for a word processing software to enable the SMG staff to flag sensitive information. The masked information was viewable by the internal staff only, and it was filtered by the document disclosure system when it was uploaded to the SICP. In addition, the training organized by the Information Disclosure Policy Division covered how to handle sensitive information and to avoid accidental disclosure.

Outcomes

Increased Use of Open Data Provided on the SODP

First introduced in early 2012 with the disclosure of 918 internal raw datasets, the SODP quickly grew. As of February 2021, the SODP provided more than 6,600 datasets across a wide range of categories, including transportation, general administration, and the environment. The number of instances in which users accessed data from the SODP increased rapidly, from 2.3 billion in 2016 to 9.3 billion in 2019 (as of October 2019; SMG 2019).

Researchers and developers based in Seoul used the SODP to improve public services. For example, the Seoul Digital Foundation relied on those data to create an efficient and user-friendly system to manage public bicycles. Unexpectedly, the open data initiative helped the SMG staff better understand data, which improved data use for city administration. For instance, SMG staff members analyzed patterns of traffic collisions to pinpoint the best locations for speed bumps to reduce collisions between cars and pedestrians (SMG 2015).

The private sector also benefited from the open data platform. Firms used the data to develop smartphone applications that help citizens use public transportation and other city services more efficiently. Combining public data with private sector data also led to useful insights for public service management. For example, the SMG and Korea Telecom more accurately calculated Seoul’s de facto population by combining their data on the number of registered residents and telecommunication signals.

Greater Transparency Facilitated by Information Disclosure Through the SICP

As of February 2021, the SICP provided access to more than 20 million administrative documents. The introduction of the SICP improved Seoul’s information disclosure rate from 73.4 percent in 2013 to 96 percent in 2019. Roughly 58,000 cases were set for nondisclosure every month in 2015; however, by 2019, that number had declined to 7,800, indicating increased transparency. The number of visitors also increased from 82,000 in 2013 to more than 6.4 million2 in 2020 (Yim 2018; SMG 2020). The SICP became a go-to source for public information among information intermediaries such as journalists who sifted through Seoul’s vast inventory of disclosed reports for key findings to bring to the attention of the general public.

Lessons Learned

Establishing an Open Data System Is a Cyclical, Iterative Process Rather Than a Linear Series of Steps

Faced with the dilemma of whether to standardize the internal operational system and external information system simultaneously or to sequence those systems, the SMG took a two-track approach by focusing on the standardization
of the external information system in the short term and initiating a longer-term process to standardize the internal operational system. This approach later proved prudent as the SMG realized that the process of public data disclosure required multifaceted processes of (a) communicating with different departments, (b) standardizing data according to internal guidelines, (c) correcting errors, and (d) updating data over time.

**Persuading the Staff and Reorienting the Institutional Culture Are Crucial First Steps**

The essential role played by SMG staff members in information disclosure made their initial reluctance to embrace the idea of open information sharing a major early challenge. To reorient the institutional culture toward openness, the Information Disclosure Policy Division provided tailored training sessions to all grades of staff members. The training helped staff members understand the importance of information sharing and familiarized them with the information disclosure system. Frequent communication of guidelines, use cases, and protocols that addressed staff concerns and feedback also helped expedite the transition.

**Creating a Multilayered Monitoring System Can Ensure Safe, Long-Term Operations**

Because administrative documents contained sensitive information that is inappropriate for public disclosure, a strong monitoring system was indispensable. To ensure safe, long-term operations, the SMG developed a multilayered monitoring system that included both automatic filtering by special software and a manual review by trained personnel. As an extra check to avoid the accidental disclosure of sensitive information, the Information Disclosure Policy Division separately managed certain data sources with particular risks for unintended leakage.

**References**


