# Characterizing Data Sharing in Civil Infrastructure Engineering: Current Practice, Future Vision, Barriers, and Promotion Strategies

Yanyu Wang, S.M.ASCE[1]; Pingbo Tang, M.ASCE[2]; Kaijian Liu, A.M.ASCE[3]; Jiannan Cai, A.M.ASCE[4]; Ran Ren, Ph.D., S.M.ASCE[5]; Jacob J. Lin, A.M.ASCE[6]; Hubo Cai, M.ASCE[7]; Jiansong Zhang, A.M.ASCE[8]; Nora El-Gohary, A.M.ASCE[9]; Mario Berges, A.M.ASCE[10]; and Mani Golparvar Fard, A.M.ASCE[11]

**Abstract:** Data sharing between different organizations is critical in supporting decision making in civil infrastructure engineering projects (e.g., transportation projects). Understanding the characteristics of data-sharing-related factors in civil infrastructure engineering is crucial for the civil engineering community to identify the priorities in promoting data sharing. The ASCE Data Sensing and Analysis (DSA) Committee initiated an investigation on data-sharing barriers in civil infrastructure engineering. This work produced (1) a classification of data-collection and data-use contexts in the domain of civil infrastructure engineering; (2) a comprehensive classification of data-sharing barriers and data-sharing promotion strategies in this domain; and (3) a synthesis of the opinions about the current practices, future vision, barriers, and data-sharing promotion strategies from civil infrastructure professionals and researchers using an online survey. The survey results revealed that (1) more data-sharing activities are occurring within the organizations, such as academia, government agencies, and for-profit corporations, compared with data exchanges among these communities; (2) "if there is a cost sharing in the data sharing," is the condition that makes data providers more willing to share data with others; and (3) compared with technology improvements, the protocol of data sharing in the business domain has more potential in promoting data sharing. Meanwhile, this research identified the top five priorities among 20 data-sharing promotion strategies, which are protocol in industry and organization, data-standard regulation, data validity, data security protection, and large-scale data-sharing technology. These five priorities play critical roles in a roadmap showing how policy, business models, and technical solutions should cooperate to promote data sharing in civil infrastructure engineering. **DOI: 10.1061/JCCEE5.CPENG-5077.** *© 2023 American Society of Civil Engineers.*

**Author keywords:** Data sharing; Data-sharing promotion; Civil infrastructure engineering.

## Introduction

Data sharing in civil infrastructure is increasingly beneficial to decision making based on big-data analysis (Cooper et al. 2019; Olszak 2016). Recent studies showed that big data can improve the performance of business organizations by, on average, about 5.7% (Müller et al. 2018). Meanwhile, as a project-driven industry, the uniqueness of each civil infrastructure project makes the generated knowledge tightly related to specific projects' contexts (Agarwal and Dhar 2014; Antunes et al. 2017; Famakin and Abisuga 2016). Sharing data across projects and organizations in civil infrastructure would be essential to knowledge discovery.

[1]Ph.D. Candidate, Dept. of Civil and Environmental Engineering, Carnegie Mellon Univ., 500 Forbes Ave., Pittsburgh, PA 15213. Email: yanyuwan@andrew.cmu.edu

[2]Associate Professor, Dept. of Civil and Environmental Engineering, Carnegie Mellon Univ., 500 Forbes Ave., Pittsburgh, PA 15213 (corresponding author). ORCID: https://orcid.org/0000-0002-4910-1326. Email: ptang@andrew.cmu.edu

[3]Assistant Professor, Dept. of Civil, Environmental, and Ocean Engineering, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, NJ 07030. Email: Kaijian.Liu@stevens.edu

[4]Assistant Professor, School of Civil and Environmental Engineering, and Construction Management, Univ. of Texas at San Antonio, San Antonio, TX 78207. ORCID: https://orcid.org/0000-0001-6110-5293. Email: jiannan.cai@utsa.edu

[5]Automation and Intelligent Construction (AutoIC) Lab, School of Construction Management Technology, Purdue Univ., 401 N. Grant St., West Lafayette, IN 47907. Email: ren153@purdue.edu

Note. This manuscript was submitted on July 8, 2022; approved on November 9, 2022; published online on January 5, 2023. Discussion period open until June 5, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Computing in Civil Engineering***, © ASCE, ISSN 0887-3801.

[6]Assistant Professor, Dept. of Civil Engineering, National Taiwan Univ., No. 1, Section 4, Roosevelt Rd., Da'an District, Taipei 10617, Taiwan. Email: jacoblin@ntu.edu.tw

[7]Professor, Lyles School of Civil Engineering, Purdue Univ., 550 Stadium Mall Dr., West Lafayette, IN 47907. Email: hubocai@purdue.edu

[8]Director and Assistant Professor, Automation and Intelligent Construction (AutoIC) Lab, School of Construction Management Technology, Purdue Univ., 401 N. Grant St., West Lafayette, IN 47907. ORCID: https://orcid.org/0000-0001-5225-5943. Email: zhan3062@purdue.edu

[9]Associate Professor, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign, Mathews Ave., Urban, IL 16801. Email: gohary@illinois.edu

[10]Professor, Dept. of Civil and Environmental Engineering, Carnegie Mellon Univ., 5000 Forbes Ave., Pittsburgh, PA 15213. Email: marioberges@cmu.edu

[11]Professor, Dept. of Civil and Environmental Engineering, Computer Science and Tech Entrepreneurship, Univ. of Illinois at Urbana–Champaign, 205 N Mathews Ave., Urbana, IL 16801. Email: mgolpar@illinois.edu

Civil infrastructure projects are complex undertakings, often involving information flow across several parties (Ogunlana et al. 2003). Different types of data and multiple involved organizations can trigger various data-sharing barriers (Pesek et al. 2019). However, current research efforts lack a holistic synthesis of the current practice, future vision, and data-sharing barriers in civil infrastructure engineering. More importantly, the literature study of the authors found a lack of a synthesis of opinions about data-sharing promotion strategies from civil infrastructure engineers and project stakeholders. This research investigated data sharing in civil infrastructure engineering to explore the following four research questions:

- What are the current data-sharing practices in civil infrastructure engineering, including the data content, data format, and data-sharing approaches?
- What is the future vision of the data sharing in civil infrastructure engineering, such as the conditions where data providers tend to share data?
- What are the most challenging data-sharing barriers in civil infrastructure engineering?
- What are the preferred strategies for promoting data sharing in civil infrastructure engineering?

The investigation includes summarizing government reports, organization standards, related papers, face-to-face interviews with domain experts, and online surveys. The authors developed a survey that provides a holistic list of data-sharing-related factors and promotion strategies based on the summarization and interviews. The survey results help explore the data-sharing-related practices and opinions from practitioners in the civil infrastructure industry.

The organization of the remaining sections of the paper is as follows. The second section reviews the related data-sharing research in the civil engineering domain and summarizes the characteristics, future vision, barriers, and data-sharing promotion strategies in the civil engineering infrastructure domain. The third section defines the research goal and presents the research methodology. The fourth section illustrates the survey design process. The fifth section presents the survey results to reveal the current practice, future vision, barriers, and priorities of promotion strategies of data sharing according to what the civil infrastructure engineering practitioners expressed. The "Roadmap of Data Sharing" section provides a roadmap for promoting data sharing in civil infrastructure engineering. The last three sections discuss and summarize the limitations, future work, and main findings.

## Literature Review

### Data-Sharing-Related Studies in Civil Engineering

Several studies have shown the value of data sharing in civil infrastructure engineering, proving that data-sharing activities between different organizations could support high-quality decision making and help organizations achieve high performance. Sharing data between the building information modeling (BIM) and GIS data has improved site planning (Irizarry and Karan 2012), supply chain management (Irizarry et al. 2013), and energy optimization (Wang et al. 2019). The integration of radio frequency identification (RFID) and BIM allows tracking personnel in real-time, enhancing the safety of the construction site (Costin et al. 2015).

Meanwhile, from a long-term perspective, data analytics, decision making, and action feedback shared between different organizations would support the industry in summarizing domain knowledge (Cooper et al. 2019). A long-term collaboration among academics, industry, and government led by AbouRizk et al. (2011) aimed at extensive process modeling and simulation development in construction management. The modeling and simulation methods developed and tested could speed up modeling and simulation in other projects. The knowledge generated from long-term collaboration could benefit the development of the construction management industry.

Although shared data could improve organizations' performance (Manyika et al. 2011; Müller et al. 2018; Wang et al. 2016), a holistic investigation of civil infrastructure engineering's data-sharing activities is missing. Few data-sharing studies in the civil engineering domain discussed current practices, barriers, future vision, and promotion strategies. Although some research explored the current practice of data sharing in civil engineering, their explorations mainly focused on protecting the content of the data. The format of data and the data-sharing approaches received little attention in these studies. Meanwhile, most data-sharing research focused on the technical improvement of promoting data sharing, and limited research has discussed the policy and business aspects of encouraging data sharing. Also, little research has surveyed practitioners' opinions on the proposed promotion strategies. This research examines detailed data characteristics in civil infrastructure engineering, summarizes the potential promotion strategies on technology, policy, and business aspects, and analyzes different organizational practitioners' opinions about data sharing as a system for project executions. Table 1 summarizes the aspects of data-sharing efforts discussed in the published study and compares them with this research.

### Characteristics of Data in Civil Infrastructure Engineering

Many characteristics of data and data-use contexts can influence data sharing and data-sharing approaches. We reviewed the literature to answer the following questions:

- What data contents are captured in what data-use contexts?
- What data formats are used in civil infrastructure engineering practice?
- What data-sharing approaches are used in civil infrastructure engineering practice?

A comprehensive review can help the community understand data sharing with details such as data formats, data contents, data-use contexts, data-sharing approaches in current practice, and engineers' vision about future data-sharing practice.

This research formulates a framework for classifying data contents and data-use contexts based on information from the official websites and data sets released from agencies working on transportation asset management (AKDOT 2019; AZDOT 2019; BOTS 2015; ILDOT 2019; MAG 2019; NTAD 2011; ODOT 2019; USDOTFHA 2019; WSDOT 2019). We first identified the ontology of the data-related information in the cited government websites. Then, we used the nominal group technique (NGT) to discuss and group this data-related information into five categories of data-use contexts: (1) infrastructure design, (2) geotechnical and hydraulic design, (3) infrastructure contract and financial administration, (4) infrastructure safety and maintenance, and (5) infrastructure data inventory. Each category of data-use context involves specific data contents, as listed in Table 2. Table 2 also summarizes the data content and data-use contexts identified.

The data format is also a factor influencing the technical feasibility of data sharing. Table 3 summarizes the data format identified from literature and data standards used by civil engineers (ASTM 2019; BuildingSmart 2019; Chen 2019; OGC 2019).

Heterogeneous data formats may have different requirements for data-sharing approaches and raise diverse challenges in the data-sharing process. For example, video data could overload cyberinfrastructure. This research synthesizes the data-sharing

**Table 1.** Data-sharing-related research in civil engineering domain

| References | Data-sharing domain | Data-sharing current practice | Data-sharing barriers | Data-sharing future vision | Data-sharing promotion strategy on technology | Data-sharing promotion strategy on policy | Data-sharing promotion strategy on business | Practitioners' priority on promotion strategy | Different organizational practitioners' priority on promotion strategy |
|---|---|---|---|---|---|---|---|---|---|
| Thu and Wehn (2016) | Hydrology | X | X | X | X | — | — | — | — |
| Wang et al. (2020) | Hydrology | — | — | — | X | — | — | — | — |
| Petri et al. (2017) | Construction | — | — | — | X | — | — | — | — |
| Mutis and Mehraj (2022) | Construction | X | — | — | — | X | — | X | — |
| Karaguzel et al. (2019) | Building energy | — | — | — | X | — | — | — | — |
| Costin and Eastman (2019) | Urban systems | — | X | X | X | X | — | — | — |
| Cooper et al. (2019) | Civil engineering academic research | X | X | X | X | X | — | — | — |
| This research | Civil infrastructure engineering | X | X | X | X | X | X | X | X |

Note: X = discussed; and dash = not applicable.

**Table 2.** Data-use contexts and data contents in the civil infrastructure engineering

| Data-use context | Data content |
|---|---|
| Infrastructure design | • Steel structures<br>• Concrete structures<br>• Load rating structures<br>• Seismic and multihazard resilience structure<br>• Security structures<br>• Utility design |
| Geotech and hydraulic design | • Earth and rock works<br>• Earth-retaining structures<br>• Geohazards<br>• Ground improvement<br>• Structural foundations<br>• Subsurface investigation<br>• Hydraulics planning<br>• Hydrology<br>• Drainage design and analysis<br>• Drainage erosion and sediment control<br>• Hydraulic design and analysis<br>• Aquatic organism passage design<br>• Asset management |
| Infrastructure construction | • Steel structures<br>• Concrete structures<br>• Utility construction<br>• Construction material records<br>• Three-dimensional engineered models |
| Infrastructure contract and financial administration | Contract and financial administration documents |
| Infrastructure safety and maintenance | • Construction safety records (accident/incident)<br>• Inspection records (including Structural Health)<br>• Crash data in operation<br>• Traffic data in operation<br>• Maintenance records |
| Infrastructure data inventory | • Current and historical<br>• General information inventory<br>• Human services transportation<br>• Land use<br>• Future general information inventory<br>• Future human services transportation<br>• Future land use<br>• Border and boundaries information<br>• Weather |

techniques used in practice by interviewing six infrastructure engineering domain experts. Table 4 presents typical data-sharing approaches identified.

The authors highlight two types of metadata in civil infrastructure engineering as critical for data sharing: (1) data dictionaries, and (2) data ontology. Data shared between different organizations may contain various attributes. Data dictionaries specify the meaning of various attributes in data. Data ontology describes the hierarchal structure and relationships between these attributes. Data dictionaries and data ontology could help prevent inconsistency in data and assist data users in understanding the meaning of data (Benoît et al. 2004; El-Gohary and El-Diraby 2010; Noy and Hafner 1997; Noy and McGuinness 2001). Data dictionaries have appeared

**Table 3.** Data formats in the civil infrastructure engineering

| Data formats | Format descriptions |
|---|---|
| Text, documentation, and/or scripts | Examples of such data formats include TXT, PDF, HTML, and extensible markup language (XML) |
| Still image | Examples of such data formats include TIFF, JPEG2000, PNG, JPEG/JFIF, DNG, and BMP |
| Geospatial | Examples of such data formats include Shapefile (SHP, DBF, and SHX), GeoTIFF, NetCDF, GML, and MDB |
| Graphic image-raster formats | Examples of such data formats include GIF, TIFF, JPEG2000, PNG, JPEG/JFIF, DNG, and BMP |
| Graphic image-vector formats | Examples of such data formats include AutoCAD Drawing Interchange Format, Scalable vector graphics, Encapsulated Postscripts, and Shapefiles |
| Graphic image-cartographic | Examples of such data formats include GeoTIFF, GeoPDF, GeoJPEG2000, and Shapefile |
| Audio | Examples of such data formats include WAVE, AIFF, MP3, MXF, and FLAC |
| Video | Examples of such data formats include MPEG-4, AVI, MOV, and MXF |
| Metadata (data about data) | Describe the content, quality, condition, and other characteristics of data, such as data dictionary and data ontology. Examples of such data formats include XML, JSON, GML, IFC, and OWL |
| Database (including tabular data) | Examples of such data formats include CSV, TAB, SPSS portable format (POR), XLS, and MDB/ACCDB (MS Access) |
| Three-dimensional (3D) imaging (3D point data) | Examples of such data formats include LAS, OBJ, PLY, XYZ, E57, PTX, ASC, FBX, PCG, RCS, RCP, and XML |
| BIM model (or building data) | Examples of such data formats include RVT, IFC, NWD (Navisworks), DXF, and DWG |

**Table 4.** Data-sharing approaches in the civil infrastructure engineering

| ID | Data-sharing approaches |
|---|---|
| 1 | FTP server: different from cloud server, FTP server stands for file transfer protocol |
| 2 | Cloud server: cloud server is a general term for a variety of services like cloud storage, e.g., Amazon Web Services (AWS) Cloud and Google Cloud Platform (GCP) |
| 3 | Project management solutions, such as BIM360, Procore, CMIC |
| 4 | Data-sharing centers, such as SharePoint |
| 5 | Open platforms, such as indianamap.org and NRC Human Reliability database |
| 6 | Paper-based documents |

in academic research and government and professional organizations' open platforms. For example, Benoît et al. (2004) proposed a data dictionary and formatting standard to disseminate geotechnical data. The Department for Transport of England (England Highway 2021) published a data dictionary for the highway system. BuildingSmart (2015) developed a data dictionary for promoting BIM applications (ISO 2007).

However, data dictionaries alone could hardly support data-sharing applications because data dictionaries only explain the corresponding data attributes but not the structure or relationship between different attributes. Some recent studies improved the efficiency of data retrieving and processing based on data ontologies that can utilize relationships between different data attributes for processing data sets with various data attributes. For example, Zhang and Issa (2013) proposed an ontology-based model extraction methodology. Liu and El-Gohary (2017) proposed an ontology-based bridge condition and maintenance information extraction methodology. Zhong et al. (2020) proposed an ontology-based hazard identification methodology with computer vision technology.

## Future Vision of Data Sharing in Civil Infrastructure Engineering

A list of conditions for achieving an ideal data-sharing system could lead to a tool for collecting a shared vision of engineers about data sharing for supporting practical civil infrastructure engineering. The interview results and the summarization from government reports, organization standards, and related papers helped the authors compile a list of possible conditions that motivate data-sharing behaviors.

**Table 5.** Conditions of data sharing from data providers in civil infrastructure engineering

| ID | Conditions of data sharing from data providers in civil infrastructure engineering |
|---|---|
| 1 | If sharing the data enabled me to get into contact with other partners |
| 2 | If I had enough time beforehand to process my data |
| 3 | If my employer supported me actively (e.g., by providing technical support and time) |
| 4 | If I knew the intention of requesters very clearly |
| 5 | If I knew who would be able to access the data |
| 6 | If I would share the benefit from the outcomes of requesters |
| 7 | If there is a cost-sharing in the data sharing |
| 8 | If I get the financial benefit for my effort |

The online survey asked respondents to rate their agreement level on each item in this list to explore under what conditions the organizations would be willing to share certain types of data. Table 5 presents the compiled condition list.

A summary of data-sharing practices in other fields could benefit the development of data-sharing guidelines in civil infrastructure engineering. Open data is among the least restrictive forms of data sharing, in contrast to managed access mechanisms, which typically have terms of use and, in some cases, oversight by the data generators (Serwadda et al. 2018). The biomedical domain has been promoting open data sharing. For example, more than 8,000 shared neuroimage data sets are available online. Data exchanges between countries and organizations contributed to the battle against the COVID-19 pandemic (Poldrack and Gorgolewski 2014; Rios et al. 2020). The practice in the biomedical domain shows that various data shared have different potentials for reuse and require varied sharing costs. The raw and preprocessed data would have the highest potential for reuse, whereas the aggregated data have less potential for reuse. The cost of sharing raw and preprocessed data would be higher than that for aggregated data (Poldrack and Gorgolewski 2014).

Data requiring less effort to avoid privacy and proprietary issues in the civil domain have more open data sets. Geospatial data have fewer privacy concerns, so multiple open geospatial data sets, such as Esri Open Data Hub, Natural Earth Data, and OpenStreetMap, are available. Civil engineers are trying to extend the benefits of data sharing by adopting advanced technology to protect privacy and proprietary. For example, Sangogboye et al. (2018) proposed

**Table 6.** ISO standards for promoting data sharing

| ISO standard | Description |
|---|---|
| ISO 20614:2017: *Information and documentation—data exchange protocol for interoperability and preservation* | ISO 20614:2017 specifies a standardized framework for the various data (including data and related metadata) exchange transactions between an archive and its producers and consumers (ISO 2017). |
| ISO/IEC 27010:2015: *Information technology—security techniques—information security management for inter-sector and inter-organizational communications* | ISO/IEC 27010:2015 provides guidelines in addition to the guidance given in the ISO/IEC 27000 family of standards for implementing information security management within information sharing communities (ISO 2015a). |
| ISO 8000: *Data quality and enterprise master data* | ISO 8000 defines the quality requirements of the data exchange in the business process (ISO 2011). |
| ISO/IEC/IEEE 15288:2015: *Systems and software engineering—system life cycle processes* | ISO/IEC/IEEE 15288:2015 establishes a common framework of process descriptions for describing the life cycle of systems created by humans. It defines a set of processes and associated terminology from an engineering viewpoint (ISO 2015b). |
| ISO 15926: *Industrial automation systems and integration—integration of life-cycle data for process plants including oil and gas production facilities* | ISO 15926 specifies a representation of information associated with the engineering, construction, and operation of process plants (ISO 2003). It provides a standard for data integration, sharing, exchange, and hand-over between computer systems. |
| ISO 19650-5:2020: *Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM)—information management using building information modelling—Part 5: security-minded approach to information management* | ISO 19650-5:2020 specifies the principles and requirements for security-minded information management at a stage of maturity described as "building information modeling (BIM) according to the ISO 19650 series," and as defined in ISO 19650-1, as well as the security-minded management of sensitive information that is obtained, created, processed, and stored as part of, or in relation to, any other initiative, project, asset, product or service (ISO 2020a). |
| ISO 16739:2013: *Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries* | ISO 16739:2013 specifies a conceptual data schema and an exchange file format for BIM data. The conceptual schema is defined in EXPRESS data specification language. The standard exchange file format for exchanging and sharing data according to the conceptual schema uses the Clear text encoding of the exchange structure. Alternative exchange file formats can be used if they conform to the conceptual schema (ISO 2013). |
| ISO 37156:2020: *Smart community infrastructures—guidelines on data exchange and sharing for smart community infrastructures* | ISO 37156:2020 gives guidelines on principles and the framework to use for data exchange and sharing for entities with authority to develop and operate community infrastructure (ISO 2020b). |

a framework for promoting data sharing with privacy preservation in building systems. Schwee et al. (2019) proposed a method for anonymizing building data for cross-organizational data analytics. Karaguzel et al. (2019) developed a data-sharing platform to support building designers using empirical data from multiple locations with privacy protection. Janghyun et al. (2022) reviewed preserving privacy in data collected from buildings and analyzed different privacy implementation approaches. Further integration of civil engineering data research and privacy computing is an emerging topic.

General data-sharing principles provide guidelines for designing data-sharing agreements or standards among different stakeholders. In 2016, Wilkinson et al. (2016) proposed the findable, accessible, interoperable, and reusable (FAIR) guiding principles for scientific data management and stewardship. Data sharing between machines and people can adopt this principle as well. Meanwhile, the emphasis on data sharing between machines is a specific focus of this principle because the support from computational tools becomes increasingly important in this age (Wilkinson et al. 2016). In 2018, Carroll et al. proposed collective benefit, authority to control, responsibility, and ethics (CARE) principles for indigenous data governance (Carroll et al. 2020). These principles complete the rights and interests of indigenous people with a focus on power differentials and historical contexts (Carroll et al. 2020). In 2019, the Australian government published a best-practice guide to applying data-sharing principles from multiple aspects, including project assessment, people authority, data-sharing settings, data protection, and output from the data sharing (Australian Government 2019).

Consistent data exchange is critical for fusion and integrated decision making with data from multiple organizations. Existing data-sharing-related standards may serve as a beacon in guiding the future development of data-sharing standards in civil infrastructure engineering. ISO provides a series of data managing and exchanging standards semantic for various stakeholders participating in the infrastructure life cycle (ISO 2003, 2011, 2013, 2015a, b, 2017, 2020a, b). We researched and summarized data managing and sharing related ISO standards in Table 6.

ISO is trying to develop more standards in data sharing, such as ISO 37172:2002, a new data exchange and sharing standard on geographic information (ISO 2022). Besides ISO, the United Engineering Foundation delivered its *Infrastructure Architecture Framework* in 2021, which provides suggestions for the development of data sharing in civil infrastructure (Costley et al. 2021). Currently, limited civil infrastructures have specific international data-exchanging standards. Although BIM and smart community infrastructures have international data-exchanging standards, innovative community infrastructure has only a guideline on data exchanging instead of a detailed standard for defining data quality and data security requirements. More specific international data-sharing standards are thus necessary for promoting civil infrastructure development.

## Five Categories of Data-Sharing Barriers

Two purposes related to understanding data-sharing barriers guided the survey design related to these barriers. First, summarizing the data-sharing barriers could help list the potential options for the survey participants to rate the severity of each barrier by marking the level of challenge. Second, the summarization of data-sharing barriers could help the authors identify data-sharing promotion

**Table 7.** Data-sharing barriers in the civil infrastructure engineering

| ID | Data-sharing barriers in the civil infrastructure engineering | Barrier category |
|---|---|---|
| 1 | There is a lack of awareness of existing platforms/avenues for sharing or requesting data | Data-sharing awareness |
| 2 | Others would use my data to publish before me | Social and organizational factors |
| 3 | Others would criticize my work | |
| 4 | Others would falsify my work | |
| 5 | The data could be misinterpreted | |
| 6 | Data are sensitive | |
| 7 | Data are private | |
| 8 | Data are of competitive edges | |
| 9 | Lack of standard, policy, or consensus regarding data sharing | |
| 10 | Ethical and legal issues causing concerns about data ownership, legal implications, and intellectual property (both for individuals and organizations) | |
| 11 | Lack of comprehensive data security technology support | |
| 12 | The data structure is too complex to the extent of making data sharing worthless (e.g., others do not know how to use the data) | Data and software technical issues |
| 13 | Lack of quantification of data quality | |
| 14 | Lack of automatic technology for extraction, integration, and analysis of information from various data formats, such as integrated analysis of images and texts | |
| 15 | Cyberinfrastructure could not support big-data use (e.g., limited internet bandwidth, storage space, computing power, and so on) | Hardware and cyberinfrastructure issues |
| 16 | Data access is not available | |
| 17 | Major efforts are required to collect the data | Finance and workforce issues |
| 18 | Major efforts are required to share the data | |
| 19 | Engineering/technical support for data sharing is limited | |
| 20 | Financial and human resources with a computational mindset for data sharing are limited | |

strategies that overcome those barriers. Each data-sharing promotion strategy should address one or more data-sharing barriers. The classification of data-sharing barriers would help organize those strategies into categories.

The authors summarized five categories of data-sharing barriers in civil infrastructure engineering as detailed in Table 7, including (1) the lack of data-sharing awareness, (2) social and organizational factors, (3) software and data technology issues, (4) hardware and cyberinfrastructure issues, and (5) finance and workforce issues (Attard et al. 2015; Balaji 2018; Beaulieu-Jones et al. 2019; Becker et al. 2016; Bradley et al. 2016; Dikeocha et al. 2018; DOE 2014, 2018; Dyke and Hubbard 2011; EUROPA 2016; Guo et al. 2017; Halfawy 2008, 2010; Le and David Jeong 2017; Lee et al. 2006; Ng et al. 2017; Nguyen et al. 2017; Pryor 2009; Rathje et al. 2017; Sangogboye et al. 2018; Wall et al. 2005; Zakaib 2019; Zuiden 2019). Table 7 lists the data-sharing barriers under each category.

### *Three Categories of Data-Sharing Promotion Strategies*

Summarizing data-sharing promotion strategies can help design a survey comprehensively studying perceptions from a wide range of respondents who evaluated these strategies' effectiveness in overcoming data-sharing barriers. The authors summarized various solutions discussed in the literature for overcoming specific data-sharing obstacles. The solutions proposed in the literature formed 50 actions, which can belong to 15 solution groups and three categories, as shown in Fig. 1 (Attard et al. 2015; Balaji 2018; Beaulieu-Jones et al. 2019; Becker et al. 2016; Bradley et al. 2016; Dikeocha et al. 2018; DOE 2014, 2018; Dyke and Hubbard 2011; EUROPA 2016; Guo et al. 2017; Halfawy 2008, 2010; Le and David Jeong 2017; Lee et al. 2006; Ng et al. 2017; Nguyen et al. 2017; Pryor 2009; Rathje et al. 2017; Sangogboye et al. 2018; Wall et al. 2005; Zakaib 2019; Zuiden 2019).

Each solution group in Fig. 1 has an ID, such as T.1, to represent a set of strategies that can form a group of associated solutions to address data-sharing barriers in the corresponding category. Each solution group consists of multiple capsules representing actions needed in data-sharing promotion. Codes for capsules show the corresponding categories of data-sharing barriers (as listed in Table 7) that the capsules could address. Notably, strategies in the business category are related to marketing strategies, and social support, such as creating education and training programs for big-data sharing listed in B.6 in Fig. 1. The education and business system must provide integrated support in training graduates and workers with a computational mindset. An updated curriculum is essential to overcome two categories of barriers: (1) finance and workforce problems, and (2) social and organizational factors, as detailed in Fig. 1.

Based on the content of each solution group and interviews with domain experts, the research team compacted the 50 actions listed in Fig. 1 into high-level approaches under three categories of strategies to represent technical, policy, and business solution groups, as listed in Table 8. The compacted data-sharing strategies are more visual-friendly for survey participants to provide their opinions on each strategy.

The results of each strategy's importance level can help identify critical solutions to promote data sharing from the business, policy, and technology perspectives. Table 9 presents the relationships between data-sharing promotion strategies and the 15 solution groups. Each solution group in Fig. 1 corresponds to one or multiple strategies in Table 8.

## Research Methodology

The overall goals of the research are to (1) explore the current practice and future vision of data sharing in civil infrastructure engineering, and (2) characterize data-sharing barriers and identify the practical priority of data-sharing promotion strategies for overcoming those barriers in civil infrastructure projects. To achieve this goal, the project team designed a survey based on a summarization from government reports, organization standards, and related papers of civil infrastructure–related government public materials, organization standards, and research papers. To capture


**Fig. 1.** Data-sharing policy, technology, and business solution groups.

**Table 8.** Data-sharing promotion strategies in the civil infrastructure engineering

| ID | Data-sharing promotion strategies in the civil infrastructure engineering | Group of data-sharing promotion strategies |
|---|---|---|
| 1 | Data security protection | Technology strategies |
| 2 | Data anonymization and privacy-preservation | |
| 3 | Data quality measurement and information loss measurement | |
| 4 | The technology of translating heterogeneous data into consistent data | |
| 5 | Enhancement of data labeling and ground truth | |
| 6 | Large-scale data-sharing technology | |
| 7 | Real-time data-sharing technology | |
| 8 | The intelligent data-sharing method | |
| 9 | Policy on data liability | Policy strategies |
| 10 | Data standard regulation | |
| 11 | Incentive on data-related research (e.g., establish testbeds to create artificial data) | |
| 12 | Dissemination of available resources and the value of data sharing (e.g., workshops and best practices) | |
| 13 | Data security and privacy protection (e.g., intellectual property agreements, privacy protection) | |
| 14 | Data value demonstration (i.e., quantitative and qualitative) | Business strategies |
| 15 | Protocol in industry and organization (e.g., policy and regulation) | |
| 16 | Data validity (e.g., standard and type of data for data owner and receiver) | |
| 17 | Repository and platform for data sharing | |
| 18 | Risk analysis and risk mitigation | |
| 19 | Skills of workers, engineers, and stakeholders | |
| 20 | Financial, human, and other resources | |

most factors potentially influencing data sharing in practice, the authors conducted a few expert interviews before the online survey to cross-check summarized factors that potentially influence data-sharing behaviors in civil infrastructure projects. Based on the summarization and expert interviews, the research team created a set of questions used in a survey to comprehend those data-sharing behaviors. Fig. 2 shows the research approach and corresponding outcomes of each research activity.

The research team developed a survey to solicit opinions and feedback from a broader range of civil infrastructure engineering practitioners about data-sharing practices in their work. The research team summarized data characteristics from the aspects of

**Table 9.** Data-sharing solution groups integrating data-sharing promotion strategies

| Technology solution groups | Technology strategy | Policy solution groups | Policy strategy | Business solution groups | Business strategy | Business solution groups | Business strategy |
|---|---|---|---|---|---|---|---|
| T.1 | 1 | P.1 | 11 | B.1 | 14, 17, 20 | B.5 | 18 |
| T.2 | 2 | P.2 | 12 | B.2 | 14 | B.6 | 19, 20 |
| T.3 | 3, 4 | P.3 | 9, 10, 13 | B.3 | 15, 17 | B.7 | 15, 18, 19 |
| T.4 | 5, 6, 7, 8 | P.4 | 13 | B.4 | 16 | — | — |

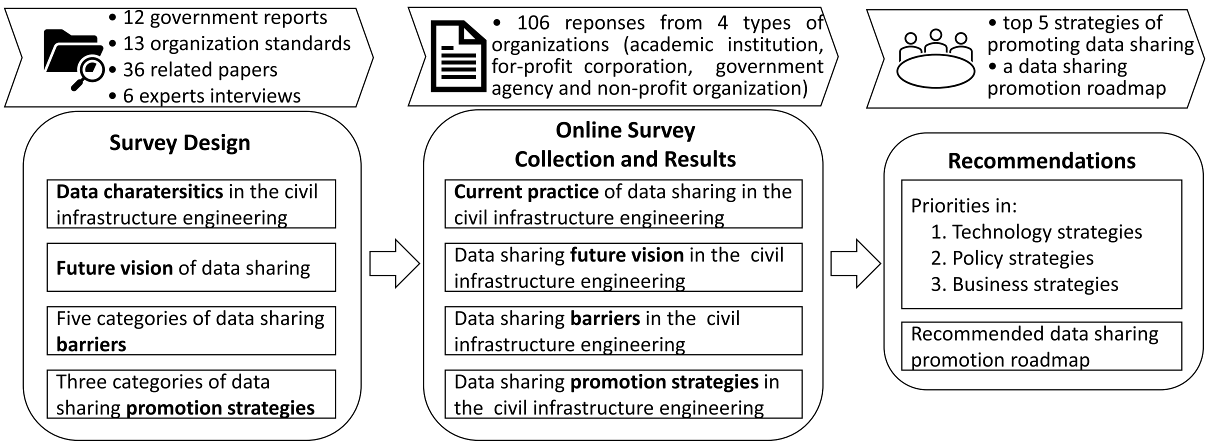


**Fig. 2.** Overview of research methodology.

data-use contexts, data contents, and data formats based on all the available government public documents, including data classifications on 12 government reports and 13 organization standards. After summarizing 36 data-sharing-related papers and face-to-face interviews with six experts who have worked in the civil engineering domain for more than 20 years, the research team synthesized a list of conditions that motivate data providers and barriers to data sharing. Further, we organized technology, policy, and business categories of strategies to overcome data-sharing barriers. The outcomes of the summarized data characteristics, future wishes, data-sharing barriers, and data-sharing promotion strategies formed the survey.

Using the Qualtrics platform (Qualtrics 2020), the research team has carried out a half-year-long online survey with more than 200 practitioners in civil engineering and collected a total of 106 responses. Analyzing these responses enabled us to (1) explore the current practice of data sharing in civil infrastructure engineering; (2) formulate a future vision of data sharing in civil infrastructure engineering; (3) examine five categories of data-sharing barriers among different data-sharing activities; and (4) prioritize data-sharing strategies under three categories (technology, policy, and business strategies) for promoting data sharing in civil infrastructure engineering. The identified priorities in data-sharing promotion strategies from the survey results helped the research team propose a roadmap to improve the current data sharing in civil infrastructure engineering.

## Survey Design

Surveys are a widely used approach for perception studies (Azeez et al. 2019). This survey consisted of four sections to explore the opinion of civil infrastructure engineering practitioners about data sharing. The four sections are (1) the current practice of data sharing in civil infrastructure; (2) a future vision of the data-sharing system for the domain of civil infrastructure engineering; (3) data-sharing barriers in practice; and (4) data-sharing promotion strategies under three categories, including technology, business, and policy.

We asked respondents about data-sharing-related opinions in each section through multiple choice or rating questions. The literature review results formed the choices or rating options in the survey. We used NGT to discuss and summarize the literature review findings into choices and rating options. The NGT process enabled us to collect extensive ideas and summarize comprehensive results than standard group discussions (Burrows et al. 2011; Robinson and Shepard 2011). Each team member presented their literature review findings on one aspect of the data-sharing activities, and the team summarized and discussed all findings together. We also interviewed six experts to finalize the choices in the survey. Specifically, in the "Current Practice" section of the survey, we asked respondents to report their current data-sharing activities using multiple choice questions. We provided the choices of data format, data content, data context, and data-sharing approach based on the literature review results, as presented in Tables 2–4.

In the Future Vision section of the survey, we asked respondents about their wanted or willingness to share data format, data content, data context, and data-sharing approaches based on their role in the data-sharing activities (i.e., data provider or data requester). We also asked data providers to rate the conditions in which they are willing to share data. The conditions are summarized in Table 5. In the Data-Sharing Barriers section of the survey, we asked respondents to rate the data-sharing barriers summarized in Table 7. In the Data-Sharing Promotion Strategies section of the survey, we asked respondents to rate the data-sharing strategies as summarized in Table 8.

**Table 10.** Different numbers of responses in different sections of survey

| Section name | Response count |
|---|---|
| Current practice | 45 |
| Future vision | 56 |
| Barriers | 20 |
| Data-sharing promotion strategies | 25 |

## Survey Results

This section provides the survey results about data sharing in the civil infrastructure domain, including the aspects of (1) current practice, (2) future vision, (3) barriers, and (4) promotion strategies in data sharing. The Current Practice section revealed the present data sharing between different organizations and the corresponding characteristics of the shared data. The Future Vision section asked about the perceived gap between data types wanted by the data receiver, the data provider's willingness to share such data, and the conditions where the data providers would agree to share. The Barriers section allowed survey participants to express the severities (levels of the challenges) of data-sharing barriers they perceive. The last subsection captured the importance level of different data-sharing promotion strategies perceived by the survey participants.

From February 24 to August 26, 2020, the research team reached out to more than 200 engineering practitioners and collected 106 responses from various professionals working in the civil infrastructure domain (Wang 2019). Respondents were from academic institutions, government agencies, for-profit corporations, and nonprofit organizations. The number of respondents varied in different sections because the survey allowed responders to skip some less relevant questions according to their roles in data sharing. Table 10 illustrates the number of respondents in each section of the survey.

### *Current Practice of Data Sharing in the Domain of Civil Infrastructure Engineering*

The survey collected 45 responses for the Current Practice section, 25 of which were from academic institutions. Fig. 3 visualizes the current practice of data-sharing activities between different organizations. The width of the link indicates the strength of data-sharing activities between different organizations. The number in the brackets shows the number of responses from the corresponding organizations.

In the survey results, 43% of data providers came from academic institutions, 29% from government agencies, 14% from for-profit corporations, and 14% from nonprofit organizations. In the survey results, most of the data receivers were in academic institutions (51%), followed by government agencies (25%), for-profit corporations (16%), and nonprofit organizations (8%). The results showed that (1) most organizations share data with the same type of organizations; for example, academic institutions mostly share data with other academic institutions; (2) nonprofit organization is the only organization type that shares data with government agencies more actively than sharing with of the organizations of their own type; and (3) compared with other organization types, government agencies share data with different organizations more equally.

Fig. 4 shows the details of data-use context in the data-sharing activities among different organizations, with different types of data context indicated. The radius of circles shown in Fig. 4 represents the count from the survey about how often the respondents selected a particular type of data using context and the corresponding data receiving organization.

Fig. 4 indicates that infrastructure construction is the most active context of data-sharing. The contexts of infrastructure safety and maintenance and infrastructure design also have many data-sharing indicators per survey results. Besides, data sharing among nonprofit and for-profit organizations are limited. Fig. 4 also shows that government agencies are active in data-sharing activities with the most diverse contexts. In addition, infrastructure contracts and financial administration data exchange are mostly between government agencies and for-profit corporations.

Table 11 presents the formats and approaches of data sharing between different organizations. The results show that (1) BIM model data have limited data-sharing activities compared with other types of data, although BIM models have been promoted for years; (2) paper-based documents are still popular, even though high-tech internet connections exist today; and (3) currently, data-sharing

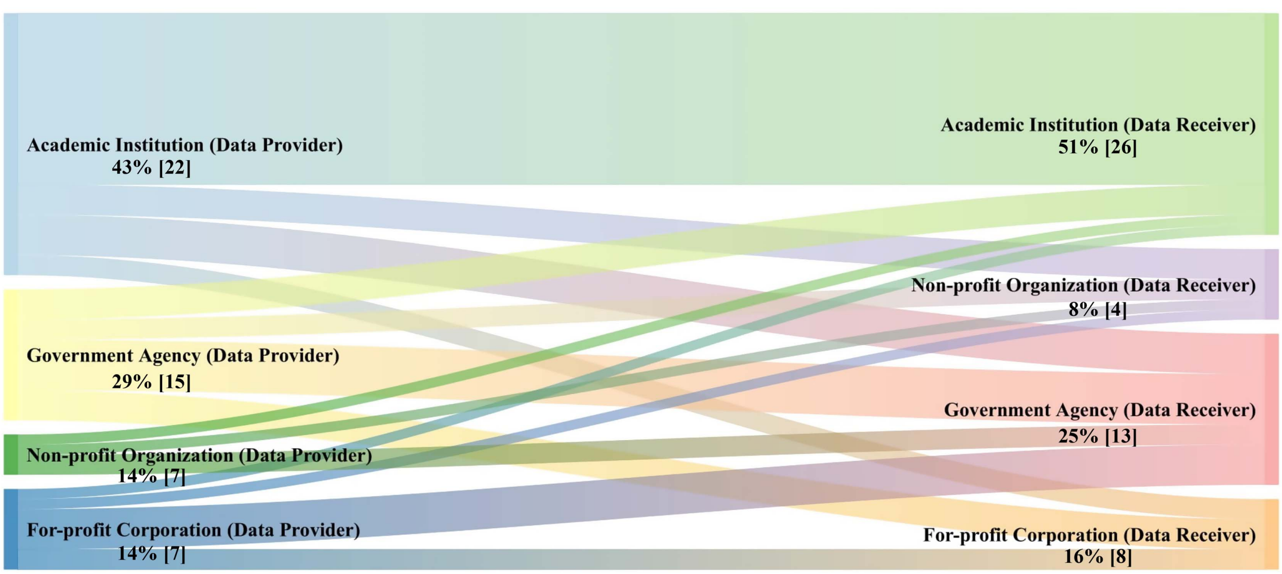


**Fig. 3.** Strength of current data-sharing practice among different organizations.

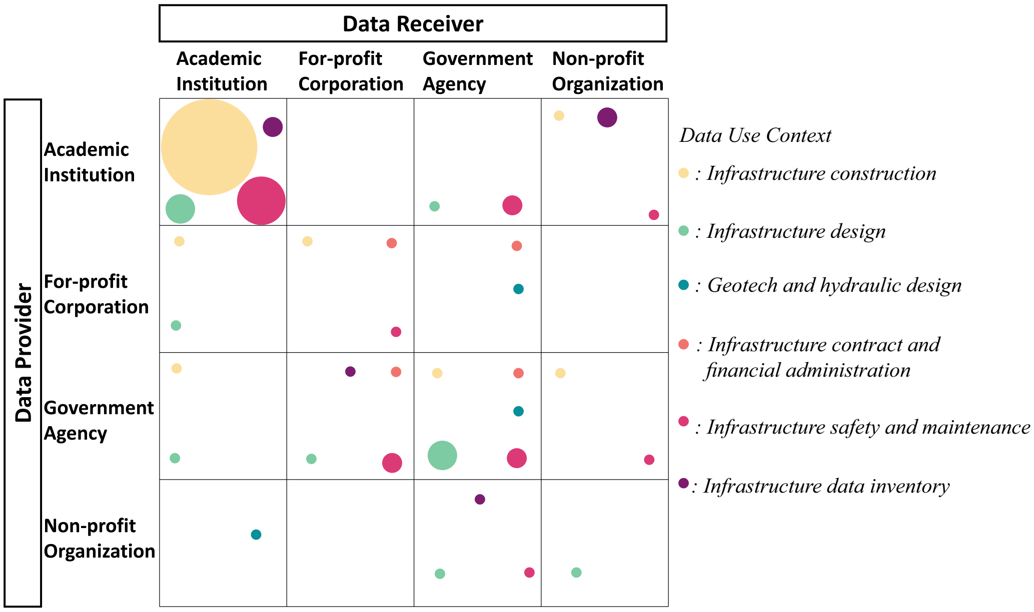


**Fig. 4.** Number of data-use contexts in responses from different types of organizations participating in data sharing.

**Table 11.** Number of data formats and data-sharing approaches in current practice responses from all organizations participating in data sharing

| Data format | Count of the data format responses | Data format | Count of the data format responses | Data-sharing approach | Count of the data-sharing approach responses |
|---|---|---|---|---|---|
| Text | 22 | Audio | 8 | FTP server | 15 |
| Still image | 20 | Video | 8 | Cloud server | 15 |
| Geospatial | 11 | Database | 18 | Open platforms | 7 |
| Graphic image-raster formats | 6 | 3D imaging (3D point data) | 8 | Data-sharing centers | 11 |
| Graphic image-vector formats | 10 | BIM model (or building data) | 6 | Project management solutions | 14 |
| Graphic image-cartographic | 9 | Metadata (data about data) | 14 | Paper documents | 15 |

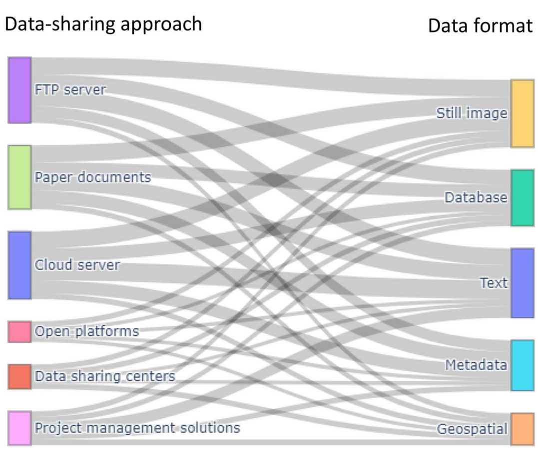


**Fig. 5.** Current data-sharing activities using various data-sharing approaches for data in multiple formats.

open platforms seem to contribute less, as revealed in the survey results.

Fig. 5 further analyzes the relationship between the data-sharing approaches and the top-five data formats in the current data-sharing practice. The results show that no significant preferences for different data formats may prefer a specific data-sharing approach. Data in the still-image, text, database, and metadata formats are mostly shared through the FTP server, paper documents, and cloud servers.

### Future Vision about Data Sharing in the Domain of Civil Infrastructure Engineering

There are 56 responses in the Future Vision section. The survey results of the Future Vision focus on (1) the conditions where the data providers would be willing to share data with other organizations; and (2) the data types that data receivers want and the types that data providers would be willing to share. Specifically, the survey results in the Future Vision section have two parts. Part 1 is a ranking of the conditions of sharing data from different data provider organizations. Such ranking results could help the data receivers understand the data providers' needs and motivate them to share data. Part 2 is a comparison between the wanted data from data receivers and willing to share data from data providers from the aspects of data-use context, data format, and data-sharing approaches.

Fig. 6 shows the responses about the conditions the stakeholders would like to share the data, thereby providing a vision of future data-sharing systems. The results show that (1) "if there is a cost-sharing in the data sharing" is the condition that achieves a relatively high consistent agreement level among all organizations;

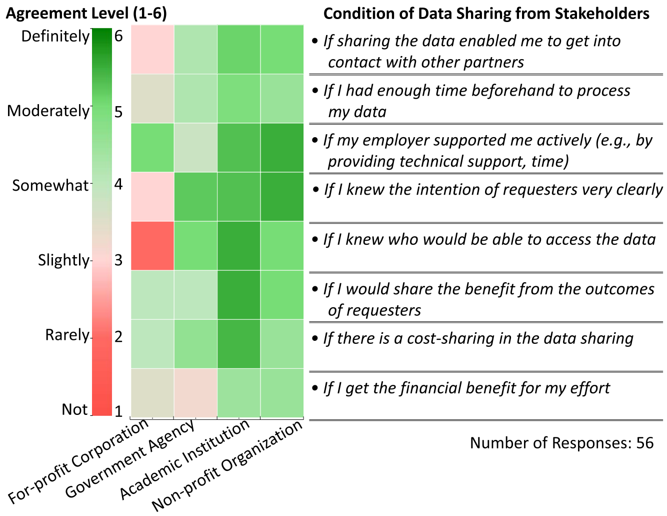


**Fig. 6.** Future vision of a better data-sharing system: data-sharing conditions perceived by different organizations (corresponding to conditions listed in Table 5).

(2) for for-profit corporations, the conditions that "if I knew who would be accessing the data" and "if I knew the intention of requesters very clearly," are relatively lower motivations for them to share data, whereas all other organization types view these two conditions as strong motivations; and (3) academic institutions achieved the highest average agreement level and viewed all conditions as good motivations, but for-profit corporations achieved the lowest average agreement level and viewed many conditions as ineffective motivations. For-profit corporations seem only to emphasize "if my employer supported me actively" as a compelling motivation. There is a perception gap about data-sharing motivation strategies between for-profit and nonprofit organizations. Sharing from for-profit corporations is more conservative than academic institutions because they have limited compelling motivation in data sharing.

Fig. 7 shows the data-sharing contexts, format, and approach that data providers would be willing to share and what data receivers would need. The results showed the following:

- Infrastructure construction is the context where providers are most willing to share, and the data receiver needs data the most, whereas the interest in data-sharing activity focusing

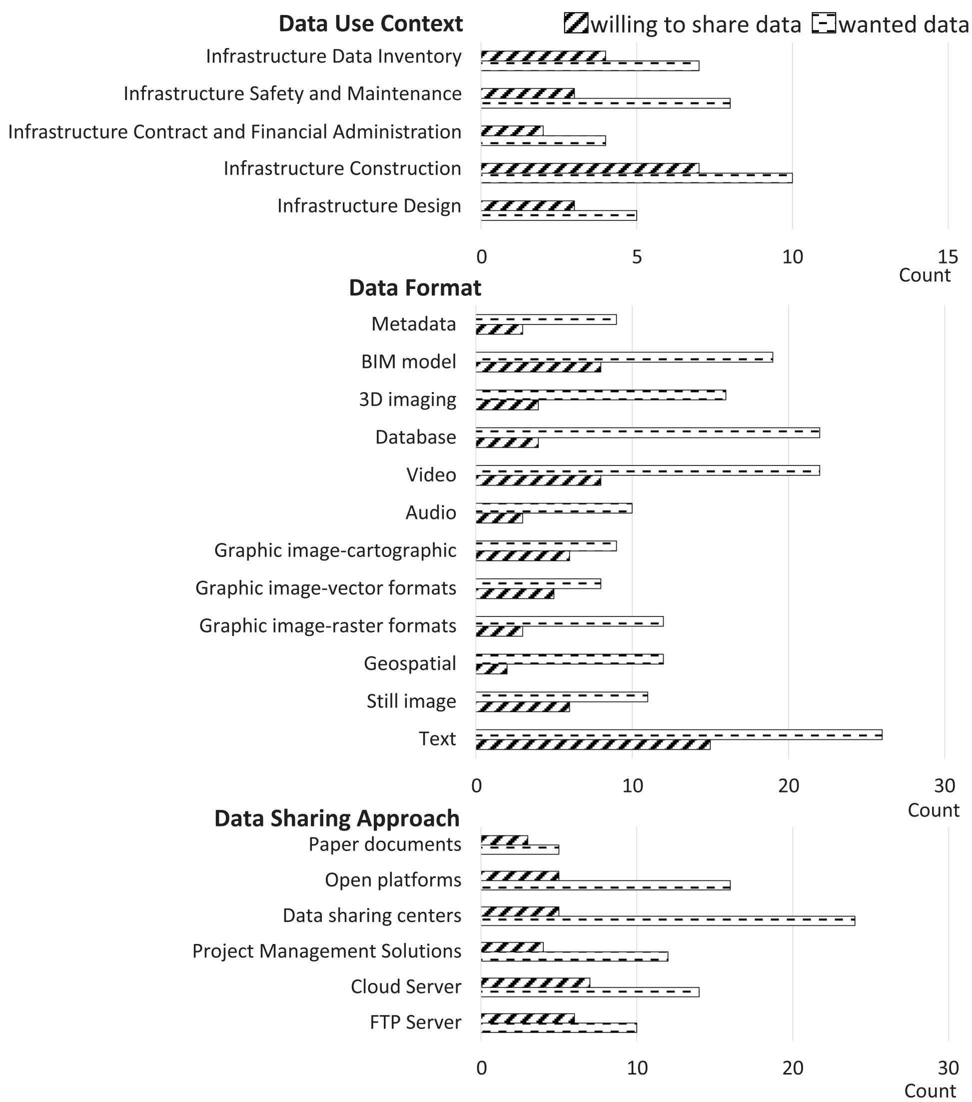


**Fig. 7.** Number of data-sharing contexts, formats, and approaches in responses from data providers and data receivers wished for in future data sharing.

**Table 12.** Top five challenging barriers corresponding to different data characteristics

| Data characteristics | | Top 5 barrier IDs (average challenge level[a]) | | | | | Cronbach's alpha |
|---|---|---|---|---|---|---|---|
| Level 1 | Level 2 | 1st | 2nd | 3rd | 4th | 5th | |
| Data content | Infrastructure data inventory | $SB_6$ (5.33) | $SB_5$ (4.67) | $SB_{10}$ (4.50) | $FB_{20}$ (4.50) | $SB_8$ (4.42) | 0.86 |
| | Infrastructure construction | $SB_7$ (5.10) | $SB_{10}$ (5.10) | $FB_{17}$ (4.95) | $SB_9$ (4.70) | $SB_8/SB_{11}/FB_{20}$ (4.60) | 0.86 |
| | Infrastructure safety and maintenance | $SB_8$ (5.22) | $FB_{19}$ (5.11) | $FB_{20}$ (5.11) | $FB_{17}$ (4.94) | $SB_7$ (4.83) | 0.77 |
| Data format | Database/metadata | $FB_{19}$ (5.10) | $SB_5$ (5.00) | $SB_9$ (4.73) | $HB_{16}$ (4.73) | $FB_{17}$ (4.73) | 0.82 |
| | Image-based[b] | $FB_{19}$ (5.20) | $SB_5$ (4.97) | $SB_3$ (4.65) | $DB_{14}$ (4.61) | $SB_8/SB_9$ (4.61) | 0.78 |
| | BIM model | $SB_5$ (5.20) | $SB_9$ (5.00) | $FB_{19}$ (5.00) | $HB_{15}$ (4.90) | $DB_{14}$ (4.80) | 0.74 |
| | Text | $SB_5$ (5.11) | $FB_{19}$ (5.11) | $SB_9$ (4.78) | $FB_{17}$ (4.78) | $SB_8$ (4.72) | 0.71 |

Note: $AB_1$ = data-sharing awareness; $SB_2$–$SB_{11}$ = social and organization factors; $DB_{12}$–$DB_{14}$ = data and software technical issues; $HB_{15}$–$HB_{16}$ = hardware and cyberinfrastructure issues; $FB_{17}$–$FB_{20}$ = finance and workforce issues; $SB_5$ = data could be misinterpreted; $FB_{19}$ = engineering/technical support for data sharing is limited; and $FB_{20}$ = financial and human resources with a computational mindset for data sharing is limited.
[a]Challenge level: 1 = not; 2 = rarely; 3 = slightly; 4 = somewhat; 5 = moderately; and 6 = definitely.
[b]Image-based: 3D imaging/graphic image-raster formats/graphic image-vector formats/still image/video/geospatial shapefiles.

on infrastructure contract and financial administration is less compared with other types of data.

- Although data-sharing activities based on video and BIM models are few in the current practice (Table 11), these two data types have high demands from receivers. In contrast, data providers would prefer sharing data in text format.
- Although paper-based documents are popular in the current practice of data sharing (Table 11), both data receivers and data providers have less interest in sharing data through paper-based documents in the future.
- The most significant gaps between data receivers and providers in the data-use context and data format and data-sharing approaches are infrastructure safety and maintenance data, database, and the use of data-sharing centers, respectively. Whereas the data receivers are more interested in data-sharing centers, the data provider leans toward using a cloud server to provide the data.

### *Data-Sharing Barriers in the Domain of Civil Infrastructure Engineering*

Different data content and formats may raise the different challenging levels of data-sharing barriers. Understanding such differences helps the data-sharing practitioners allocate resources to addressing corresponding barriers. This section explores such differences based on the challenging level rating results from responders who share the corresponding data content or formats in the current practice. The consistency level of the rating results was measured by the criteria of Cronbach's alpha (Cronbach 1951). If the value of Cronbach's alpha is larger than 0.7 (Barbera et al. 2020), the consistency of the corresponding rating results is reliable.

There were 27 responses collected in the Data-Sharing Barriers survey section. Table 12 listed the top five different data content and format barriers with Cronbach's alpha larger than 0.7 (corresponding to the barriers listed in Table 7). The results indicate that (1) most of the data-sharing barriers are in the Social and Organizational Factors category, followed by the Finance and Workforce Issues; (2) the barrier consistently listed in the top five most challenging barriers for different data content is "business support (e.g., financial and human resources) for data sharing is limited ($FB_{20}$)," which implies the business support is vital concerning data content; and (3) the barriers constantly among the top three most challenging barriers for data format are "engineering/technical support for data sharing is limited ($FB_{19}$)" and "the data could be misinterpreted ($SB_5$)."

Understanding the different opinions about data-sharing barriers from the data provider and receiver is critical to promoting data-sharing activities. Revealing these differences could help data providers and receivers understand the concerns from the other side. Such understanding would also help explain different practitioners' current challenges in data-sharing practice.

As shown in Fig. 8, data receivers considered that the challenging level of "there is a lack of awareness of existing platforms/avenues for sharing or requesting data ($AB_1$)" is relatively higher than the data providers did. The revealed differences between the opinions of data requesters and data providers can help practitioners allocate support in data sharing more efficiently. Data providers and data receivers had significantly different views on the challenge level of "the data could be misinterpreted ($SB_5$)" and "others would falsify my work ($SB_4$)." Data providers had more concerns about these barriers than the data receivers did.

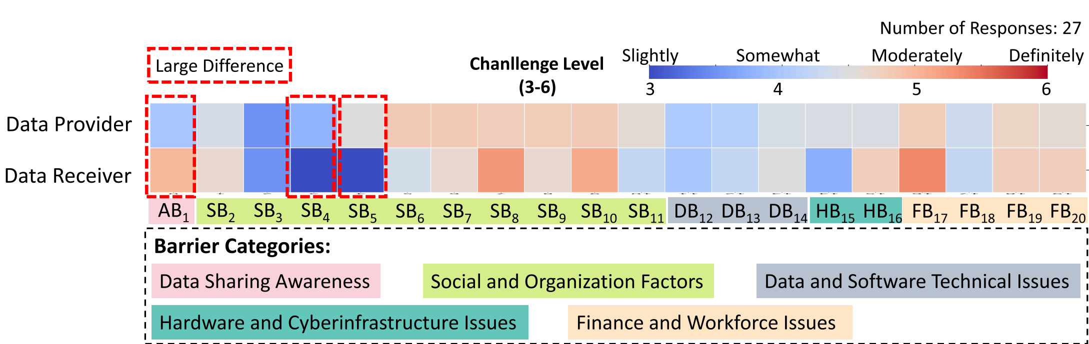


**Fig. 8.** Challenge level of each data-sharing barrier perceived by data providers and data receivers (corresponding to data-sharing barriers listed in Table 7).

An explicit data-sharing protocol such as a data-use agreement could help address the problems of the data providers by clarifying data uses and responsibilities because both sides could negotiate their concerns and achieve a consistent agreement to address these concerns.

Fig. 9 groups the responses about data-sharing barriers from different types of organizations. Labels 1–20 on the $x$-axis correspond to the numbering of data-sharing barriers listed in Table 7. The $y$-axis represents the average challenge level of each barrier as rated by survey participants. The result shows that (1) most of the listed barriers in the survey are above the slightly challenging level; and (2) different organizations have different opinions on the challenge levels of data-sharing barriers. Their most profound concern for respondents from academic institutions was that "others would use my data to publish before me ($SB_2$)." For respondents from the government agencies, their most profound concern was that "data are sensitive ($SB_6$)". For respondents from for-profit corporations and nonprofit organizations, they agreed that "data are private ($SB_7$)," "data are of competitive edges ($SB_8$)," and "ethical and legal issues causing concerns about data ownership, legal implications, and intellectual property (both for individuals and organizations) ($SB_{10}$)" are the barriers with the highest challenge level.

## Data-Sharing Promotion Strategies in the Domain of Civil Infrastructure Engineering

This section examines the survey results of data-sharing promotion strategies. The survey results of data-sharing promotion strategies aim to provide insights on the priority of allocating resources in promoting data sharing. Respondents rated the importance level of each strategy under corresponding categories, i.e., technology, policy, and business. Even though we provided six levels for the responders to rate the importance level of each strategy, including not important at all, low important, slightly important, moderately important, very important, and extremely important, no listed strategy received an important level less than moderately important. Such rating results reflect the importance of data-sharing promotion strategies summarized in this research.

Fig. 10 highlights the data-sharing promotion strategies that receive the top five highest importance level ratings using dash-line rectangles. The circle in Fig. 10 highlights the strategy that received

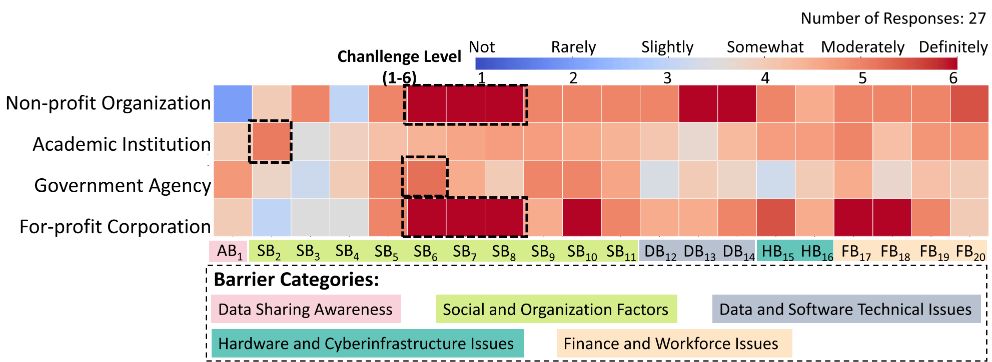


**Fig. 9.** Challenge level of each data-sharing barrier perceived by different organizations (corresponding to data-sharing barriers listed in Table 7).

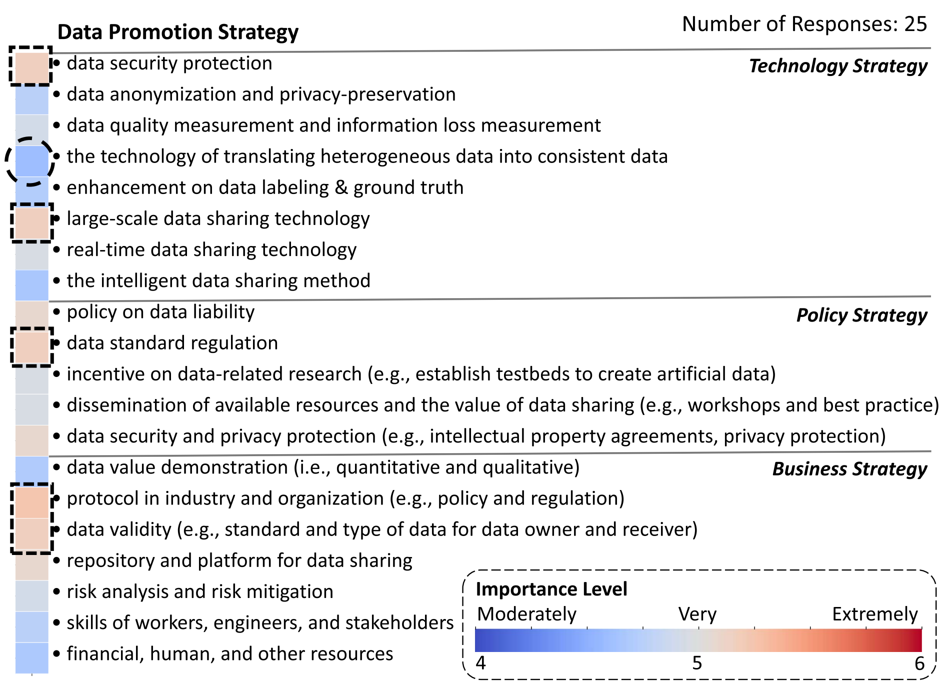


**Fig. 10.** Average importance level of each strategy in promoting data sharing (corresponding to strategies listed in Table 8).

the lowest rating (least important). Among all data-sharing promotion strategies, "protocol in industry and organization" received the highest importance level rating. "Data security protection," "large-scale data-sharing technology," "data standard regulation," and "data validity" received the secondary important level rating. "The technology of translating heterogeneous data into consistent data" received the minimum importance level rating. Overall, the average importance rating on the policy strategies was significantly higher than the average on technology and business strategies.

The survey results indicate that standards and protocols in data sharing are the most important promotion strategies. The related strategies listed under the policy and business categories both receive relatively high importance level ratings. In contrast, some advanced technologies, such as the translating technology of heterogeneous data, data anonymization and privacy preservation, intelligent data-sharing method, and enhancement of data labeling and ground truth, were less important. However, they received significant attention from engineering researchers.

## Roadmap of Data Sharing

Based on the survey results, this section first discusses potential reasons underlying the priority queue of data-sharing strategies under each category (i.e., technology, policy, and business). These analyses provide several recommendations that form a roadmap for systematically addressing the data-sharing barriers identified in this study. Primarily, this section focuses on the top five strategies that received the highest rating results and the strategies that receive the lowest rating results under the corresponding category. Comprehending these priority queues and their reasons is critical for promoting data sharing with targeted resource allocations.

### *Priorities in the Data-Sharing Promotion Strategies*

#### Technology Strategies

Among the series of data-sharing promotion strategies, "data security protection" and "large-scale data-sharing technology" belong to the top five data-sharing promotion strategies. "Data security protection" received a high importance rating because sharing data between different organizations may bring risks to the data providers. As one of the responders who works in the national laboratory annotated: "How to share data with the industry partners, researchers, governmental regulation agencies without sharing vulnerabilities? That vulnerability would be the organization's vulnerabilities or the vulnerability of critical civil infrastructures for public safety."

Meanwhile, with the emergence of monitoring and tracking technologies and ever-advancing data storage power, the data generated from civil infrastructure will grow exponentially. "Large-scale data-sharing technology" has become more critical in the current practice. However, the current practice has not yet fully used large-scale data-sharing techniques or defined what technology can satisfy the practical needs of working with large data sets. As one of the responders annotated: "working with large data is a bit cumbersome. Moreover, there is no technology tool to process it."

In contrast, "the technology of translating heterogeneous data into consistent data" had the lowest importance rating among all technical strategies. It implies that heterogeneous data could be an unusual phenomenon in current practice. People with the same working content in different organizations may frequently use data of similar formats or contents. Possibly only a few people need to handle multiple data sources of different formats for some unique situations under the current practice. However, with the growing diversity of data in civil engineering, this strategy may become more critical in the future.

#### Policy Strategies

Among the policy strategies for promoting data sharing, "data standard regulation" belongs to the top five important data-sharing promotion strategies. The possible reasons for this result could be that existing policies may not fully meet the needs of data-sharing activities in practice. Existing data standards focus more on structured digital data, such as industry foundation classes (IFC) files. However, many data generated in the civil infrastructure industry are unstructured, which means the semantics of the data are missing (Sun et al. 2015). To address this issue, some experts proposed ontology to organize civil infrastructure data (Zhou and El-Gohary 2016). Such ontologies may help the government or industry community to start data standard regulation and define the data liability on these numerous unstructured data.

In contrast, "incentive on data-related research" had the lowest rating among all policy strategies. The reason is that responders from nonacademic institutions gave it a relatively low importance level rating. Because current data-sharing activities between academic and nonacademic institutions are limited, the value of data-related research is not apparent to nonacademic institutions. In the future, by increasing data-sharing promotion activities, the collaboration between academic and nonacademic institutions could reduce such disagreements.

#### Business Strategies

Among the business strategies for promoting data sharing, "the protocol used in industry and organization" and "data validity" belonged to the top five important data-sharing promotion strategies. These two strategies, along with "data standard regulation," focused on the administration issues of data sharing. Also, the high importance levels perceived for these two strategies are consistent with the perceived high importance levels on the policy strategies regarding "data standard regulation." These results reflect the urgency of clear administration-related guidelines in sharing civil infrastructure data.

"Data value demonstration" had the lowest importance level among all business strategies. The reason could be that organizations participating in data sharing could see data value and measure it by themselves. The demonstration of the data value could hardly change the opinion of the organizations from self-judgment and self-explanation activities. Thus, other business strategies seem more critical than "data value demonstration."

### *Recommendations on Data-Sharing Promotion for a Roadmap of Implementation*

Based on the analysis of the priorities in data-sharing strategies under each category, "protocol in industry and organization," "data security protection," "large-scale data-sharing technology," "data standard regulation," and "data validity" are the most critical strategies. Fig. 11 provides recommendations on data-sharing promotion, focusing on the relationship between these strategies and providing several assisting tools for implementing each strategy. Assisting tools can be technical, policy, or business actions that could help implement the strategies. These recommendations form a roadmap for promoting data sharing in civil infrastructure engineering.

Fig. 11 uses boxes to represent business strategies, policy strategies, and technology strategies. Fig. 11 lists each strategy's possible challenges and assisting tools in the horizontal direction. There is a lack of customizable data-use agreements for adapting

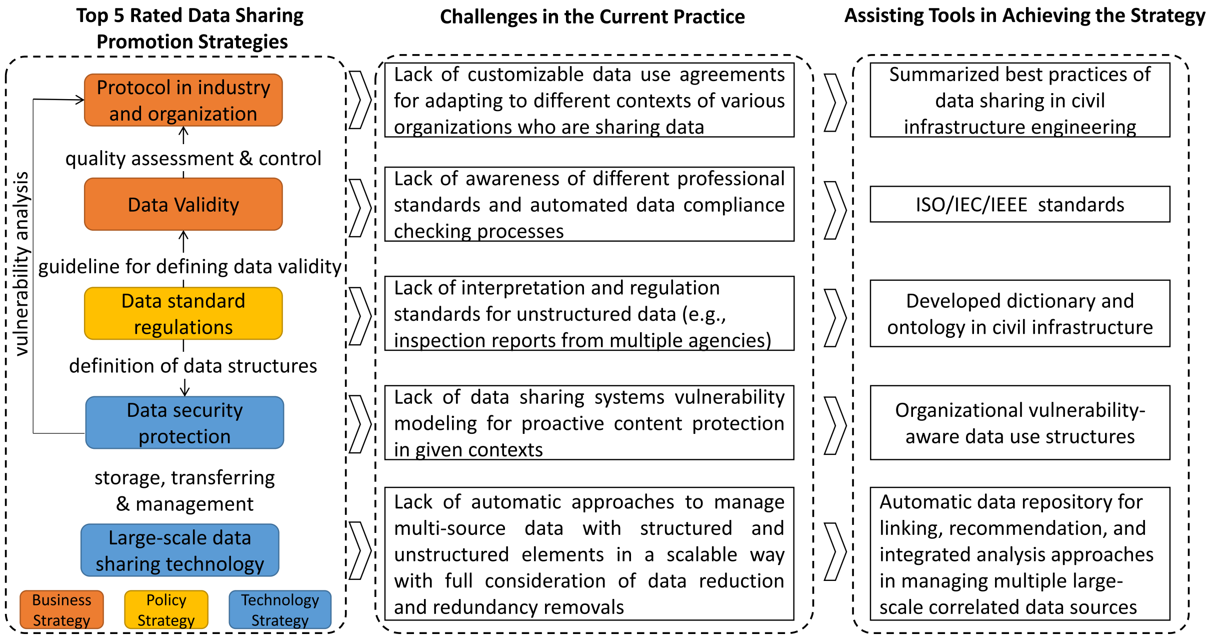


**Fig. 11.** Recommended roadmap for promoting data sharing in civil infrastructure engineering.

to different contexts of various organizations sharing data. To achieve the business strategy of improving the protocol in industry and organization, a summary of the best data-sharing practices in civil infrastructure engineering may help develop a standard data-sharing protocol between organizations under various contexts with different data contents and formats. In the future, the summarized best practices can support the development of an automatic system for generating the data using protocol templates based on the data-sharing organizations' types, context, content, and format of the shared data. These templates will enhance the protocol and negotiation processes with much higher efficiency, comprehensiveness, and lower risks for all organizations engaged in data sharing.

The strategy of data validity, which means the standard and type of data for data owner and receiver, could provide quality assessment and control (QA/QC) to support the protocol of data sharing between different organizations. The current challenge of data validity in the business domain is a lack of awareness of different professional standards and automated data compliance checking processes. Although the existing standards from ISO, International Electrotechnical Commission (IEC), and IEEE may help, such as the several listed in Table 6, the improvements in the policy strategy of data standard regulations would provide more specific guidelines for defining data validation processes.

For example, as mentioned by Whittle et al. (2011), few standards exist for condition assessment data collection with buried pipes, and none are sufficiently comprehensive to enable the development of performance-based design or verifiable implementation of artificial intelligence–based prediction models. With the development and adoption of such standards to ensure data integrity, data collected by various operators, working for various companies, using various types of equipment and software at different times, and even across multiple utilities can be scientifically analyzed. Meanwhile, the developed dictionary and ontology in civil infrastructure could support the community in designing the standard regulations in unstructured data. The basic idea is that the developed ontologies could help the community organize different data types (Pundt and Bishr 2002). Thus, the community could design and implement the data standard regulations over multiple data types.

Meanwhile, a comprehensive data standard regulation policy could support technological improvements in data security protection. Data standard regulation can help establish a consistent data structure among different organizations. Data sharing based on such a consistent data structure would be easier to keep the data integrity, thereby reducing data leaks that bring cybersecurity issues (Yu and Wen 2010). The improvements in cybersecurity protection could also help the community to recognize the vulnerability in data management and thus improve the protocol in industry and organizations (Alhazmi et al. 2005).

At last, large-scale data-sharing technology provides the base of data storage, transferring, and management in the era of big data (Agarwal and Dhar 2014). Some advanced data-sharing repository technologies can assist in fulfilling large-scale data sharing. For example, Liu et al. (2012) proposed a platform that can help integrate heterogeneous data models of the building and HVAC systems and the dynamic data from embedded sensors and controllers. Wu et al. (2022) proposed a method that relieves the workload of extending the data repository with new data. Large-scale data-sharing technology is fundamental in promoting data sharing in civil infrastructure engineering, as shown in Fig. 11.

## Discussion

Data characteristics and data-sharing modes influence data-sharing activities. This section first discusses the influence of data characteristics in the infrastructure domain with a comparison of data characteristics in the other civil engineering domain. Then, this section discusses the differences between sharing data with limited organizations and sharing data with the public, and discusses the potential motivations to encourage sharing data with the public.

The differences in data content and data-use context influence the opinions of data sharing in different domains. For example, data anonymization and privacy-preservation technology seem not to be the top priority in data-sharing promotion strategies related to public service–providing infrastructure facilities. Still, privacy shows some importance underlying some results that connect with privacy without explicitly mentioning privacy. For example, technology on data security protection has been rated as a preferred data-sharing promotion strategy, and privacy has a strong connection with data security protection. Indeed, privacy preservation is the most needed technology in other civil engineering domains, such as building energy (Luo et al. 2022; Janghyun et al. 2022).

The potential reason behind this less explicit statement of privacy issues in civil infrastructure engineering projects could be that the data content and use context in civil infrastructure engineering are less private than in building energy data. Building energy data may reveal the identity of living individuals because people are working or living in the building (Sangogboye et al. 2018). In contrast, data in civil infrastructure engineering (e.g., infrastructure data inventory) are less likely to be linked with a specific person. According to the definition from the European Commission, data containing information that relates to an identified or identifiable living individual is personal data (European Commission 2016). Personal data will have higher levels of requirements of privacy concerns. Thus, priority levels of different promotion strategies varied in various data-sharing activities that happen in different domains.

The types and numbers of entities involved in data-sharing activities impact the value and challenging levels of data sharing. Data sharing between various organizations and individuals can promote the value in the data more than the data-sharing between a few limited organizations. Open data sharing can extensively enable board organizations or individuals to produce innovations. Whereas data sharing with limited organizations could use nondisclosure agreements to precisely control the background of the data requesters and the intentions for the data-sharing request, the open data-sharing mode has less control over the liabilities of the data requesters. In contrast, in the current open data-sharing mode, the data providers use licenses with exemptions for protecting themselves, such as the Creative Commons (CC0) license (Creative Commons 2017).

The motivation of data providers to share data publicly is limited in the current practice. The survey results in the Future Vision part showed that the data providers would be willing to share data if there is a cost-sharing in the data sharing. Although business support for data sharing can promote open data, credit to the data providers may also incentivize data sharing. For example, Nature Portfolio established a peer-reviewed journal, *Scientific Data* (Scientific Data 2014), which publishes data sets with specific descriptions of the collecting process and data contents in the data sets. Such open-access journal publications can incentivize data providers through the potential citation crediting.

## Limitations and Relevant Ongoing and Future Work

Given various civil infrastructure engineers' working backgrounds, we collected different civil infrastructure engineering stakeholders' current practice in data sharing and their opinions about how they view data-sharing barriers, future vision, and promotion strategies. Based on the synthesized list of data-sharing-related factors, the research team conducted a half-year-long online survey. One major limitation of this study is that most of the survey participants are from the construction engineering and management domain, despite having some responses from other areas (transportation and bridge engineering). Unavoidably, the survey results reflect the opinions majorly from academic construction areas. However, insights obtained from the survey results about data-sharing protocols and data security issues have driven the authors to initiate two ASCE tasks that could help the authors reach out to diverse groups of researchers and industrial professionals in the future. These two ASCE tasks are (1) a white paper on the protocol of data sharing in civil engineering; and (2) a virtual symposium on data sharing and cybersecurity in civil engineering. Using these two initiatives, the authors could reach out to diverse stakeholders in the civil infrastructure domain in the future with more specific data-sharing-related tasks.

## Conclusions

This section summarizes the contributions of this research and the main findings of (1) the current practice; (2) future vision and barriers of data sharing in civil infrastructure engineering; and (3) strategies for addressing data-sharing barriers, respectively. These findings can guide the development of an implementation plan that integrates multiple data-sharing promotion strategies into actional solutions for encouraging data sharing in civil infrastructure engineering, as presented in Fig. 11.

The research presented in this paper contributes the following:

- We summarized characteristics of data in civil infrastructure engineering from three aspects: data-use context, data format, and data-sharing approaches, thereby providing a systematical list for reviewing the current practice of data sharing in civil infrastructure engineering in handling different data contents and formats with various contexts.
- We summarized lists of conditions expressed by data providers for approaching a vision of better data sharing between various organizations and synthesized multiple stakeholders' needs and willingness to share data in the domain of civil infrastructure engineering.
- We classified the barriers identified into five categories and explored how specific barriers influence various data-sharing activities in different contexts.
- We synthesized the available data-sharing promotion strategies and organized these strategies into three groups of solutions, i.e., technology-, policy-, and business-related solution groups, and prioritized different promotion strategies in typical infrastructure engineering contexts.
- We analyzed data-sharing promotion strategies and formed a roadmap based on the analysis that can guide the development of a healthy data-sharing system in civil and infrastructure.

### *Current Practice of Data Sharing in the Domain of Civil Infrastructure Engineering*

Currently, data-sharing activities happen mainly between the same type of organizations. Compared with academic institutions, for-profit organizations, and nonprofit organizations, government agencies are the most active organization in data-sharing activities with the most diverse shared data context. Infrastructure construction data are the most frequently shared data context between organizations. Infrastructure contracts and financial administration data are rarely shared between different organizations. BIM models are rarely shared between various organizations in the current data-sharing activities. Most shared data are in text format, followed by still images. Meanwhile, paper-based documents are the most frequent data-sharing approach.

### *Future Vision and Barriers to Data Sharing in the Domain of Civil Infrastructure Engineering*

Infrastructure construction data are the most wanted data from data receivers. Data providers are mainly willing to provide infrastructure construction data as well. Data-sharing activities based on video and BIM models have high demand from data receivers. Although data providers have no apparent preference for data-sharing approaches, data receivers prefer to use data-sharing centers and open platforms to receive data. "If there is a cost-sharing in the data sharing" is the condition that would make data providers more willing to share data. The civil industry could potentially see more data-sharing activities if this condition is met.

Stakeholders of different roles may face different barriers and prefer different approaches to promote data sharing. Academic institutions are more concerned that "others would use my data to publish before me." Government agencies view "data are sensitive" as the most formidable data-sharing barrier. For-profit corporations rated "data are private," "data are of competitive edges," and "ethical and legal issues causing concerns about data ownership, legal implications, and intellectual property (for both individuals and organizations)" at the highest challenge level.

### *Priority Queue of Data-Sharing Promotion Strategies in the Civil Infrastructure Engineering*

There are numerous strategies for promoting data sharing. However, the promotion strategies with high priority deserve more attention and resources. This research explored the priority of data-sharing promotion strategies and synthesized a roadmap for implementing the priority strategies in the domain of civil infrastructure engineering. Meanwhile, limited research discussed the policy and business strategies to promote data sharing in civil engineering; this research provides a synthesized list of policy and business strategies.

This research identified that "protocol in industry and organization," "data-standard regulation," "data validity," "data security protection," and "large-scale data-sharing technology" are the top five priorities in data sharing of civil infrastructure engineering. This research further developed a roadmap for integrating these priorities and relevant assisting tools to implement a better data-sharing system in civil infrastructure engineering.

## Data Availability Statement

All data that support the findings of this study are available from the corresponding author upon reasonable request.

## Acknowledgments

The research was sponsored by the American Society of Civil Engineers (ASCE). The support is gratefully acknowledged.

## References

AbouRizk, S., D. Halpin, Y. Mohamed, and U. Hermann. 2011. "Research in modeling and simulation for improving construction engineering operations." *J. Constr. Eng. Manage.* 137 (10): 843–852. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000288.

Agarwal, R., and V. Dhar. 2014. "Big data, data science, and analytics: The opportunity and challenge for IS research." *Inf. Syst. Res.* 25 (3): 443–448. https://doi.org/10.1287/isre.2014.0546.

AKDOT (Alaska Department of Transportation). 2019. "Road weather information system." Accessed August 17, 2019. https://www.roadweather.alaska.gov/iways/roadweather/forms/IndexForm.html.

Alhazmi, O., Y. Malaiya, and I. Ray. 2005. "Security vulnerabilities in software systems: A quantitative perspective." In *Proc., IFIP Annual Conf. on Data and Applications Security and Privacy*, 281–294. New York: Springer.

Antunes, R., V. A. González, K. Walsh, and O. Rojas. 2017. "Dynamics of project-driven production systems in construction: Productivity function." *J. Comput. Civ. Eng.* 31 (5): 04017053. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000703.

ASTM. 2019. "ASTM Committee E57 on 3D imaging systems subcommittee E57.94 on data interoperability." Accessed April 17, 2019. https://compass.astm.org/EDIT/html_annot.cgi?E2807+11(2019)#s00001.

Attard, J., F. Orlandi, S. Scerri, and S. Auer. 2015. "A systematic review of open government data initiatives." *Gov. Inf. Q.* 32 (4): 399–418. https://doi.org/10.1016/j.giq.2015.07.006.

Australian Government. 2019. "Best practice guide to applying data sharing principles." Accessed August 29, 2022. https://www.pmc.gov.au/sites/default/files/publications/data-sharing-principles-best-practice-guide-15-mar-2019.pdf.

AZDOT (Arizona Department of Transportation). 2019. "Data and information." Accessed April 17, 2019. https://azdot.gov/planning/data-and-information.

Azeez, M., J. Gambatese, and S. Hernandez. 2019. "What do construction workers really want? A study about representation, importance, and perception of US construction occupational rewards." *J. Constr. Eng. Manage.* 145 (7): 04019040. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001669.

Balaji, B. 2018. "Zodiac dataset." Accessed December 25, 2018. https://www.synergylabs.org/bharath/datasets.html.

Barbera, J., N. Naibert, R. Komperda, and T. C. Pentecost. 2020. "Clarity on Cronbach's alpha use." *J. Chem. Educ.* 98 (2): 257–258.

Beaulieu-Jones, B. K., Z. S. Wu, C. Williams, R. Lee, S. P. Bhavnani, J. B. Byrd, and C. S. Greene. 2019. "Privacy-preserving generative deep neural networks support clinical data sharing." *Circulation* 12 (7): e005122. https://doi.org/10.1161/CIRCOUTCOMES.118.005122.

Becker, T., E. Curry, A. Jentzsch, and W. Palmetshofer. 2016. *New horizons for a data-driven economy: Roadmaps and action plans for technology, businesses, policy, and society*, 277–291. New York: Springer. https://doi.org/10.1007/978-3-319-21569-3.

Benoît, J., J. I. Bobbitt, D. J. Ponti, and S. A. Shimel. 2004. "Data dictionary and formatting standard for dissemination of geotechnical data." In *Geotechnical engineering for transportation projects*, 829–836. Reston, VA: ASCE.

BOTS (Bureau of Transportation Statistics). 2015. "National transportation atlas database." Accessed April 17, 2019. https://www.bts.gov/archive/publications/national_transportation_atlas_database/index.

Bradley, A., H. Li, R. Lark, and S. Dunn. 2016. "BIM for infrastructure: An overall review and constructor perspective." *Autom. Constr.* 71 (Mar): 139–152. https://doi.org/10.1016/j.autcon.2016.08.019.

BuildingSmart. 2015. "buildingSMART data dictionary." Accessed April 17, 2019. https://www.buildingsmart.org/users/services/buildingsmart-data-dictionary/.

BuildingSmart. 2019. "Industry foundation classes." Accessed April 17, 2019. https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/.

Burrows, T., N. Findlay, C. Killen, S. E. Dempsey, S. Hunter, P. Chiarelli, and S. Snodgrass. 2011. "Using nominal group technique to develop a consensus derived model for peer review of teaching across a multi-school faculty." *J. Univ. Teach. Learn. Pract.* 8 (2): 8. https://doi.org/10.53761/1.8.2.8.

Carroll, S. R., I. Garba, O. L. Figueroa-Rodríguez, J. Holbrook, R. Lovett, S. Materechera, M. Parsons, K. Raseroka, D. Rodriguez-Lonebear, and R. Rowe. 2020. "The CARE principles for indigenous data governance." *Data Sci. J.* 19 (1): 43. https://doi.org/10.5334/dsj-2020-043.

Chen, A. C. R. Y. 2019. "ISO/TC 59/SC 13 organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM)." Accessed April 17, 2019. https://www.iso.org/committee/49180.html.

Cooper, D., R. Springer, J. G. Benner, D. Bloom, E. Carrillo, A. Carroll, B. Chang, X. Chen, D. Daix, and E. Dommermuth. 2019. *Supporting the changing research practices of civil and environmental engineering scholars*. New York: Ithaka S+R. https://doi.org/10.18665/sr.310885.

Costin, A., and C. Eastman. 2019. "Need for interoperability to enable seamless information exchanges in smart and sustainable urban systems."

*J. Comput. Civ. Eng.* 33 (3): 04019008. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000824.

Costin, A. M., J. Teizer, and B. Schoner. 2015. "RFID and BIM-enabled worker location tracking to support real-time building protocol and data visualization." *J. Inf. Technol. Constr.* 20 (29): 495–517.

Costley, L. J., M. Salvato, C. Maestre, M. Bryner, and S. Ewing. 2021. "Infrastructure architecture framework: A multi-sector approach to enterprise systems engineering and management." Accessed December 23, 2021. https://www.nyas.org/media/24240/nyas-infra-arch-framework-v10-1.pdf.

Creative Commons. 2017. "CC0 1.0 universal (CC0 1.0) public domain dedication." Accessed September 10, 2022. https://creativecommons.org/publicdomain/zero/1.0/.

Cronbach, L. J. 1951. "Coefficient alpha and the internal structure of tests." *Psychometrika* 16 (3): 297–334. https://doi.org/10.1007/BF02310555.

Dikeocha, L. U., B. N. Nwaiwu, C. C. Nwagu, and L. Amaechi. 2018. "Promoting small scale business: Road map for a rapid national development." *Nigerian J. Bus. Educ.* 5 (2): 132–140.

DOE. 2014. "Public access plan." Accessed December 15, 2018. https://www.energy.gov/sites/prod/files/2014/08/f18/DOE_Public_Access%20Plan_FINAL.pdf.

DOE. 2018. "Privacy." Accessed December 15, 2018. https://www.energy.gov/about-us/web-policies/privacy.

Dyke, S. O., and T. J. Hubbard. 2011. "Developing and implementing an institute-wide data sharing policy." *Genome Med.* 3 (9): 60–68. https://doi.org/10.1186/gm276.

El-Gohary, N. M., and T. E. El-Diraby. 2010. "Domain ontology for processes in infrastructure and construction." *J. Constr. Eng. Manage.* 136 (7): 730–744. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000178.

England Highway. 2021. "Asset data management manual v13.0." Accessed December 20, 2021. https://www.standardsforhighways.co.uk/ha/standards/admm/index.htm.

EUROPA. 2016. "EU regulation for data protection." Accessed April 17, 2019. https://eur-lex.europa.eu/eli/reg/2016/679/2016-05-04.

European Commission. 2016, "General data protection regulation." Accessed September 8, 2022. https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en#answer.

Famakin, I. O., and A. O. Abisuga. 2016. "Effect of path-goal leadership styles on the commitment of employees on construction projects." *Int. J. Constr. Manage.* 16 (1): 67–76. https://doi.org/10.1080/15623599.2015.1130601.

Guo, F., C. T. Jahren, Y. Turkan, and H. David Jeong. 2017. "Civil integrated management: An emerging paradigm for civil infrastructure project delivery and management." *J. Manage. Eng.* 33 (2): 04016044. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000491.

Halfawy, M. R. 2008. "Integration of municipal infrastructure asset management processes: Challenges and solutions." *J. Comput. Civ. Eng.* 22 (3): 216–229. https://doi.org/10.1061/(ASCE)0887-3801(2008)22:3(216).

Halfawy, M. R. 2010. "Municipal information models and federated software architecture for implementing integrated infrastructure management environments." *Autom. Constr.* 19 (4): 433–446. https://doi.org/10.1016/j.autcon.2009.11.013.

ILDOT (Illinois Department of Transportation). 2019. "Transportation system." Accessed April 19, 2019. https://idot.illinois.gov/transportation-system/index.

Irizarry, J., and E. P. Karan. 2012. "Optimizing location of tower cranes on construction sites through GIS and BIM integration." *J. Inf. Technol. Constr.* 17 (23): 351–366.

Irizarry, J., E. P. Karan, and F. Jalaei. 2013. "Integrating BIM and GIS to improve the visual monitoring of construction supply chain management." *Autom. Constr.* 31 (Jun): 241–254. https://doi.org/10.1016/j.autcon.2012.12.005.

ISO. 2003. "Industrial automation systems and integration—Integration of life-cycle data for process plants including oil and gas production facilities." Accessed March 23, 2021. https://www.iso.org/standard/29556.html.

ISO. 2007. "Building construction: Organization of information about construction works, Part 3: Framework for object-oriented information." Accessed March 23, 2021. https://www.iso.org/standard/38706.html.

ISO. 2011. "Data quality—Part 1: Overview." Accessed March 23, 2021. https://www.iso.org/standard/50798.html.

ISO. 2013. "Industry foundation classes (IFC) for data sharing in the construction and facility management industries." Accessed March 23, 2021. https://www.iso.org/standard/70303.html.

ISO. 2015a. "Information technology—Security techniques—Information security management for inter-sector and inter-organizational communications." Accessed March 23, 2021. https://www.iso.org/standard/68427.html.

ISO. 2015b. "Systems and software engineering—System life cycle processes." Accessed March 23, 2021. https://www.iso.org/standard/63711.html.

ISO. 2017. "Information and documentation—Data exchange protocol for interoperability and preservation." Accessed March 23, 2021. https://www.iso.org/standard/68562.html.

ISO. 2020a. "Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM)—Information management using building information modelling—Part 5: Security-minded approach to information management." Accessed March 23, 2021. https://www.iso.org/standard/74206.html.

ISO. 2020b. "Smart community infrastructures—Guidelines on data exchange and sharing for smart community infrastructures." Accessed March 23, 2021. https://www.iso.org/standard/69242.html.

ISO. 2022. "Smart community infrastructures—Data exchange and sharing for community infrastructures based on geographic information." Accessed December 23, 2022. https://www.iso.org/standard/81292.html.

Janghyun, K., H. Barry, and H. Tianzhen. 2022. "A review of preserving privacy in data collected from buildings with differential privacy." *J. Build. Eng.* 5 (2): 104724. https://doi.org/10.1016/j.jobe.2022.104724.

Karaguzel, O. T., M. Elshambakey, Y. Zhu, T. Hong, W. J. Tolone, S. Das Bhattacharjee, I. Cho, W. Dou, H. Wang, and S. Lu. 2019. "Open computing infrastructure for sharing data analytics to support building energy simulations." *J. Comput. Civ. Eng.* 33 (6): 04019037. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000857.

Le, T., and H. David Jeong. 2017. "NLP-based approach to semantic classification of heterogeneous transportation asset data terminology." *J. Comput. Civ. Eng.* 31 (6): 04017057. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000701.

Lee, H.-K., C.-M. Liu, and M.-H. Lee. 2006. "Aligning the business roadmap to technology roadmap for managing the new product development project." *J. Chin. Inst. Ind. Eng.* 23 (6): 449–457. https://doi.org/10.1080/10170660609509341.

Liu, K., and N. El-Gohary. 2017. "Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports." *Autom. Constr.* 81 (Jun): 313–327. https://doi.org/10.1016/j.autcon.2017.02.003.

Liu, X., B. Akinci, M. Berges, and J. H. Garrett Jr. 2012. "An integrated performance analysis framework for HVAC systems using heterogeneous data models and building automation systems." In *Proc., 4th ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, 145–152. New York: Association for Computing Machinery. https://doi.org/10.1145/2422531.2422558.

Luo, N., Z. Wang, D. Blum, C. Weyandt, N. Bourassa, M. A. Piette, and T. Hong. 2022. "A three-year dataset supporting research on building energy management and occupancy analytics." *Sci. Data* 9 (1): 1–15. https://doi.org/10.1038/s41597-022-01257-x.

MAG (Maricopa Association of Governments). 2019. "Maricopa association of governments open GIS." Accessed April 17, 2019. https://geodata-azmag.opendata.arcgis.com/.

Manyika, J., M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Hung Byers. 2011. *Big data: The next frontier for innovation, competition, and productivity.* Washington, DC: McKinsey Global Institute.

Müller, O., M. Fay, and J. Vom Brocke. 2018. "The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics." *J. Manage. Inf. Syst.* 35 (2): 488–509. https://doi.org/10.1080/07421222.2018.1451955.

Mutis, I., and I. Mehraj. 2022. "Cloud BIM governance framework for implementation in construction firms." *Pract. Period. Struct. Des. Constr.* 27 (1): 16. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000652.

Ng, S. T., F. J. Xu, Y. Yang, and M. Lu. 2017. "A master data management solution to unlock the value of big infrastructure data for smart, sustainable and resilient city planning." *Procedia Eng.* 196 (Aug): 939–947. https://doi.org/10.1016/j.proeng.2017.08.034.

Nguyen, V. M., J. L. Brooks, N. Young, R. J. Lennox, N. Haddaway, F. G. Whoriskey, R. Harcourt, and S. J. Cooke. 2017. "To share or not to share in the emerging era of big data: Perspectives from fish telemetry researchers on data sharing." *Can. J. Fish. Aquat. Sci.* 74 (8): 1260–1274. https://doi.org/10.1139/cjfas-2016-0261.

Noy, N. F., and C. D. Hafner. 1997. "The state of the art in ontology design: A survey and comparative review." *AI Mag.* 18 (3): 53–74. https://doi.org/10.1609/aimag.v18i3.1306.

Noy, N. F., and D. L. McGuinness. 2001. *Ontology development 101: A guide to creating your first ontology.* Stanford knowledge systems laboratory Technical Rep. No. KSL-01-05. Redwood, CA: Stanford Univ.

NTAD (National Transportation Atlas Databases). 2011. *National transportation atlas databases: 2011*. Washington, DC: USDOT.

ODOT (Oregon Department of Transportation). 2019. "ODOT data & maps." Accessed April 17, 2019. https://www.oregon.gov/odot/Data/Pages/index.aspx.

OGC (Open Geospatial Consortium). 2019. "OGC standards and resources: Made by our member community for use by the global community." Accessed April 17, 2019. https://www.opengeospatial.org/standards.

Ogunlana, S. O., H. Li, and F. A. Sukhera. 2003. "System dynamics approach to exploring performance enhancement in a construction organization." *J. Constr. Eng. Manage.* 129 (5): 528–536. https://doi.org/10.1061/(ASCE)0733-9364(2003)129:5(528).

Olszak, C. M. 2016. "Toward better understanding and use of Business Intelligence in organizations." *Inf. Syst. Manage.* 33 (2): 105–123. https://doi.org/10.1080/10580530.2016.1155946.

Pesek, A. E., J. B. Smithwick, A. Saseendran, and K. T. Sullivan. 2019. "Information asymmetry on heavy civil projects: Deficiency identification by contractors and owners." *J. Manage. Eng.* 35 (4): 04019008. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000694.

Petri, I., T. Beach, O. F. Rana, and Y. Rezgui. 2017. "Coordinating multisite construction projects using federated clouds." *Autom. Constr.* 83 (Jun): 273–284. https://doi.org/10.1016/j.autcon.2017.08.011.

Poldrack, R. A., and K. J. Gorgolewski. 2014. "Making big data open: Data sharing in neuroimaging." *Nat. Neurosci.* 17 (11): 1510–1517. https://doi.org/10.1038/nn.3818.

Pryor, G. 2009. "Multi-scale data sharing in the life sciences: Some lessons for policy makers." *Int. J. Digital Curation* 4 (3): 71–82. https://doi.org/10.2218/ijdc.v4i3.115.

Pundt, H., and Y. Bishr. 2002. "Domain ontologies for data sharing–an example from environmental monitoring using field GIS." *Comput. Geosci.* 28 (1): 95–102. https://doi.org/10.1016/S0098-3004(01)00018-8.

Qualtrics. 2020. "Qualtrics." Accessed May 12, 2018. https://www.qualtrics.com/.

Rathje, E. M., C. Dawson, J. E. Padgett, J.-P. Pinelli, D. Stanzione, A. Adair, P. Arduino, S. J. Brandenberg, T. Cockerill, and C. Dey. 2017. "DesignSafe: New cyberinfrastructure for natural hazards engineering." *Nat. Hazard. Rev.* 18 (3): 06017001. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000246.

Rios, R. S., K. I. Zheng, and M.-H. Zheng. 2020. "Data sharing during COVID-19 pandemic: What to take away." *Expert Rev. Gastroenterol. Hepatol.* 14 (12): 1125–1130. https://doi.org/10.1080/17474124.2020.1815533.

Robinson, P., and R. Shepard. 2011. "Outreach, applied research, and management needs for Wisconsin's great lakes freshwater estuaries: A cooperative extension needs assessment model." *J. Ext.* 49 (1): 13.

Sangogboye, F. C., R. Jia, T. Hong, C. Spanos, and M. B. Kjærgaard. 2018. "A framework for privacy-preserving data publishing with enhanced utility for cyber-physical systems." *ACM Transact. Sensor Network* 14 (3–4): 1–22. https://doi.org/10.1145/3275520.

Schwee, J. H., F. C. Sangogboye, and M. B. Kjærgaard. 2019. "Anonymizing building data for data analytics in cross-organizational settings." In *Proc., Int. Conf. on Internet of Things Design and Implementation*, 1–12. New York: Association for Computing Machinery. https://doi.org/10.1145/3302505.3310064.

Scientific Data. 2014. "Scientific data." Accessed September 12, 2022. https://www.nature.com/sdata/.

Serwadda, D., P. Ndebele, M. K. Grabowski, F. Bajunirwe, and R. K. Wanyenze. 2018. "Open data sharing and the Global South—Who benefits?" *Science* 359 (6376): 642–643. https://doi.org/10.1126/science.aap8395.

Sun, Y., H. Bae, S. Manna, J. White, and M. Golparvar-Fard. 2015. "Bridging semantics with physical objects using augmented reality." In *Proc., 2015 IEEE 9th Int. Conf. on Semantic Computing (IEEE ICSC 2015)*, 344–349. New York: IEEE.

Thu, H. N., and U. Wehn. 2016. "Data sharing in international transboundary contexts: The Vietnamese perspective on data sharing in the Lower Mekong Basin." *J. Hydrol.* 536 (May): 351–364. https://doi.org/10.1016/j.jhydrol.2016.02.035.

USDOTFHA (US Department of Transportation Federal Highway Administration). 2019. "Office of infrastructure." Accessed April 17, 2019. https://www.fhwa.dot.gov/infrastructure/.

Wall, B., H. Jagdev, and J. Browne. 2005. "An approach to developing an eBusiness roadmap." *Prod. Plann. Control* 16 (7): 701–715. https://doi.org/10.1080/09537280500290136.

Wang, H., Y. Pan, and X. Luo. 2019. "Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis." *Autom. Constr.* 103 (Jul): 41–52. https://doi.org/10.1016/j.autcon.2019.03.005.

Wang, H., Z. Xu, H. Fujita, and S. Liu. 2016. "Towards felicitous decision making: An overview on challenges and trends of big data." *Inf. Sci.* 367 (Nov): 747–765. https://doi.org/10.1016/j.ins.2016.07.007.

Wang, J., M. Chen, G. Lü, S. Yue, Y. Wen, Z. Lan, and S. Zhang. 2020. "A data sharing method in the open web environment: Data sharing in hydrology." *J. Hydrol.* 587 (Aug): 124973. https://doi.org/10.1016/j.jhydrol.2020.124973.

Wang, Y. 2019. "Survey design and survey results." Accessed August 18, 2022. https://illinoisaces.co1.qualtrics.com/jfe/form/SV_a9OxZCMKZfLESoJ;https://drive.google.com/file/d/1Nr5a-k_qwxSshxnmcXLKC4wICzCpnMjE/view.

Whittle, L. G., K. Steiner, and S. Sinha. 2011. "The need for national data standards for condition assessment." In *Proc., Pipelines 2011: A Sound Conduit for Sharing Solutions*, 382–394. Reston, VA: ASCE.

Wilkinson, M. D., M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, and P. E. Bourne. 2016. "The FAIR Guiding Principles for scientific data management and stewardship." *Sci. Data* 3 (1): 1–9. https://doi.org/10.1038/sdata.2016.18.

WSDOT (Washington State Department of Transportation). 2019. "Maps & data." Accessed April 17, 2019. https://www.wsdot.wa.gov/mapsdata.htm.

Wu, J., T. Akanbi, and J. Zhang. 2022. "Constructing invariant signatures for AEC objects to support BIM-based analysis automation through object classification." *J. Comput. Civ. Eng.* 36 (4): 04022008. https://doi.org/10.1061/(ASCE)CP.1943-5487.0001012.

Yu, X., and Q. Wen. 2010. "A view about cloud data security from data life cycle." In *Proc., 2010 Int. Conf. on Computational Intelligence and Software Engineering*, 1–4. New York: IEEE.

Zakaib, G. 2019. "Alberta nonprofit data strategy roadmap." Accessed April 17, 2019. https://static1.squarespace.com/static/5aef5b46cef3728571e6c46c/t/5c745650eb393140a620215e/1551128147129/ANDS+Roadmap+-+February+2019.pdf.

Zhang, L., and R. R. Issa. 2013. "Ontology-based partial building information model extraction." *J. Comput. Civ. Eng.* 27 (6): 576–584. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000277.

Zhong, B., H. Li, H. Luo, J. Zhou, W. Fang, and X. Xing. 2020. "Ontology-based semantic modeling of knowledge in construction: Classification and identification of hazards implied in images." *J. Constr. Eng. Manage.* 146 (4): 04020013. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001767.

Zhou, P., and N. El-Gohary. 2016. "Ontology-based multilabel text classification of construction regulatory documents." *J. Comput. Civ. Eng.* 30 (4): 04015058. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000530.

Zuiden, B. V. 2019. "Universal data sharing agreement." Accessed April 17, 2019. https://github.com/Clever/policies/blob/master/universal-data-sharing-agreement.md#universal-data-sharing-agreement.