Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
18/1 (2025), 47-66. DOI: http://dx.doi.org/10.21609/jiki.v18i1.1310

Optimizing Data Quality in Interagency Data Sharing: A Framework

Monica Vivi Kurniawati[1], Mohamad Faisal Zulmy[2], and Yova Ruldeviyani[3]

Faculty of Computer Science, Universitas Indonesia
Kampus UI Salemba, 10430, Indonesia

E-mail: [1]*monica.vivi@ui.ac.id,* [2]*mohamad.faisal31@ui.ac.id*, [3]*yova@cs.ui.ac.id*

Abstract

As governments strive for openness, inclusivity, and collaboration to deliver public value and inform policy with a data-driven approach, interagency data sharing (IDS) becomes increasingly critical. However, challenges remain, particularly regarding data quality. This study addresses this gap by proposing a novel framework for IDS within government agencies. This framework goes beyond traditional approaches by proactively managing data quality throughout its lifecycle. It aims to provide insights for fostering effective IDS practices that support collaborative and evidence-based decision-making, a cornerstone of modern government operations. This research employs a Systematic Literature Review (SLR) to examine current IDS practices and associated challenges. Subsequently, a case study conducted at Indonesia's Directorate General of Taxes utilizes semi-structured in-depth interviews with practitioners to explore practical implementation scenarios. The framework developed through this research aims to mitigate various challenges related to IDS, particularly focusing on enhancing data quality assurance. By ensuring high-quality exchanged data, this framework empowers governments to break down silos, fostering a more coordinated and efficient approach to public service delivery.

Keywords: *data quality, government, data sharing, data exchange, framework*

1. Introduction

Recently, governments are driven by public demand to move away from segmented service and instead focus on fostering collaboration between government agencies to deliver integrated services, with a priority on achieving public value [1]-[6]. This shift towards integrated public services underscores the importance of government entities working together, as well as partnering with non-governmental entities like businesses or communities, to address complex societal issues and improve outcomes for public services as well as resource sharing [2], [6]-[7]. Central to these efforts is the foundation of shared, timely, and actionable interagency data sharing [5], [6], [8], [9].

Interagency Data Sharing (IDS) refers to the collaborative practice of exchanging information and data among agencies. Successful IDS offers several advantages, including enhanced efficiency through reduced data management costs, shared resources and technical knowledge, decreased duplication in data collection and storage, evidence-based policymaking, and better delivery of public services [6], [10]-[18].

Yet, obstacles pose a barrier to the successful implementation of IDS, potentially impeding the achievement of its intended objectives. These challenges, which may arise in technical, organizational, regulatory, and political forms [5], [10], [13], [17], [19], [20], [21], need to be carefully considered and addressed when initiating IDS, as they can influence the willingness of agencies to participate and engage in the program. Additionally, previous research indicates that participants' reluctance to participate in IDS might arise from concerns about data quality [17], [22].

Furthermore, beyond their impact on IDS participation initiatives, issues with data quality also diminish the potential value derived from the data collected through IDS [20], as effective management, standardization, and integration of data are necessary to realize value and foster efficient collaboration [20], [23]-[25]. Poor data quality can lead to ineffective decision-making, hinder evidence-based policy formulation, and compromise the delivery of essential services and resource allocation to citizens [26], [27]. This extends to adversely affecting the well-being of the population, exacerbating educational access disparities, compromising public safety and security, and eroding public trust [28]-[30].

Previous research has produced several frameworks aimed at facilitating data sharing. These frameworks, address various aspects of data sharing, such as organizational activities benefiting from shared data, standardized data collection methods, trust establishment in partnerships, and voluntary information sharing between businesses and government [31]. However, a notable gap exists in these frameworks regarding data quality considerations. Despite recognizing the challenge and impact of poor data quality in IDS, there has been a noticeable lack of effort to address this issue [31]. This study seeks to help to fill that gap. The main questions we aim to explore in this study are:

1. What are the challenges and issues related to data quality in IDS?
2. What are the governance process for addressing these challenge and data quality issue?

To address these questions, we have constructed a conceptual framework for IDS practices within government context, referred to as the Government Interagency Data Sharing Framework (GIDS Framework).

The research proceeds as follows: First, a Systematic Literature Review (SLR) is conducted to explore IDS practices and comprehend the associated challenges and data quality issues. Specific challenges encountered in IDS within government agencies are pinpointed, along with potential solutions to help organizations surmount these obstacles. Additionally, a case study will be undertaken to further enhance the initial framework proposed.

The case study will be carried out at the Directorate General of Taxes, Ministry of Finance Indonesia, chosen for its heavy reliance on IDS for operational functions and routine IDS interactions with other governmental and non-governmental agencies [36]-[39].

This case study will involve conducting semi-structured interviews with data sharing practitioners located at the case study site. The findings derived from both the SLR and case study will serve as the foundation for the initial GIDS Framework. Subsequently, this framework will be presented and validated to the case study participants to gather additional insights and refine it into the final proposed GIDS Framework.

2. Theoretical Basis

A. Interagency Data Sharing and Data Quality Issue

The evolution of E-Government has significantly reshaped the landscape of governmental operations, emphasizing technological advancement and information exchange between agencies [2]. The significant benefits of IDS in e-government particularly reside in the context of electronic integration among government and nongovernment agencies, bringing positive impact on different aspect such as productivity, performance, policy-making processes, and overall government service delivery [17].

Effective IDS plays a crucial role in internal administration, external service delivery, and addressing complex social issues [22]. It can facilitate one-stop services, scientific decision-making, precise governance, and efficient public service delivery. This is due to collaborative efforts that enable one agency to surpass the capabilities of its individual agencies. Information sharing also reduces duplicate data collection and processing efforts, leading to decreased operating costs and increased service efficiency. Moreover, obtaining business-related information from other agencies enhances the government's overall informativeness, helping agencies navigate environmental complexity and uncertainty. This, in turn, improves decision-making capabilities and contributes to solving complex social issues, ultimately promoting social development [22].

Data issues can sneak into every phase of the data life cycle, starting from initial data creation and procurement/collection through various touch points like Internet, call centers, sensors, mailings, and branch offices, to name a few, to data processing, transfer, and storage, to archiving and purging. Thus, data quality is impacted by activities in all of the phases in the data life cycle [31]

The challenges associated with data quality in the realm of IDS encompass the informatics capability gap, technical compatibility challenges, and concerns about the reliability and trustworthiness of shared data [15] , [17] , [19], [20], [21], [22], [28], [36]. The data quality concerns even contribute to agencies' reluctant to share data due to uncertainties about the quality of their data, potentially risking their image and the fear of disseminating incorrect data [17], [22].The literature underscores the urgency of addressing these issues for successful interagency collaboration.

More research has contributed to issues related to IDS challenges. On the technical aspect, the significance of metadata, data standards, and evaluation mechanisms for data quality adds another layer to the coherence of these findings [28]. This emphasized by the essential elements of data consistency, interoperability, and the need to address legal barriers to ensure effective inter-

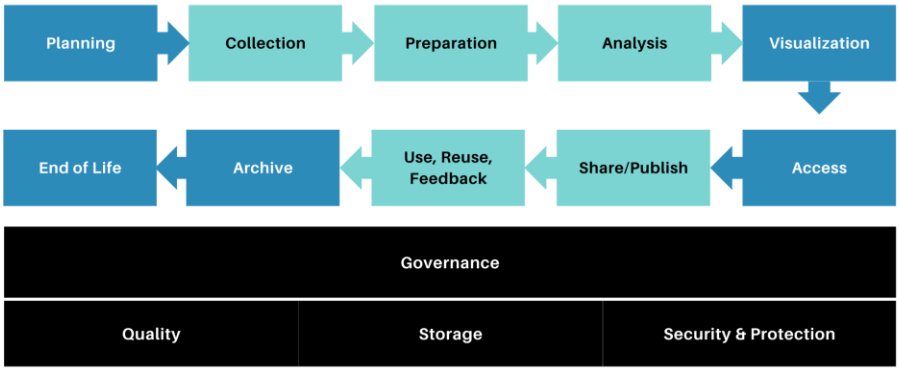

Figure 1. DaLif Data Life Cycle Framework.

agency data sharing [21]. On the organizational relationship aspect, lack of interagency trust contributes to the hesitancy in information sharing, with poor data quality cited as a reason for data being deemed unwanted [15]. The rapid availability of complex data further emphasize that data value extends beyond mere collection and requires correct interpretation [20]. The varying degrees of informatization and digital literacy among public officials also add additional challenges to achieving standardized and consistent data across sectors [36].

Together, these studies paint a comprehensive picture of the multifaceted issues surrounding data quality and its pivotal role in successful information exchange among agencies.

B. Data Quality

Recognizing the pivotal role of data as a valuable and strategic asset, DAMA-DMBOK [29] underscores the critical necessity of managing data quality throughout its entire lifecycle. This dimension lays the groundwork for comprehending the significance of data quality in the specific context of IDS within government agencies. The emphasis is on adopting a proactive and integrated approach to ensure that data maintains its reliability and value from its creation to eventual disposal.

In this study, the authors utilized the Data Management Body of Knowledge (DM-BOK) offered by DAMA-Data Management Association, considering it as an industry standard for data management. DM-BOK stands out as a comprehensive data-oriented framework, distinguishing itself from TOGAF (The Open Group Architecture Framework) and COBIT (Control Objectives for Information and Related Technologies.

DMBOK, managing data throughout its lifecycle, offers a detailed framework supporting the development and implementation of data management processes and procedures. However, it's worth noting that such data management frameworks often provide generic guidelines. The literature on data quality serves as a foundational pillar in understanding the intricacies of managing data in the context of IDS. It establishes the groundwork for the subsequent exploration of challenges and solutions associated with data quality issues in interagency collaboration.

C. Data Lifecycle

Examining the lifecycle of IDS is essential for a comprehensive understanding of its dynamics. Numerous benefits arise from consistently designing and implementing a data lifecycle for public administrations, with an emphasis on ensuring data quality [30]. These benefits include simplified data management, identifying all essentials activities related to data, user support, better data understanding, effective data and metadata gathering, better data analysis, and better data quality [30], [41]-[49]. Furthermore, data lifecycle models provide an effective solution for data management, offering a high-level framework to plan, organize, and manage all aspects of data during its life stages, with an inherent focus on maintaining data quality [38].

Therefore, to effectively address data-related issues in IDS practices, it is crucial to map these issues within the data lifecycle. By doing so, the challenges and complexities associated with IDS can be better understood and managed.

DaLif is a data lifecycle framework formed for data-driven organization (see Figure 1). This framework was derived by analyzing 76 model data lifecycles. Consequently, this study will utilize the data lifecycle framework put forward by Shah et al. [30].

Figure 1. illustrates the mandatory phases in green, with optional phases highlighted in blue. Additionally, the black-colored phases denote horizontal activities executed across the entire

lifecycle. Notably, the storage phase is both mandatory and horizontal.

The DaLif framework is constructed through the examination of research from IEEE, ACM, ScienceDirect, and Springer. However, DaLif recognizes that there may be valuable research in other libraries as well. The GIDS Framework seeks to broaden the framework's scope by integrating materials from additional libraries and incorporating case studies. Additionally, DaLif primarily focuses on data in data-driven government without specifically addressing issues related to the IDS field. Therefore, this research aims to fill this gap by incorporating studies and case studies pertaining to IDS.

D. Previous IDS Framework

Previous research has endeavored to create frameworks pertaining to data sharing. These prior studies are outlined as follows.

- Higgins et al.,[31].
  This framework identifies organizational activities for which external (shared) data would be beneficial, defines shared data for these activities, determines the availability of shared data, addresses data acquisition changes, establishes secure data transfer methods, integrates new data into existing databases, analyzes shared data for decision making, and examines changes in organizational activities resulting from data sharing.

- Thompson et al. [46]
  This research explores the creation of shared data elements and a novel data repository to facilitate data sharing and analysis. It underscores the crucial role of standardized data collection methods and quality assurance protocols in interagency data sharing endeavors. The document offers perspectives on instituting shared data elements and a data repository for collaborative data sharing, aligning with the interagency data sharing framework. It underscores the importance of standardized data collection methods and quality assurance protocols, integral elements of a robust framework for interagency data sharing.

- Tungkasthan et al. [47]
  The presented framework offers a structured approach to government data sharing, encompassing considerations related to dataset categorization, user classification, Data-Information-Knowledge-Wisdom (DIKW) modeling, data requirement levels, and access control. It provides valuable insights for policymakers and practitioners seeking to establish secure and efficient public data sharing practices.

- Metcalfe et al. [48]
  This study investigates the establishment of trust to expand and strengthen partnerships in both industry-to-industry and industry-to-government collaborations, crucial for the future success of information exchange. The framework proposes that trust is influenced by four factors: capability, humanity, transparency, and reliability.

- Rukanova et al. [49]
  This study introduces a framework for examining the obstacles, motivators, and facilitators of voluntary information sharing between businesses and the government, along with the governance mechanisms that facilitate such voluntary sharing of information. The authors' analysis indicates that successful voluntary business-government information sharing is achievable when there are compelling motivators and a proactive government agency willing to spearhead the initiation of the process.

A noticeable gap remains in these frameworks, they have not extensively addressed considerations related to data quality. To bridge this gap, it is essential to integrate a dedicated focus on data quality assurance within the broader framework of data sharing initiatives. This ensures that the shared data is not only comprehensive and secure but also maintains high quality standards. The proposed knowledge governance framework, as discussed earlier, can serve as a foundation for addressing this crucial aspect and enhancing the overall effectiveness of data sharing initiatives.

3. Methodology

This research study utilizes a three-stage research design model comprising a literature review and a case study.

A. Literature Review

The research methodology concludes by combining findings from the Systematic Literature Review (SLR) and case study to provide valuable insights into the intricate dynamics of Interagency D ata Sharing (IDS) in government. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline is employed for the SLR, acknowledged for its widespread recognition and respect.

The authors developed a literature selection technique, specifying databases, keywords, and search criteria (inclusion and exclusion). In the initial stage, a review was conducted across various database sources to obtain comprehensive

and diverse data, utilizing reputable sources such as ACM Digital Library, IEEE Xplore, JSTOR, ProQuest, ScienceDirect, and Taylor & Francis.

Following that, automated searches were conducted on selected databases utilizing categorized keywords. These keywords were classified into three categories to refine the search, facilitating the identification of synonyms and interchangeable terms from previous studies that are pertinent to the research focus. The keywords categorization as indicated in Table 1.

Table1. Search criteria.

Categories	Keywords Lists
Object	Data, information
Activity	Sharing, Exchange
Subject	Interagency, intergovernmental, interinstitutional, interorganizational, cross-sector, cross-border, cross-boundary, cross-institution, cross-organization, cross-sectoral, public sector, public organization, government
Scope	Data quality, practices, issue, challenge, data lifecycle

Boolean operators, guided by the formulated keywords in Table 1 ensured a comprehensive retrieval of relevant studies. The formulated keywords were structured into a Boolean search query: ("DATA" OR "INFORMATION") AND ("SHARING" OR "EXCHANGE") AND ("INTER" OR "CROSS" OR "PUBLIC" OR "GOVERNMENT") AND (“DATA QUALITY” OR “PRACTICE” OR “ISSUE” OR CHALLENGE” OR “LIFECYCLE”).

To focus on recent studies, the authors applied a time filter from 2019 to 2024. The query was then employed on the publication title, abstract, and keywords to identify relevant publications. The search criteria, comprising inclusion and exclusion criteria, guided the SLR process (Table 2). Quality assessment questions (see Table 3) then applied to evaluate selected articles based on object discussed in the research, clarity of research objectives, unambiguous presentation of results, interpretation and discussion of results, and conclusiveness in addressing research questions.

The article selection flow, depicted in Figure 2, illustrates this systematic process.

Following selection, the chosen articles are meticulously read by the authors for a minimum duration of 30 minutes. During this phase, particular attention is paid to gathering information relevant to addressing the research questions. Extracted data from the articles includes various aspects such as the context of data sharing, practices, associated activities, challenges or barriers, issues regarding data quality, and potential solutions or enablers for effective IDS practices. Subsequently, this gathered information

Table 2. Inclusion and exclusion criteria.

Type	Criteria	Code
Inclusion	Articles published between 2019-2024	IN1
	Articles written in English	IN2
	Articles focus on government or public administration	IN3
	Articles related to IDS practices	IN4
Exclusion	Full-text access not available	EX1
	Working papers, presentation	EX2
	Duplicated studies	EX3

Table 3. Quality assessment questions.

Code	Question
Q1	Does the article discuss real-life IDS practices or data-quality issue?
Q2	Are the research objectives clearly stated?
Q3	Are the presented result unambiguous?
Q4	Are the results interpreted and discussed?
Q5	Does the conclusion answer the research questions?

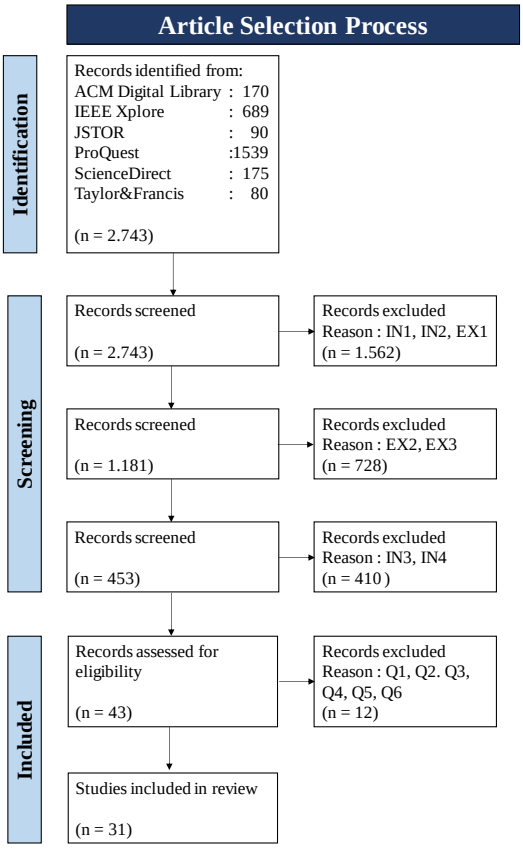

Figure 2. Article selection process.

is synthesized and organized to construct an initial comprehensive framework aimed at tackling challenges related to IDS practices and addressing data quality concerns.

B. Case Study: Directorate General of Taxes, Ministry of Finance Indonesia

In accordance with the mandate of Law Number 7 of 2021 Article 2 [50], the taxation

system in Indonesia operates on a self-assessment basis. Under this system, taxpayers are entrusted with the responsibility of calculating, reporting, and paying their taxes based on the information they possess. In this context, the Directorate General of Taxes (DGT) plays a critical role in ensuring the integrity and compliance with tax obligations.

A primary challenge in implementing the self-assessment system is to ensure that the information provided by taxpayers is accurate and aligns with their actual tax liabilities [32]. In this regard, the DGT needs to gather information from other institutions that can be used as benchmark data for taxpayers' tax transactions and compliance [32], [51], [52]. Given that the DGT is responsible for more than 70% of the nation's total revenue [53], it's crucial to recognize that any lapses in ensuring taxpayer compliance could lead to potential revenue loss for the country.

To meet this need for information from third parties, the Ministry of Finance subsequently issued Ministerial Regulation No. PMK-228/PMK.03/2017 concerning the Details of Types of Data and Information as well as Procedures for Submitting Data and Information Related to Taxation (hereafter referred to as PMK-228) [35].

PMK-228 outlines the obligations of Government Institutions, Agencies, Associations, and Other Parties (ILAP) to periodically submit data to the DJP. The data from ILAP then grouped into six types of data and information: wealth or property, debt, income earned or receive, expense incurred and/or cost, financial transactions, and economic activity. The gathered data then undergo further processing to be utilized by the DJP, not only for assessing taxpayer compliance but also as raw material for the innovation of DJP's e-government services.

According to PMK-228, 69 ILAPs are required to regularly provide their data to the DGT, totaling 349 data. Additionally, this requirement does not encompass additional institutions obligated to submit their data to the DGT through other regulations, such as the Country-By-Country-Report (CBCR), Automatic Exchange of Information (AEoI), data from interoperable services, and ad hoc Interagency Data Sharing Agreements.

In implementing PMK-228, the Directorate of Data and Information of Taxation (DIP) is assigned by the DGT as the unit responsible for conducting the entire process of data exchange, encompassing planning, collection, processing, analysis, and dissemination of data. The management of this data is then governed by the data governance framework outlined in Director General Decision No KEP-215/PJ/2021 regarding the Data Governance in the DGT (hereafter referred to as KEP-215) and Circular Letter No SE-12/PJ/2023 regarding the Guidelines for the Implementation of Data Governance in DGT (hereafter referred to as SE-12) [33], [34].

To gain insight into data sharing practice in DGT, the author conducted interviews with data-sharing practitioners from the DIP, DGT. A semi-structured interview comprising five key questions was devised for this purpose (see Table 4)

Table 4. Semi-structured questions.

No	Questions
1	How is the data sharing process within your organization's interagency collaborations?
2	What specific challenges or issues have been encountered in the current data sharing practices within your organization?
3	How does your organization ensure that shared data complies with relevant regulatory frameworks and standards?
4	What is the data quality aspect that need to be considered within your organization’s interagency collaborations?
5	How is the quality of shared data managed and monitored within your organization's interagency collaborations?

Five individuals were selected for the interviews, comprising the Head of External Data Quality Assurance, Data Analyst, Senior Data Engineer, Data Engineer, and Data Gathering Staff. These individuals were chosen based on their extensive expertise and experience in data sharing, spanning at least five years. Each interview lasted a minimum of 45 minutes, allowing for comprehensive discussions and effective capturing of insights.

The findings from this case study provide valuable insights into the practices, challenges, and data quality issues within DGT's IDS practices. These results will be incorporated into the proposed framework to enhance its comprehensiveness and depth.

C. Framework Formulation, Refinement, and Validation

The findings from SLR process and the semi-structured interviews conducted as part of the case study were utilized to map out and create a preliminary proposed GIDS Framework. This initial framework was then presented to respondents in DGT to gather additional insights and validate its practicality and to further refine the framework. The entire process of formulating, refining, and validating the framework is illustrated in Figure 3**.**

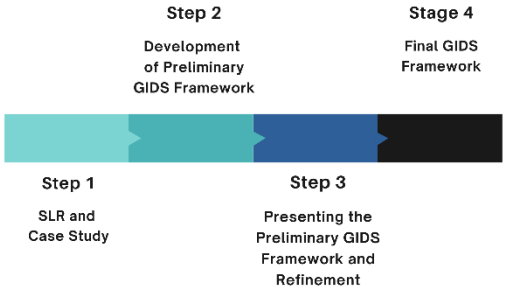

Figure 3. Research methodology.

The DGT played a crucial role in validating the initial GIDS Framework. By presenting the framework to DGT representatives, the author gathered valuable insights through a feedback session. This session focused on key questions to assess the framework's effectiveness and practicality:

1. Does the framework address and potentially mitigate critical challenges encountered in GIDS implementation?
2. Does the framework include all essential phases of the GIDS lifecycle?
3. Does the framework adequately address the data quality aspects within GIDS practices?
4. Does the framework address and potentially mitigate critical challenges encountered in GIDS implementation?

The feedback from the DGT used to refine the framework, which will be presented in the next section.

4. Result and Analysis

A. Identified Challenges and Enablers

This analysis looked at 31 studies. Almost half (47.5%) were done in Europe, with the research taking place in Netherlands, Portugal, and the UK being the top contributors. Research from the America made up 22.2%, mostly from the United States. Sharing data between government was the most common focus (58.5%), followed by government-business (34.1%) and government-citizen (7.3%) interactions. Over a third of the studies (34%) explored data sharing within public agencies for e-government purposes. Other areas of interest included criminal justice and customs (20%), healthcare (17.1%), and infrastructure, transportation, social welfare, culture, energy, and geospatial data (remaining 29%). The distribution graph from the previous study as shown in Figure. 4, Figure. 5, and Figure. 6.

As previously mentioned, this framework aims to tackle the issue of identifying challenges or obstacles that could hinder the effective implementation of IDS.

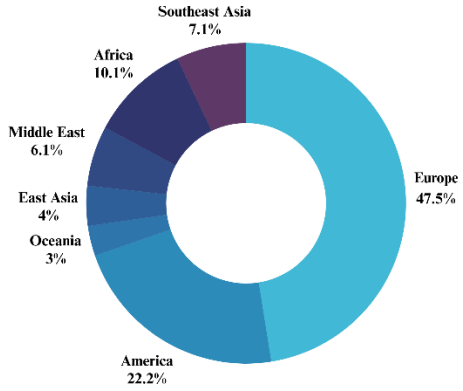

Figure 4. SLR result based on location.

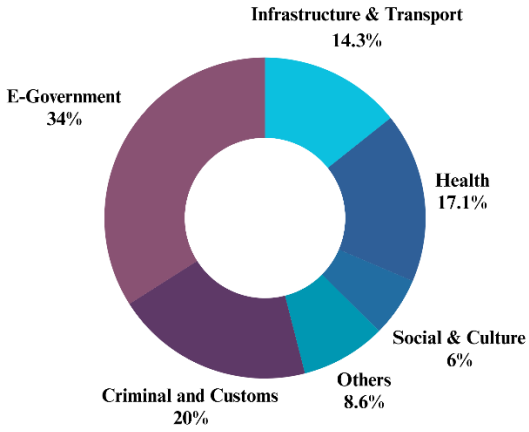

Figure 5. SLR result based on theme.

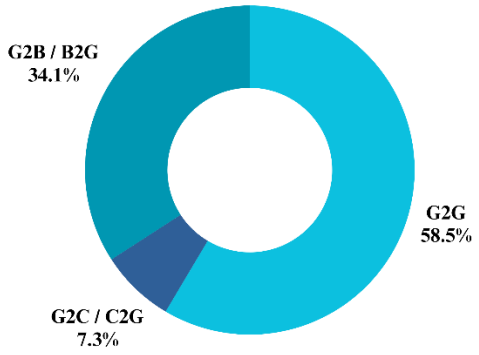

Figure 6. SLR result based on data sharing context.

Through the use of the PRISMA SLR Method, 31 studies were analyzed to pinpoint existing challenges and potential barriers. The SLR results provided insights into the distribution of IDS domain goals and suggested solutions. The mapping of the research based on challenges/barriers, data quality issue, and enabler/solution are shown in Table 5, Table 6, and Table 7.

Table 5. Challenge/barriers from previous study.

Challenges/Barriers	Research
Cost constrains	[17], [19], [54], [55], [56], [57]
Data governance in participating agencies	[17], [19], [56], [57], [58]
Data Quality	[17], [20], [49], [54], [59]
Knowledge gap between participating agencies	[56], [60], [61]
Lack of clarity of roles and responsibility	[17], [54], [55], [56], [60], [61], [62], [63]
Lack of resources	[56], [64], [65]
Organizational structure, goal, norms, and value	[17], [54], [56], [60], [66], [67]
Organizational technical capability and capacity	[19], [48], [49], [54], [57], [67], [68], [69]
Perception towards data or sistem's confidentiality, security, and privacy	[17], [19], [54], [56], [58], [64], [65], [66], [70], [71]
Perception towards data's value	[17], [54], [56], [65]
Political and legal constrain	[49]
Trust	[17], [20], [48], [54], [57], [61], [64], [66], [69], [72], [73]

Table 6. Data quality issue from previous study.

Data Quality Issue	Research
Data Accuracy and Reliability	[56], [60], [74]
Data Availability	[54], [60], [75]
Data Completeness	[74]
Data Interoperability	[54], [56], [61], [63], [66], [76]
Data Standardisation	[54], [56], [59], [64], [69], [74], [76]
Data Timeliness	[60], [74]
Data Update	[56], [59]

Table 7. Enablers and solutions from previous study.

Enablers and Solutions	Research
Coordination and communication	[17], [20], [36], [49], [55], [62], [65]
Data sharing agreeement	[17], [20], [54], [55], [56], [58], [64], [69], [72], [74], [76]
Data sharing platform	[20], [55], [60], [61], [62], [63], [64], [74]
Establishing data standard	[20], [54], [57], [61], [64], [72], [76]
Establishing metadata system and standard	[54], [72], [76]
Financial support	[19], [36], [49], [54], [56], [64], [67], [69]
Leadership and executive support	[17], [20], [49], [56], [60], [61], [76]
Regulatory support	[17], [20], [56], [61], [64], [71], [72]
Shared goals and benefits	[17], [20], [57], [65]
Training/technical support	[57], [59], [60], [70]

B. DGT's IDS Framework Overview

DGT has adopted KEP-215 as the framework for its entire data governance, encompassing data received from a third party called ILAP. In managing this data, KEP-215 follows DM-BOK as its guiding principle for framework development. The framework is divided into three main components: data governance organization, data lifecycle, and data governance.

For the data governance organization, DGT has divided responsibilities into two parts: data governance administrators and those involved in data governance implementation. The administrators include a directing committee, Chief Data Officer (CDO), Data Governance Office (DGO), Working Group, and Stakeholder group. Meanwhile, those involved in implementation consist of business owners, data steward, ICT unit, data producers, and data users.

Regarding the data lifecycle regulated by DGT, it comprises planning data policy needs, data collection and preparation, data storage, data utilization, and monitoring and evaluation. Further details regarding each data lifecycle phase are provided below:

Planning: DGT's planning phase begins when data steward formulates Data Discovery Policy based on the DGT's strategic plan, unit needs within the DGT environment, and coordination results with directorates/regional offices. Data Discovery Policy are formulated annually, with a release deadline at the end of January of the following year.

Collection: The second phase in the data lifecycle is data collection. In this phase, DGT collects data through IDS based on established regulations, such as those set by the Ministry of Finance.

Preparation: In this phase, the collected data undergoes preparation. Data preparation activities include researching the quantity and types of data, data completeness, data redundancy, and the data period received. The results of data research are then recorded in the external database. The recorded data is then identified using Taxpayer Identification Numbers (NPWP) or National Identification Numbers (NIK). Subsequently, quality control is applied to the external data based on the identification results.

Use: In this phase, data is utilized by DJP units and external parties of DGT. The utilized data results from matching and analysis conducted by the relevant directorates.

Storage: This phase describes activities related to archiving electronic and physical documents, as well as borrowing physical documents from external data collection activities.

From the interview results, it was found that the challenges in managing IDS practices at DGT include a **lack of documentation** regarding data requirement planning, leading to a lack of

oversight regarding the history and analysis of data requirements with the necessary data use cases. Additionally, **collaboration** processes among internal units involved in data processing are crucial to obtaining comprehensive information about the stages of data processing conducted, especially regarding data identification processes and data quality assurance. Another crucial issue is the absence of assessment matrices for **monitoring and evaluation** purposes, which could be used to assess IDS performance at each stage of the process. This is essential because the evaluation results can inform improvement measures for the framework.

C. IDS Data Lifecycle

The authors constructed the IDS data lifecycle by integrating the data lifecycles observed in the case study, activities found in previous research, and DaLif data lifecycle framework proposed by Shah et al. [30]. This comparison is presented through the comparison depicted in Table 8.

Table 8. Data lifecycle formulation.

Lifecycle	DaLif	DGT	Previous Study	GIDS
Planning	√	√	[77]	√
Collection	√	√	[60], [67], [68], [75], [76]	√
Preparation	√	√	[75]	√
Analysis	√	√	[67], [75]	√
Visualization	√			
Access	√			
Share/Publish	√		[60]	√
Use, re(use), feedback	√	√		√
Archive	√		[65], [74]	
End of Life	√		[74]	√

In this part, the proposed data lifecycle is investigated by grouping and analysing phase titles, activities, and process. This is conducted by using the following the logical rules in DaLif formulation such as grouping phase with similar meaning (e.g., data collection, data gathering), removal of phases that that too generic or specific, and combining phases with similar activities.

D. GIDS Framework

This section presents our final proposed framework called GIDS Framework that has been presented and validated by the case study participant (see Figure. 7). This framework consists of three main elements: **Data Governance**, **Data Lifecycle**, and **Promoting Activities**.

These components are derived from the DaLif Framework by aligning challenges and solutions, identifying key aspects of data governance, and delineating the data lifecycle through a process of SLR and case study analysis.

1) Data Governance

This section is derived from DAMA-DMBOK, which asserts that the management of data quality, data security, and data storage aspects is anticipated throughout the entirety of the data lifecycle. DaLif framework similarly aligns with this concept, incorporating data governance by integrating quality, security, and storage aspects throughout the entire data lifecycle. Sharing information effectively relies not only on willingness of parties but also on an organization's ability to manage information. This means having the necessary resources, like technology and skilled staff, to gather high-quality information, analyze it, and then share it with others [67], [78], [79].The subsequent explanations delve into the details of each facet of data governance.

a) Data and Metadata Quality

Data Quality Management plays a pivotal role in the broader context of data management. When data is of low quality, it tends to represent not just a cost but also a risk, rather than adding value. The challenge for organizations lies in effectively managing data quality, primarily because data often emerges as a byproduct of operational processes, and explicit quality standards are frequently lacking. Given that the quality of data can be influenced by various events throughout its lifecycle, it becomes essential to incorporate quality planning as an integral part of the data lifecycle.

To guarantee data quality at each phase of the data lifecycle, organizations must define data quality dimensions and establish methods for measurement and assessment. Data quality dimensions are attributes of data quality that can, when measured correctly, indicate the overall quality level of data [27]. The identification of relevant quality dimensions can be seen as a starting point to the subsequent assessment phase and builds the basis for various improvement activities [80].

The most common dimensions are completeness, timeliness, and accuracy, followed by consistency and accessibility. The definitions of these dimensions according to [81] are as follows:

- Completeness: The extent to which data are of sufficient breadth, depth and scope for the task at hand.
- Accuracy: The extent to which data are correct, reliable and certified.
- Timeliness: The extent to which the age of the data is appropriate for the task at hand.
- Consistency: The extent to which data are

presented in the same format and compatible with previous data.

- Accessibility: The extent to which information is available, or easily and quickly retrievable.

In addition to delineating data quality dimensions, it is imperative to articulate the steps for measuring and assessing data quality. The subjective measurement of data quality can involve obtaining user ratings on the quality levels of various dimensions [27]. Alternatively, data quality metrics, involving computations that offer an indication of the data quality level, can be established.

Regarding data quality assessment, DAMA-DMBOK asserts that such assessments provide insights into existing issues and obstacles, as well as the repercussions and risks associated with subpar data quality. These assessments can pinpoint business processes at risk when operating with low-quality data, along with delineating the financial and other benefits of integrating a data quality program into data governance initiatives.

Moreover, as metadata represents a form of data crucial for organizational data management, its quality must be treated with the same level of attention and management as the data itself. Organizations derive more value from their data assets when the data is of high quality. The quality of data hinges on governance, with metadata playing a critical role in data governance by elucidating the data and processes that drive organizational functionality [29]. It is required to possess metadata for the purpose of describing and elucidating data, as it contributes to enhancing data accuracy. The specifics of metadata vary for each infrastructure organization based on the nature of the data they handle [20].

b) Data Security

Moving on to the second data governance aspect, data security plays a crucial role in data governance. It involves implementing measures to safeguard sensitive information during sharing. This comprehensive approach is essential for managing data risks, ensuring the confidentiality and privacy of data throughout its lifecycle. The complexities associated with information chains, involving diverse organizations and data sources, underscore the need for a robust framework addressing privacy, security, and confidentiality during data sharing.

Three key aspects of security issues in data sharing: encryption, access control & authentication, and physical security [82]. The quantity, variety, and sensitivity of big data, managed through heterogeneous technological solutions, emphasize the criticality of data security and privacy protection [30]. Security is seen as as a crucial factor in encouraging businesses to share data with the government, as clear business rules outlining data access permissions are essential for building trust and facilitating data sharing [71].

Legal and policy aspects in data sharing focus on protecting shared information, introducing stewardship and usage concepts. Stewardship entails organizations' responsibility for ensuring data integrity and quality, while usage emphasizes information's usefulness to intended users. Formal mechanisms, complemented by informal measures like social control and trust-building, address coordination needs and collaboration problems in data governance and security. The increased importance of data security in autonomous data

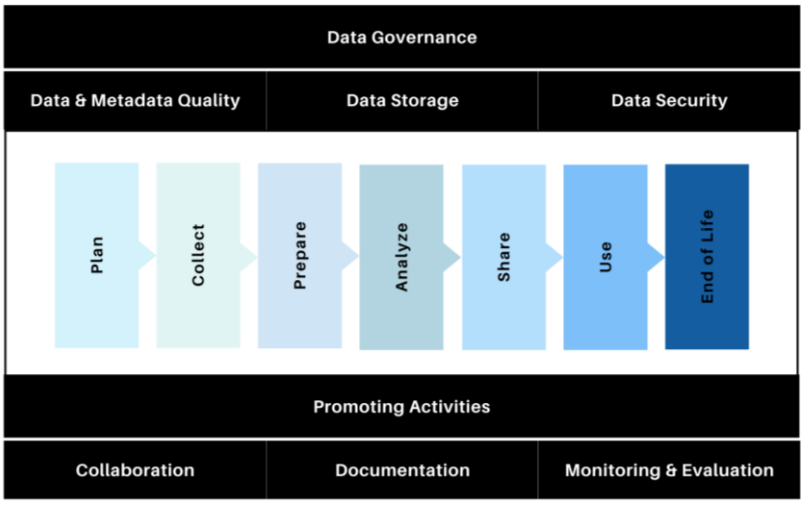

Figure 7. GIDS Framework.

sharing then delineates the necessity of open and standardized infrastructure, environments, and procedures to protect the integrity, trust, and confidentiality of data [69]. Some research has underscore utilizing advance technology blockchain and electronic certificates to mitigate information security issues [22].

c) Data Storage

The third element, data storage, underscores the vital importance of establishing robust storage protocols to maintain data integrity within the broader context of data governance. The primary goal during the storage phase is to securely preserve data throughout its lifecycle [30]. This necessitates a comprehensive approach encompassing the design, implementation, and support of stored data, ultimately aiming to maximize its value throughout its entire lifecycle, as there is there is a potential for data loss across different phases of the lifecycle [20]. Moreover, within an organization, each department claims ownership of its data, and upon completing one phase, they often initiate a new one without storing the relevant data, leading to data loss.

Another point of data storage is to make sure that the data is not only stored but stored properly. The issue of improper data storage resulted in some data being stored, but it was difficult to retrieve due to the unclear whereabouts, names, and formats [20]. Although there are internal platforms for data storage and sharing within the infrastructure organization, they suffer from outdated systems and slow data connections. Consequently, individuals often resort to storing data on their hard drives, leading to data loss for future projects. Moreover, the requirement for specific data formats on a platform poses challenges, particularly in collaborations with other infrastructure owners with their unique data formats and standards.

2) Data Lifecyle

Aligned with the stages identified in the literature, the data lifecycle component unfolds in seven sequential steps as follows.

a) Plan

The planning phase delineates the commencement of the IDS, encompassing the establishment of the data sharing project's objectives, formulation of essential policies and procedures for data management, specification of data sources and methods, and the identification and resolution of potential data quality issues associated with the incoming data from IDS.

Before initiating IDS, understanding how data is generated in IDS partner institutions is essential. Since data creation does not occur within one's own institution, full control over the process is limited. Instead, the focus is on assessing potential data quality problems associated with specific data creation methods. DAMA-DMBOK emphasizes that data entry processes often lead to data quality issues, influenced by factors such as the proficiency of front-office staff, inconsistent business process execution, and data entry interfaces.

Activities related to planning phase include:

- Prepare for the resource [30], [82]
- Analyzing data requirements based on business goal and strategy [30], [33], [34], [77]
- Provide detailed description of the data requirements, consist of data source, data elements, data format and data type, methods of collecting the data, and periodicity [30], [33], [34]
- Initiating IDS with collaborating agencies,
- Assigning roles and responsibilities [20], [30], [34]
- Developing data sharing agreement [30], [64]

In planning on initiating IDS, several key concepts play crucial role in ensuring the quality and successful IDS initiatives, such as the need of clarity or roles and responsibility, trust, interagency acquittance, and data sharing agreements.

Clarity of roles and responsibility (CRR) essential for making sure that the right people have the right authorities in IDS, such as who should participate in data preparation process and who will be responsible in case of data breaches or misuse [54], [60]. Reduced ambiguity, complexity, and information asymmetry in collaborative efforts result from clear roles and responsibilities, crucial for successful outcomes [83]. The use of boundary objects, such as plans, formal meetings, diagrams, maps, and information systems, plays a critical role in achieving CRR for organizations involved in data sharing initiatives [62].

Trust is a critical factor in successful information-sharing initiatives [2], [18], [48], [69]. Conversely, a deficiency in trust among government departments has been pinpointed as a cause for inactive data sharing [36]. This concept of trust extends beyond simply believing in the absence of malicious intent. It encompasses the additional dimension of trusting the other party's competence in handling information deemed valuable [84]. When this trust in competence is eroded, the willingness to remain engaged in collaborative efforts and share information diminishes significantly [18], [84], [85]. This lack of trust, encompassing apprehensions about the future utilization of shared data, imposes

constraints on data sharing, including cross-border transfers [64]. Hence, it becomes imperative for participating institutions to reach a consensus on these aspects and formalize them into an agreement, contract, or policy [30], [83].

In terms of data, it becomes essential in planning phase to formulate a standardized format for shared data, as a lack of standardization poses a hindrance to effective cooperation [64]. Challenges often arise in data sharing schemes due to inadequately defined and overly generic data standards [15]. The study emphasizes that poorly crafted data standards for data sharing can result in significant complications. To address this issue, it is crucial to develop data standards collaboratively involving both the data owner and the data receiver. This collaborative approach is vital because it is intricately linked to the data ability, indicating an organization's capability to comply with or provide the required data. Neglecting this aspect may result in shared data failing to meet the needs of the data requester, because the agency operates based on its own set of data standards.

The development of interagency acquaintance proves to be a crucial element in improving the clarity and expressiveness of data [15]. The research indicates that individuals from diverse agencies who already have established relationships are more prone to effectively prepare, articulate, and share data. Conversely, a deficiency in interagency acquaintance can present obstacles, complicating the identification of appropriate contacts for data requests and potentially leading to the misdirection of data to an incorrect agency [15].

There are three main types of data: structured data, semi-structured data, and unstructured data [27], [86]. Structured data is a generalization or aggregation of items with elementary attributes within a domain item, while semi-structured data has a structured format with some degree of flexibility. Unstructured data is described as a generic sequence of symbols [27], [86]. Addressing data quality issues is crucial, and obtaining information about the data type during the planning phase is essential. This is because the processing and assessment of data may vary based on its type. Unstructured data presents a significant challenge, as many techniques used for assessing structured and semi-structured data are not applicable to unstructured data [87]. Therefore, mitigating the potential risks to data quality associated with unstructured data should be done proactively.

Data sharing agreements involve a collaborative understanding between the entity providing the data and the one receiving it. Essential aspect should be addressed and explicitly ruled in data sharing agreement, such as responsibilities of both the data provider and recipient, the intended purpose of the data, a comprehensive description of the data, the agreed-upon time frame for the arrangement, accessibility to the data, and compliance with legal requirements for lawful data processing [64].

The Singaporean Government's Trusted Data Sharing Framework outlines seven key principles for crafting data sharing agreements. These principles aim to ensure responsible data sharing practices, encompassing the delineation of permitted data uses, establishment of data ownership safeguards, clear liability assignment for breaches, robust measures for maintaining data confidentiality, a defined agreement duration, adherence to relevant legal frameworks, and careful consideration of technical infrastructure for secure data exchange [69], [88]. The OECD Council also established recommendations to facilitate data access and sharing. These recommendations provide initial guidance for structuring data-sharing arrangements between diverse stakeholders and actors involved in the data ecosystem [89]. These agreements are intricately connected to the previously mentioned concept of trust and clarity of roles and responsibility.

b) Collect

Collection phase consists of set of activities through which data is gathered from different internal and external sources and in different formats i.e. structured, unstructured and semi-structured forms [30].

Activities related to collection phase include:

- Gathering data from sources [20], [21], [30], [48], [90]
- Gathering data reference, data dictionary, and metadata [30]
- Conduct initial data quality assessments and communicate the findings to the related party [33], [34], [91]

In IDS data collection phase, several key concepts play crucial role in ensuring the quality of the data, such as initial data quality assessment, data dictionary, and data standardization.

The primary objective of an initial data quality assessment is to gain insights into the data and lay the groundwork for an effective improvement plan [91]. This stage involves a thorough evaluation, aiming to detect potential data quality issues early on and establish a solid foundation for comprehensive data quality management. The simplest form of initial data quality assessment can be conducted by comparing the received data with agreed-upon

data standards. Consequently, the outcomes of this assessment can, at the very least, address the completeness dimension of the data.

A data dictionary refers to a compilation of names, definitions, and attributes related to data elements [82]. It serves to elucidate the significance and purposes of data elements, offering guidance on interpretation, accepted meanings, and representation. Additionally, a data dictionary furnishes metadata about these data elements, aiding in delineating their scope and characteristics, along with establishing rules for their usage and application. The utilization of a data dictionary plays a crucial role in enforcing data standards, ensuring uniformity in the collection and utilization of data.

The term "data standardization" refers to a mutually agreed-upon set of data established by both involved parties. When establishing data standards for the purpose of data sharing, it is crucial to consider the data-sharing capabilities of the collaborating agency. Imposing data standards that are unattainable for the collaborating partner may impede collaborative initiatives rather than facilitating them [64]. Data standardization can address issues such as inconsistency in data structures, parameters, metadata, discrepancies in reporting across sources, variations in reporting mediums, flexibility in data structure, and privacy concerns [92]. Furthermore, the absence of data standardization can impact the usability and comprehensibility of the data for all involved parties [20].

c) Prepare

While the primary focus of endeavors to enhance data quality often revolves around preventing errors, DAMA-DMBOK acknowledges that data quality can also be elevated through certain types of data processing. Data processing involves transforming data into more refined forms of information. In the context of IDS, data processing is employed to ensure that data received from collaborating agents adheres to the data standard criteria established within one's organization. However, DAMA-DMBOK highlights issues arising from data processing functions, including incorrect assumptions about data sources, outdated business rules, and altered data structures.

Activities related to processing phase include:

- Conduct data cleansing, data transformation, data standardization, and data enhancement [29], [33], [34], [92], [93]
- Conduct quality assurance on data processing data and process [33], [34]

In IDS data preparation phase, several key concepts play crucial role in ensuring the quality of the data, such as data cleansing, data transformation and standardization, data enhancement, and quality assurance.

Data Cleansing or scrubbing is a process that modifies data to adhere to data standards and domain rules [29]. This involves identifying and correcting data errors to enhance data quality to a satisfactory level, as outlined in DAMA-DMBOK. It's important to note that there should be a defined level specified in the service level agreement. This phase involves systematic error identification and correction procedures, ensuring a methodical approach to rectifying errors and continuous monitoring of data quality.

The technical capacity of a government department for IDS is evident in its performance of data management tasks and the efficiency of its cross-department data-sharing platform, which includes activities such as data cleaning, metadata creation, data classification, and establishing data provenance [94].

Data transformation and standardization involves creating a cohesive and interoperable dataset that aligns with broader industry or organizational standards. Ensuring data content complies with standard Reference Data values (e.g., country codes) and formats is crucial. To kickstart data standardization, it's essential to establish a clear definition of data standards within one's agency, ensuring a shared understanding among data engineers.

Moving beyond mere conversion, the next stage ensures data transformation into a format suitable for analysis and congruent with organizational requisites. During routine processing, data rules are triggered to convert data into a readable format for the target architecture. However, readability doesn't always imply acceptability, as rules can be generated directly in a data integration stream or rely on alternative technologies within a tool. Data transformation, building on standardization techniques, guides rule-based transformations by mapping original data values and patterns into a target representation, incorporating rearrangements or corrections as dictated by the rules in the knowledge base. In essence, standardization is a specialized aspect of transformation, employing rules that capture context, linguistics, and idioms recognized as common through repeated analysis by rules analysts or tools.

Data enhancement, or enrichment, is the process of improving the quality and usefulness of a dataset by adding more details [29]. Some improvements come from combining data within an organization, and external data can also be bought for this purpose. Data enhancement improves a dataset by adding details like

timestamps to track when data is created or modified. It also involves using audit data for historical tracking, reference vocabularies for better understanding, and adding contextual information like location. Geographic information is enhanced through address standardization, and demographic details, such as age and income, enrich customer data.

The process of quality assurance in data processing involves ensuring that the data processed adheres to the established data quality standards set by an organization. The objective of data quality assurance is to assess both the quality level of the processed data and the procedures involved in the processing [38]. This entails a systematic evaluation to guarantee that the processed data meets the predefined quality criteria and that the processing methods align with the set standards [33].

d) Analyze

The analysis phase plays a crucial role in an organization's ability to manage a substantial amount of information that can impact the business. This phase is tasked with the development of all data analysis and analytics processes aimed at extracting knowledge and uncovering novel insights, as supported by references [37], [95]. The outcomes of the data analysis phase encompass knowledge, the revelation of new insights, the generation of new data, interpretations, and/or new datasets.

Activities related to analysis phase include:

- Select appropriate data analysis tools and techniques [30]
- Perform analysis data to extract knowledge or discover new insight [30], [33], [34]

In IDS data analysis phase, data mining plays a crucial role in ensuring the usefulness of the data acquired.

Data and text mining involve techniques like profiling, data reduction, association, clustering, and self-organizing maps, collectively used to extract insights from data [29]. Profiling identifies typical behavior for applications like fraud detection, and data reduction simplifies large datasets. Association uncovers relationships, clustering groups elements by shared traits, and self-organizing maps aid visualization. Additionally, data mining, a specific analysis form, uses algorithms to reveal data patterns. Unlike standard queries, data mining discovers unknown relationships and is crucial in the exploration phase for swift identification of data elements, uncovering new relationships, and providing structure for the classification of unclear or unknown data elements.

e) Share

In this stage, data obtained from IDS is disseminated, either internally or externally. Shared data can take the form of distributing data through a web portal or an e-government information system. During this phase, the sharing of data must be conducted cautiously, ensuring that no restrictions outlined in data sharing agreements are violated. For instance, if a collaborating agent specifies that data usage should not include the name of the collaborating agency as the data source, an organization must adhere to this requirement when sharing the data.

Activities related to share phase:

- Identifying the classification of data [20], [96], [97]
- Identifying authorized individuals who can access the data. [30], [71], [98]
- Choosing the appropriate method for sharing the data [30], [99], [100]

In IDS data sharing phase, several key concepts play crucial role in ensuring the usefulness and security of the data, such as data access and authorization and data classification.

In the context of data sharing, managing access and authority is crucial for control and security [71]. This involves identifying authorized individuals or entities with permission to access specific data sets. Organizations must carefully manage access rights to ensure that only relevant stakeholders can retrieve and use the shared information. Determining the appropriate level of access is vital, considering factors such as job roles, responsibilities, and data sensitivity. Additionally, access and authorization considerations must also account for time constraints, determining whether a data user will have access indefinitely or for a specific period. Regular reviews of access permissions are essential to adapt to changes, ensuring a dynamic yet secure approach to data sharing.

Data classification pertains to categorizing the level of confidentiality associated with the data. The classification system can vary across different organizations, ranging from general or public information to restricted, confidential, and highly confidential data [33], [34]. This categorization helps organizations establish set of rules for handling and protecting information based on its sensitivity, ensuring that appropriate security measures are applied to safeguard data in accordance with its classified level of confidentiality. Thus, contractual agreements are pivotal in defining the terms of data sharing, specifying the nature of data to be shared, and outlining expectations regarding data exchange among partners throughout and after a collaborative project [20].

f) Use

This phase emphasizes the utilization of data by end users. During this stage, there is also feedback from users regarding the quality of the provided data. In essence, it involves assessing how end users interact with and perceive the data, gathering insights into its usability and effectiveness. This feedback loop is valuable for refining and enhancing data quality, ensuring that it aligns with the users' expectations and requirements.

Activities related to use phase include:

- Employing data to fulfil business objectives [33], [34]
- Providing feedback concerning the quality of the data [33], [34], [101], [102]

In using data acquired from IDS, data quality feedback plays a crucial role in ensuring the quality improvement of the data. Data quality feedback is crucial for effective data management. It involves users providing insights into the accuracy, completeness, consistency, and reliability of the data they use. This feedback loop is vital for continuous improvement, helping organizations promptly address data quality issues. By actively seeking and incorporating user feedback, organizations can enhance their data quality standards, aligning them more closely with user expectations and business needs. Regular and constructive feedback contributes to the overall reliability and usability of the data, fostering a culture of continual improvement in data management.

A feedback system is essential to maintain equilibrium between administrative initiatives and demand from various sources [18]. Furthermore, to mitigate resistance regarding accountability risks, it's crucial to implement an efficient mechanism for allocating risks and responsibilities.

g) End of Life

In this stage, redundant, obsolete, and irrelevant data is eliminated from the system [103], [104], [105]. Consideration is given to the conclusion of the data's utility or its end of life. Implementing secure and ethical data disposal practices is crucial for responsible data management. When data becomes obsolete, organizations must carefully dispose of it to protect sensitive information and uphold ethical standards. This involves using secure methods, like specialized software tools, for irreversible erasure from storage devices to prevent unauthorized access. Ethical considerations include respecting privacy rights, adhering to legal requirements, and ensuring transparency in the disposal process.

Activities related to end of data's life as follows:

- Establishing plan, policies, and method for data retention and disposal [30], [49]
- Identifying data that requires disposal [33], [34]
- Ensuring the permanent removal of unnecessary data from storage mediums to prevent the inadvertent disclosure of sensitive information [30].

In the last phase of data acquired from IDS, data retention policy plays a crucial role in ensuring the data management is done properly. A data retention policy guides organizations in systematically managing data throughout their lifecycle. This policy establishes guidelines for retaining various types of data, considering legal, regulatory, and operational requirements. It specifies the duration data should be stored and the methods of storage, including secure disposal practices at the end of its useful life. The policy ensures compliance with relevant laws, addresses security concerns, optimizes storage resources, and promotes efficient data management. Regular reviews are essential to adapt to evolving regulations and technology.

3) Promoting Activity

The third component, Promoting Activity, focuses on fostering collaborative and efficient data sharing practices. This section encompasses collaboration, documentation, monitoring and evaluation.

a) Collaboration

In the context of the IDS data lifecycle, collaboration enhances information exchange efficiency, ensuring seamless transitions from data creation to disposal. It fosters resource and expertise sharing, improving data quality and reliability.

Collaboration among agencies encompasses various levels of interaction: establishing connections (networking), coordinating efforts effectively, and mutually pursuing objectives while sharing resources (cooperation), and increased engagement in activities that are advantageous to all parties involved (collaboration) [55].

In another context, cooperative initiatives can facilitate the voluntary exchange of information between businesses and governmental bodies [49]. This exchange of information becomes crucial when authorities lack a formal mechanism to access important business data, such as invoices and packing lists. Consequently, collaboration extends beyond mere data acquisition to obtaining information that can offer valuable insights for decision-making processes.

To foster collaboration among participating agencies, it is crucial to recognize that these

entities may be motivated to initiate data sharing due to factors such as interest, motivation, and anticipated benefits, like incentive mechanism [67], [79], [84], [94], [106], [107]. Conversely, inhibiting forces, including barriers and challenges such as legislation, organizational readiness, political pressure, and technical constraints, may also exist [21], [56], [67]. The attainment of data sharing essentially involves harnessing the driving forces and overcoming the challenges associated with sharing data [56].

b) Documentation

In the realm of interagency data sharing, the activity of documentation holds significant importance, ensuring transparency, accountability, and the overall quality of shared data. The establishment of comprehensive records to meticulously track the sharing process is a crucial aspect. Adhering to DAMA-DMBOK guidance, thorough documentation becomes imperative to avoid the pitfalls associated with incomplete records that might compromise data quality. Proficient documentation serves as a structured guide, aiding in the identification and correction of errors, providing repeatability in processes, and supporting compliance efforts. It acts as a navigational tool, enhancing the efficiency and reliability of interagency data sharing by offering a clear and traceable foundation throughout the collaborative endeavor.

Another challenge related to data involves the absence of documented records. Most data reside in individuals' knowledge rather than in written form [20]. To acquire the necessary data for the project, respondents relied on email communication or direct verbal inquiries with colleagues. This unstructured approach leads to data loss and confines information within a specific department of an organization. As one participant mentioned, "in infrastructure organizations, a significant amount of data is still stored in the minds of individuals or in someone's files," hindering not only intra-departmental data sharing within the infrastructure organization but also collaboration between different infrastructure owners for various projects.

c) Monitoring and Evaluation

In the context of the interagency data sharing data lifecycle, monitoring and evaluation are critical elements for ensuring the effectiveness and success of collaborative efforts [67]. Additionally, having clear performance measurements and enough staff to analyze the data makes organizations more likely to effectively share information [67], [108]. The effectiveness of performance evaluation and motivation hinges on the continuous monitoring of employee performance, aiming to attain performance management objectives while also serving as a means to inspire and incentivize employees [94].

The continuous tracking and observation inherent in monitoring provides real-time insights into the shared data processes, offering timely feedback to align activities with predefined objectives. Simultaneously, periodic evaluations allow for a comprehensive assessment of the impact and outcomes of interagency data sharing initiatives, ensuring that they remain aligned with overarching goals. Moreover, collaborative data evaluation is strongly encouraged, as it facilitates feedback from collaborating partners regarding the quality of the shared data. This iterative feedback loop not only enhances data quality but also fosters a culture of continuous improvement and mutual understanding among participating agencies, contributing to the overall success of collaborative data-sharing endeavors.

5. Conclusion

This paper explores the complex landscape of Interagency Data Sharing (IDS) in government, drawing insights from a thorough review of studies over the past six years and selected research databases. The primary goal is to address challenges and enhance data quality throughout the IDS process, and the proposed framework focuses on three key components: Data Governance, Data Lifecycle, and Promoting Activity.

The framework addresses the challenges and opportunities in collaborative information sharing among government agencies. Data Governance ensures the integrity, security, and storage of shared information, while the Data Lifecycle component guides agencies through seven stages, embedding data quality considerations at every step. The Promoting Activity component emphasizes collaboration, documentation, and continuous monitoring, crucial for sustained success in interagency collaboration.

This study's practical impact revolves around improving IDS in government via a refined framework. This framework offers practical guidance for navigating complex relationships of data quality across data life cycle phases. By addressing challenges and opportunities in collaborative information sharing, the framework contributes directly to improving the operational processes of government agencies engaged in IDS initiatives.

In terms of theoretical contributions, this research enriches the understanding of information life cycles and data quality in governmental context. The study contributes to

the broader academic discourse on information management by exploring the theoretical underpinnings of data quality considerations in the public sector.

However, it's important to acknowledge some limitations. Since the framework hasn't been widely tested in other IDS practices outside the case study, it needs further enrichment to validate the framework's effectiveness. Incorporating more case studies would enhance its suitability for application in government contexts. Future research endeavors should encompass diverse case studies and integrate recent data sources to ensure a comprehensive grasp of the evolving landscape of IDS in government.

Continued refinement of the proposed framework is essential for future improvements. This could entail gathering feedback from a wider range of government IDS expert, continuously updating the framework to keep pace with advancing technology and evolving information-sharing practices. Expanding the study's scope to include a broader range of datasets and case studies would also contribute to a more robust and widely applicable framework.

Acknowledgement

The authors thank the Ministry of Finance of the Republic of Indonesia and Australia Awards for the Split-Site Master Scholarship Program.

References

[1] L. Zhou, R. Huang, and B. Li, "'What is mine is not thine': Understanding barriers to China's interagency government data sharing from existing literature," *Libr Inf Sci Res*, vol. 42, no. 3, p. 101031, 2020, doi: 10.1016/j.lisr.2020.101031.

[2] S. Curtis and J. Edwards, "Improving public services by sharing the right information," *Public Money and Management*, vol. 39, no. 5, pp. 355–358, 2019, doi: 10.1080/09540962.2019.1611238.

[3] D. S. Sayogo, S. B. C. Yuli, and W. Wiyono, "Challenges and critical factors of interagency information sharing in Indonesia," *Transforming Government: People, Process and Policy*, vol. 14, no. 5, pp. 791–806, 2020, doi: 10.1108/TG-11-2019-0108.

[4] C. Lee, K. Chang, and F. S. Berry, "Testing the development and diffusion of e-government and e-democracy: A global perspective," *Public Adm Rev*, vol. 71, no. 3, pp. 444–454, 2011.

[5] Y. C. Chen, L. T. Hu, K. C. Tseng, W. J. Juang, and C. K. Chang, "Cross-boundary e-government systems: Determinants of performance," *Gov Inf Q*, vol. 36, no. 3, pp. 449–459, 2019, doi: 10.1016/j.giq.2019.02.001.

[6] J. R. Gil-Garcia and D. S. Sayogo, "Government inter-organizational information sharing initiatives: Understanding the main determinants of success," *Gov Inf Q*, vol. 33, no. 3, pp. 572–582, 2016, doi: 10.1016/j.giq.2016.01.006.

[7] S. Wouters, M. Janssen, V. Lember, and J. Crompvoets, "Strategies to advance the dream of integrated digital public service delivery in inter-organizational collaboration networks," *Gov Inf Q*, vol. 40, no. 1, p. 101779, 2023, doi: 10.1016/j.giq.2022.101779.

[8] H. J. Scholl and M. C. Scholl, "Smart governance: A roadmap for research and practice," *IConference 2014 proceedings*, 2014.

[9] N. Noori, T. Hoppe, and M. de Jong, "Classifying pathways for smart city development: Comparing design, governance and implementation in Amsterdam, Barcelona, Dubai, and Abu Dhabi," *Sustainability*, vol. 12, no. 10, p. 4030, 2020.

[10] S. S. Dawes, "Interagency information sharing: Expected benefits, manageable risks," *Journal of policy analysis and management*, vol. 15, no. 3, pp. 377–394, 1996.

[11] H. Kwon, T. A. Pardo, and G. B. Burke, "Interorganizational collaboration and community building for the preservation of state government digital information: Lessons from NDIIPP state partnership initiative," *Gov Inf Q*, vol. 26, no. 1, pp. 186–192, 2009.

[12] J. Fedorowicz, J. L. Gogan, and C. B. Williams, "A collaborative network for first responders: Lessons from the CapWIN case," *Gov Inf Q*, vol. 24, no. 4, pp. 785–807, 2007.

[13] J. R. Gil-Garc\'\ia and T. A. Pardo, "E-government success factors: Mapping practical tools to theoretical foundations," *Gov Inf Q*, vol. 22, no. 2, pp. 187–216, 2005.

[14] T.-M. Yang, L. Zheng, and T. Pardo, "The boundaries of information sharing and integration: a case study in Taiwan e-government," in *Proceedings of the 4th International Conference on Theory and Practice of Electronic Governance*, 2010, pp. 131–136.

[15] L. Zhou, L. Chen, and Y. Han, "'Data stickiness' in interagency government data sharing: a case study," *Journal of Documentation*, vol. 77, no. 6, pp. 1286–1303, Oct. 2021, doi: 10.1108/JD-04-2021-0087.

[16] B. A. Fiedler, "Constructing legal authority to facilitate multi-level interagency health data sharing in the United States," *Int J Pharm Healthc Mark*, vol. 9, no. 2, pp. 175–194, 2015, doi: 10.1108/IJPHM-07-2014-0045.

[17] M. A. Gharawi and H. H. Alneami, "Exploring the Influence of Organizational Context on Cross-boundary Information-Sharing Initiatives: The Case of the Saudi's Government Secure Bus," in *2020 3rd International Conference on Information and Computer Technologies (ICICT)*, 2020, pp. 183–192. doi: 10.1109/ICICT50521.2020.00036.

[18] F. Wang, "Understanding the dynamic mechanism of interagency government data sharing," *Gov Inf Q*, vol. 35, no. 4, pp. 536–546, 2018, doi: https://doi.org/10.1016/j.giq.2018.08.003.

[19] D. M. Walker, V. A. Yeager, J. Lawrence, and A. S. Mcalearney, "Identifying Opportunities to Strengthen the Public Health Informatics Infrastructure: Exploring Hospitals' Challenges with Data Exchange," *Milbank Quarterly*, vol. 99, no. 2, pp. 393–425, 2021, doi: 10.1111/1468-0009.12511.

[20] M. R. Nezami, M. L. C. de Bruijne, M. J. C. M. Hertogh, and H. L. M. Bakker, "Collaboration and Data Sharing in Inter-Organizational Infrastructure Construction Projects," *Sustainability*, vol. 14, no. 24, p. 16835, 2022, doi: https://doi.org/10.3390/su142416835.

[21] C. Schmit, K. Kelly, and J. Bernstein, "Cross Sector Data Sharing: Necessity, Challenge, and Hope," *The Journal of Law, Medicine & Ethics*, vol. 47, no. S2, pp. 83–86, 2019, doi: https://doi.org/10.1177/1073110519857325.

[22] D. Ma, J. Zhou, and M. Zuo, "Inter-agency

information sharing for Chinese e-government development: a comparison between vertical and horizontal dimensions," *Inf Technol Dev*, vol. 28, no. 2, pp. 297–318, Apr. 2022, doi: 10.1080/02681102.2020.1801566.

[23] K. I. A. Khan, R. Flanagan, and S.-L. Lu, "Managing information complexity using system dynamics on construction projects," *Construction management and economics*, vol. 34, no. 3, pp. 192–204, 2016.

[24] S. Titus and J. Bröchner, "Managing information flow in construction supply chains," *Construction innovation*, vol. 5, no. 2, pp. 71–82, 2005.

[25] S. Kurapati, I. Kourounioti, H. Lukosch, L. Tavasszy, and A. Verbraeck, "Fostering sustainable transportation operations through corridor management: A simulation gaming approach," *Sustainability*, vol. 10, no. 2, p. 455, 2018.

[26] I. Makuta and B. O'Hare, "Quality of governance, public spending on health and health status in Sub Saharan Africa: A panel data regression analysis," *BMC Public Health*, vol. 15, no. 1, pp. 1–11, 2015, doi: 10.1186/s12889-015-2287-z.

[27] C. Cichy and S. Rass, "An overview of data quality frameworks," *IEEE Access*, vol. 7, pp. 24634–24648, 2019, doi: 10.1109/ACCESS.2019.2899751.

[28] Y. Bai, S. Li, and W. Zhang, "The Cross-agency Sharing Scheme and Data Quality Evaluation —— A Case study of Geo-spatial big Data," *J Phys Conf Ser*, vol. 1955, no. 1, Jun. 2021, doi: https://doi.org/10.1088/1742-6596/1955/1/012014.

[29] Susan. Earley, Deborah. Henderson, and Data Management Association., *DAMA-DMBOK : data management body of knowledge*.

[30] S. I. H. Shah, V. Peristeras, and I. Magnisalis, "DaLiF: a data lifecycle framework for data-driven governments," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00481-3.

[31] E. Higgins, M. Taylor, P. Lisboa, and F. Arshad, "Developing a data sharing framework: A case study," *Transforming Government: People, Process and Policy*, vol. 8, no. 1, pp. 151–164, 2014, doi: 10.1108/TG-02-2013-0007.

[32] R. Matabean, V. Juwono, and others, "Implementation of Data Collection Policy and Tax Information on Directorate General of Taxes," *Jurnal Borneo Administrator*, vol. 17, no. 3, pp. 365–378, 2021.

[33] Directorate General of Taxes, *Director General Decision No. KEP-215/PJ/2021 regarding Data Governance in the Directorate General of Taxes environment*. 2021.

[34] Directorate General of Taxes, *Circular Letter No SE-12/PJ/2023 regarding Guidelines for the Implementation of Data Governance in the Directorate General of Taxes Environment*. 2023.

[35] Ministry of Finance, *Ministerial Regulation No. PMK-228/PMK.03/2017 concerning the Details of Types of Data and Information as well as Procedures for Submitting Data and Information Related to Taxation*. 2017.

[36] C. Dong, J. Liu, and J. Mi, "How to Enhance Data Sharing in Digital Government Construction: A Tripartite Stochastic Evolutionary Game Approach," *Systems*, vol. 11, no. 4, p. 212, 2023, doi: https://doi.org/10.3390/systems11040212.

[37] M. El Arass and N. Souissi, "Data Lifecycle: From Big Data to SmartData," *Colloquium in Information Science and Technology, CIST*, vol. 2018-Octob, no. 1, pp. 80–87, 2018, doi: 10.1109/CIST.2018.8596547.

[38] A. Sinaeepourfard, J. Krogstie, S. A. Petersen, and D. Ahlers, "F2c2C-DM: A Fog-To-cloudlet-To-Cloud Data Management Architecture in Smart City," *IEEE 5th World Forum on Internet of Things, WF-IoT 2019 - Conference Proceedings*, pp. 590–595, 2019, doi: 10.1109/WF-IoT.2019.8767226.

[39] V. den Eynden, "Data Life Cycle and Data Management Planning," *Looking after and managing your research data*, 2013.

[40] A. Sinaeepourfard, J. Garcia, X. Masip-Bruin, and E. Mar\'\in-Tordera, "A comprehensive scenario agnostic Data LifeCycle model for an efficient data complexity management," in *2016 IEEE 12th International Conference on e-Science (e-Science)*, 2016, pp. 276–281.

[41] W. K. Michener and M. B. Jones, "Ecoinformatics: supporting ecology as a data-intensive science," *Trends Ecol Evol*, vol. 27, no. 2, pp. 85–93, 2012.

[42] S. Orenga-Roglá and R. Chalmeta, "Framework for implementing a big data ecosystem in organizations," *Commun ACM*, vol. 62, no. 1, pp. 58–65, 2018.

[43] M. Ku and J. R. Gil-Garcia, "Ready for data analytics? Data collection and creation in local governments," in *Proceedings of the 19th annual international conference on digital government research: Governance in the data age*, 2018, pp. 1–10.

[44] A. Sinaeepourfard, J. Garcia, X. Masip-Bruin, E. Marin-Tordera, X. Yin, and C. Wang, "A data lifeCycle model for smart cities," in *2016 international conference on information and communication technology convergence (ICTC)*, 2016, pp. 400–405.

[45] A. Sinaeepourfard, J. Garcia, X. Masip-Bruin, and E. Mar\'\in-Torder, "Towards a comprehensive data lifecycle model for big data environments," in *Proceedings of the 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies*, 2016, pp. 100–106.

[46] H. J. Thompson, M. S. Vavilala, and F. P. Rivara, "Common Data Elements and Federal Interagency Traumatic System for TBI Research".

[47] A. Tungkasthan, P. Poompuang, and S. Intarasema, "Government Data Sharing Framework based on DIKW Hierarchy Model," in *2019 17th International Conference on ICT and Knowledge Engineering (ICT&KE)*, 2019, pp. 1–4. doi: 10.1109/ICTKE47035.2019.8966872.

[48] M. Metcalfe, J. Nager, and C. S. Hacker, "Trust Framework for Data Sharing between Industry and Government," in *2023 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, 2023, pp. 1–9. doi: 10.1109/ICNS58246.2023.10124290.

[49] B. Rukanova, Y.-H. Tan, R. Huiden, A. Ravulakollu, A. Grainger, and F. Heijmann, "A framework for voluntary business-government information sharing," *Gov Inf Q*, vol. 37, no. 4, p. 101501, 2020, doi: https://doi.org/10.1016/j.giq.2020.101501.

[50] Republic of Indonesia, *Law Number 7/2021 Regarding Harmonization of Tax Regulations*. 2021.

[51] Y. Prastowo, "New Perspective of Comprehensive Reform: Integrating Corruption Eradication and Tax Optimization Agenda," in *Tax and Corruption Symposium*, 2017.

[52] M. T. Nugroho, "Kepatuhan Wajib Pajak vs Pemeriksaan Pajak: Quo Vadis?" [Online]. Available: https://akurat.co/kepatuhan-wajib-pajak-vs-pemeriksaan-pajak-quo-vadis

[53] Ministry of Finance, *State Revenue and Expenditure Budget 2024*. 2024.

[54] I. Jussen, J. Schweihoff, and F. Möller, "Tensions in Inter-Organizational Data Sharing: Findings from Literature and Practice," in *2023 IEEE 25th Conference on Business Informatics (CBI)*, 2023, pp.

1–10. doi: 10.1109/CBI58679.2023.10187530.

[55] M. A. Madjid, M. Hum, and S. H. Musakkir, "Information-Sharing Pattern in Eradication of International Drugs Network through the Sea in Indonesia," in *2019 IEEE 6th Asian Conference on Defence Technology (ACDT)*, 2019, pp. 19–24. doi: 10.1109/ACDT47198.2019.9072888.

[56] I. Susha, B. Rukanova, A. Zuiderwijk, J. R. Gil-Garcia, and M. Gasco Hernandez, "Achieving voluntary data sharing in cross sector partnerships: Three partnership models," *Information and Organization*, vol. 33, no. 1, p. 100448, 2023, doi: 10.1016/j.infoandorg.2023.100448.

[57] L. Vandercruysse and A. Bourgeus, "Citizen-Centric Personal Data Storage in The Public Sector : An Exploration of Opportunities ," vol. 11, pp. 30–53, 2023.

[58] H. Jones, A. Bandyopadhyay, N. Kennedy, S. Brophy, and J. Evans, "Investigating methods of sharing data between police , health , education , and social services : Semi-structured interviews with police service areas in Wales," vol. 0, no. 0, pp. 1–11, 2023, doi: 10.1177/0032258X231220864.

[59] Q. H. Hamamurad, N. M. Jusoh, and U. Ujang, "Factors That Affect Spatial Data Sharing in Malaysia," 2022.

[60] T. Chen, J. R. Gil-Garcia, G. B. Burke, A. Dey, and D. Werthmuller, "Understanding Cross-Boundary Information Sharing in Emergency Management: Insights from Public Alert and Warning Messages in US Local Governments," in *Proceedings of the 24th Annual International Conference on Digital Government Research*, in DGO '23. New York, NY, USA: Association for Computing Machinery, 2023, pp. 486–495. doi: 10.1145/3598469.3598524.

[61] D. S. Sayogo, J. R. Gil-Garcia, and S. B. C. Yuli, "Determinants of Cross-Boundary Information Sharing Success: Comparing Intra-Agency, Inter-Agency, and Cross-Sectoral Collaboration Initiatives," in *The 21st Annual International Conference on Digital Government Research*, in dg.o '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 138–146. doi: 10.1145/3396956.3396997.

[62] J. R. Gil-Garcia, A. Guler, T. A. Pardo, and G. B. Burke, "Characterizing the importance of clarity of roles and responsibilities in government inter-organizational collaboration and information sharing initiatives," *Gov Inf Q*, vol. 36, no. 4, p. 101393, Oct. 2019, doi: 10.1016/j.giq.2019.101393.

[63] A. J. Holmgren, M. Esdar, J. Hüsers, and J. Coutinho-almeida, "Health Information Exchange : Understanding the Policy Landscape and Future of Data Interoperability," pp. 184–194, 2023.

[64] D. Brand, J. A. Singh, A. G. N. McKay, N. Cengiz, and K. Moodley, "Data sharing governance in sub-Saharan Africa during public health emergencies: Gaps and guidance," *S Afr J Sci*, vol. 118, no. 11/12, pp. 1–6, 2022, doi: https://doi.org/10.17159/sajs.2022/13892.

[65] J. C. Pickering and A. M. Fox, "Enabling Collaboration and Communication Across Law Enforcement Jurisdictions : Data Sharing in a Multiagency Partnership," 2022, doi: 10.1177/08874034211066756.

[66] K. McBride, S. Kamalanathan, S.-M. Valdma, T. Toomere, and M. Freudenthal, "Digital Government Interoperability and Data Exchange Platforms: Insights from a Twenty Country Comparative Study," in *Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance*, in ICEGOV '22. New York, NY, USA: Association for Computing Machinery, 2022, pp. 90–97. doi: 10.1145/3560107.3560123.

[67] A. Y. S. Park, "What advances information sharing for sustainability performance management? Empirical evidence from U.S. local governments," *Urban Governance*, vol. 1, no. 2, pp. 72–80, 2021, doi: https://doi.org/10.1016/j.ugj.2022.01.001.

[68] F. Leoni, M. Carraro, and E. Mcauliffe, "Data-centric public services as potential source of policy knowledge . Can ' design for policy ' help ?," vol. 17, no. 3, pp. 399–411, 2023, doi: 10.1108/TG-06-2022-0088.

[69] S. Y. Tan, A. Taeihagh, and D. Pande, "Data sharing in disruptive technologies: lessons from adoption of autonomous systems in Singapore," *Policy Design and Practice*, vol. 6, no. 1, pp. 57–78, 2023.

[70] H. J. Muhasin, A. Y. Gheni, and H. A. Yousif, "Proposed model for data protection in information systems of government institutions," vol. 11, no. 3, pp. 1715–1722, 2022, doi: 10.11591/eei.v11i3.3727.

[71] S. van Engelenburg, M. Janssen, and B. Klievink, "Design of a software architecture supporting business-to-government information sharing to improve public safety and security," *J Intell Inf Syst*, vol. 52, no. 3, pp. 595–618, Jun. 2019, doi: https://doi.org/10.1007/s10844-017-0478-z.

[72] S. Chignard, "Data collaborations at a local scale : Lessons learnt in Rennes (2010 – 2021)," 2023, doi: 10.1017/dap.2023.16.

[73] P. Naghizadeh and M. Liu, "Using Private and Public Assessments in Security Information Sharing Agreements," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1801–1814, 2020, doi: 10.1109/TIFS.2019.2950125.

[74] P. Thamjaroenporn, "Big Data Analytics Framework for Digital Government," 2020, doi: 10.1109/IBDAP50342.2020.9245461.

[75] A. Karam, T. M. Illemann, K. H. Reinau, G. Vuk, and C. O. Hansen, "Towards Deriving Freight Tra ffi c Measures from Truck Movement Data for State Road Planning : A Proposed System Framework," pp. 1–26, 2020.

[76] Z. Mao, "A CIO-centric local government data-sharing leadership ecosystem in China," vol. 39, no. 4, pp. 798–810, 2023, doi: 10.1177/02666669211054176.

[77] T. M. Illemann, A. Karam, and K. Hegner Reinau, "Towards Sharing Data of Private Freight Companies with Public Policy Makers: A Proposed Framework for Identifying Uses of the Shared Data," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jan. 2021, pp. 132–136. doi: 10.1145/3463858.3463897.

[78] S. Kim and H. Lee, "The impact of organizational context and information technology on employee knowledge-sharing capabilities," *Public Adm Rev*, vol. 66, no. 3, pp. 370–385, 2006.

[79] M. Alavi and D. E. Leidner, "Knowledge management and knowledge management systems: Conceptual foundations and research issues," *MIS quarterly*, pp. 107–136, 2001.

[80] P. H. S. Panahy, F. Sidi, L. S. Affendey, and M. A. Jabar, "The impact of data quality dimensions on business process improvement," *2014 4th World Congress on Information and Communication Technologies, WICT 2014*, pp. 70–73, 2014, doi: 10.1109/WICT.2014.7077304.

[81] R. Y. Wang, "Beyond accuracy: What data quality means to data consumers," *Journal of Management Information Systems*, vol. 12, no. 4, pp. 5–34, 1996, doi: 10.1080/07421222.1996.11518099.

[82] D. Puspasari, A. N. Hadiyanto, and S. Setiawan, "Inter-Organizational Data Sharing: What Issues Should Be Considered?," in *2021 International*

Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS, 2021, pp. 181–186.

[83] F. Karaba, J. K. Roehrich, S. Conway, and J. Turner, "Information sharing in public-private relationships: the role of boundary objects in contracts," *Public Management Review*, pp. 1–25, doi: 10.1080/14719037.2022.2065344.

[84] B. Guo, X. Deng, Q. Guan, J. Tian, and X. Zheng, "An Incentive Mechanism for Cross-Organization Data Sharing Based on Data Competitiveness," *IEEE Access*, vol. 6, pp. 72836–72844, 2018, doi: 10.1109/ACCESS.2018.2882233.

[85] M. Janowicz-Panjaitan and R. Krishnan, "Measures for dealing with competence and integrity violations of interorganizational trust at the corporate and operating levels of organizational hierarchy," *Journal of Management Studies*, vol. 46, no. 2, pp. 245–268, 2009.

[86] J. A. Vayghan, S. M. Garfinkle, C. Walenta, D. C. Healy, and Z. Valentin, "The internal information transformation of IBM," *IBM Systems Journal*, vol. 46, no. 4, pp. 669–683, 2007, doi: 10.1147/sj.464.0669.

[87] G. Costantino, F. Martinelli, I. Matteucci, and M. Petrocchi, "Efficient detection of conflicts in data sharing agreements," in *Information Systems Security and Privacy: Third International Conference, ICISSP 2017, Porto, Portugal, February 19-21, 2017, Revised Selected Papers 3*, 2018, pp. 148–172.

[88] Singapore Infocomm Media Development Authority, *Trusted Data Sharing Framework*. 2019. [Online]. Available: https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf

[89] OECD, *Data Governance: Enhancing Access to and Sharing of Data*. 2021. [Online]. Available: https://www.oecd.org/sti/ieconomy/enhanced-data-access.html.

[90] D. McGilvray, *Executing data quality projects: Ten steps to quality data and trusted information (TM)*. Academic Press, 2021.

[91] Y. Liu, W. Ren, Q. Xu, and Z. Liu, "A coordination mechanism through cost sharing of corporate social responsibility and government subsidy in a two-echelon supply chain," *Journal of Business and Industrial Marketing*, vol. 37, no. 1, pp. 29–46, 2021, doi: 10.1108/JBIM-03-2020-0148.

[92] L. Gardner, J. Ratcliff, E. Dong, and A. Katz, "A need for open public data standards and sharing in light of COVID-19," *Lancet Infect Dis*, vol. 21, no. 4, p. e80, 2021, doi: 10.1016/S1473-3099(20)30635-6.

[93] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *J Manuf Syst*, vol. 48, pp. 157–169, 2018, doi: 10.1016/j.jmsy.2018.01.006.

[94] G. Xie, "Performance evaluation of cross-departmental data sharing in coastal cities based on big data era," *J Coast Res*, vol. 107, no. SI, pp. 215–217, 2020.

[95] S. Muthy, T. Sookram, and B. Gobin-Rahimbux, "Big data analytics life cycle for social networks' posts," in *Information Systems Design and Intelligent Applications: Proceedings of Fifth International Conference INDIA 2018 Volume 2*, 2019, pp. 379–388.

[96] K. Biesialska, X. Franch, and V. Muntés-Mulero, "Big Data analytics in Agile software development: A systematic mapping study," *Inf Softw Technol*, vol. 132, no. October 2020, p. 106448, 2021, doi: 10.1016/j.infsof.2020.106448.

[97] Government of Ireland, *Public service Data Strategy 2019–2023*. 2018. [Online]. Available: https://www.osi.ie/wp-content/uploads/2018/12/Public-Service-Data-Strategy-2019-2023.pdf

[98] H. D. A. dos Santos, M. I. S. Oliveira, G. de F. A. B. Lima, K. M. da Silva, R. I. V. C. Rayelle, and B. F. Lóscio, "Investigations into data published and consumed on the Web: a systematic mapping study," *Journal of the Brazilian Computer Society*, vol. 24, no. 1, 2018, doi: 10.1186/s13173-018-0077-z.

[99] A. S. Panah, A. Yavari, R. Van Schyndel, D. Georgakopoulos, and X. Yi, "Context-driven granular disclosure control for internet of things applications," *IEEE Trans Big Data*, vol. 5, no. 3, pp. 408–422, 2019, doi: 10.1109/TBDATA.2017.2737463.

[100] M. Zambetti, R. Pinto, and G. Pezzotta, "Data lifecycle and technology-based opportunities in new product service system offering towards a multidimensional framework," *Procedia CIRP*, vol. 83, no. March, pp. 163–169, 2019, doi: 10.1016/j.procir.2019.02.135.

[101] H. Elmekki, D. Chiadmi, and H. Lamharhar, "Open government data: Towards a comparison of data lifecycle models," *ACM International Conference Proceeding Series*, 2019, doi: 10.1145/3333165.3333180.

[102] W. Rang, D. Yang, D. Cheng, and Y. Wang, "Data Life Aware Model Updating Strategy for Stream-Based Online Deep Learning," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 10, pp. 2571–2581, 2021, doi: 10.1109/TPDS.2021.3071939.

[103] K. Rahul and R. K. Banyal, "Data Life Cycle Management in Big Data Analytics," *Procedia Comput Sci*, vol. 173, no. 2019, pp. 364–371, 2020, doi: 10.1016/j.procs.2020.06.042.

[104] L. Lin, T. Liu, J. Hu, and J. Zhang, "A privacy-aware cloud service selection method toward data life-cycle," in *2014 20th IEEE international conference on parallel and distributed systems (ICPADS)*, 2014, pp. 752–759.

[105] European Commission, "European Data Strategy." Accessed: Mar. 29, 2024. [Online]. Available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en

[106] A. T.-D. Framework, "Understanding Employees' Willingness to Contribute to Shared Electronic Databases," *Communic Res*, vol. 33, no. 4, pp. 1–20, 2006.

[107] A. Willem and M. Buelens, "Knowledge sharing in public sector organizations: The effect of organizational characteristics on interdepartmental knowledge sharing," *Journal of public administration research and theory*, vol. 17, no. 4, pp. 581–606, 2007.

[108] W. Cho, D. Kim, and A. YS Park, "Local Government's Resource Commitment to Environmental Sustainability: Capacity, Conservatism, and Contractual Dynamics," *Urban Affairs Review*, vol. 59, no. 2, pp. 447–475, 2023.