Data & Policy (2024), 6: e27
doi:10.1017/dap.2024.19

CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Overcoming intergovernmental data sharing challenges with federated learning

Kilian Sprenkamp[1], Joaquín Delgado Fernández[2], Sven Eckhardt[3] and Liudmila Zavolokina[1]

[1]Digital Society Initiative, University of Zurich, Zurich, Switzerland
[2]Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Esch-sur-Alzette, Luxembourg
[3]Information Management Research Group, University of Zurich, Zurich, Switzerland
Corresponding author: Kilian Sprenkamp; Email: kilian.sprenkamp@uzh.ch

Received: 30 June 2023; **Revised:** 09 December 2023; **Accepted:** 20 March 2024

Keywords: artificial intelligence; data sharing challenges; federated learning; eGovernment

Abstract

Intergovernmental collaboration is needed to address global problems. Modern solutions to these problems often include data-driven methods like artificial intelligence (AI), which require large amounts of data to perform well. As AI emerges as a central catalyst in deriving effective solutions for global problems, the infrastructure that supports its data needs becomes crucial. However, data sharing between governments is often constrained due to socio-technical barriers such as concerns over data privacy, data sovereignty issues, and the risks of information misuse. Federated learning (FL) presents a promising solution as a decentralized AI methodology, enabling the use of data from multiple silos without necessitating central aggregation. Instead of sharing raw data, governments can build their own models and just share the model parameters with a central server aggregating all parameters, resulting in a superior overall model. By conducting a structured literature review, we show how major intergovernmental data-sharing challenges listed by the Organisation for Economic Co-operation and Development can be overcome by utilizing FL. Furthermore, we provide a tangible resource implementing FL linked to the Ukrainian refugee crisis that can be utilized by researchers and policymakers alike who want to implement FL in cases where data cannot be shared. Enhanced AI while maintaining privacy through FL thus allows governments to collaboratively address global problems, positively impacting governments and citizens.

Policy Significance Statement

Artificial intelligence (AI) has become indispensable for global problem-solving, but its full potential is curtailed by concerns over sharing raw data, creating obstacles in intergovernmental relations. AI grants governments heightened decision-making prowess, which is essential for strategic governance and international diplomacy. However, balancing this advantage with the imperative to protect citizens' data rights and maintain national sovereignty over data is a geopolitical challenge. Federated learning (FL), a decentralized AI method, has the potential to redefine intergovernmental collaboration by circumventing key data-sharing challenges. FL allows governments to create their own AI models without sharing raw data. By exchanging only model parameters, they ensure data privacy while harnessing collective data-driven insights that surpass individual country capabilities. By aggregating these parameters on a central server, a robust overall model is achieved, enhancing the use of AI for global problem-solving.

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

1. Introduction

Despite approaches to allying with other countries, objectively, sovereign nation-states exercise power over a population of citizens within their territorial borders. With the increasing impact of digital technology and the rise of the internet as a “borderless space,” the role of traditional borders in the digital realm is questioned more and more often. Although the very essence of the internet is to connect users and devices beyond borders, countries attempt to preserve their sovereignty by subjecting cyberspace to their own national rules (Goldsmith, 2007). However, it is becoming increasingly clear that addressing today’s most pressing issues—such as climate change, crisis management, and international supply chain laws—requires a level of transnational cooperation that goes beyond sovereign national interests.

An area where sovereignty is widely pursued among different countries is data sharing. Specifically, when it comes to data storage, governments often rely on technologies such as private data centers secured within their own territories. The 2022 World Economic Forum underlined the problem that data sharing is impractical as data is stored in different legacy system silos (Antonio, 2022). Further, legitimate reasons for making data sharing complicated are disincentives due to the collective action theory (Olson, 1965) and data sharing being unethical (Ward and Sipior, 2010), especially when personal information is involved. In addition, legal uncertainties exist when data is shared between nations created through legislation like the Clarifying Lawful Overseas Use of Data Act (CLOUD Act, 2018). However, sharing data allows enhanced analytics of this data and the use of technological innovation.

Governments need to keep up with waves of technological innovation for economic, social, and political reasons. One such technological innovation is artificial intelligence (AI). In the industry, AI is used in various settings, such as price predictions (Eckhardt et al., 2022). The technological advancement in the industry is often coined as Industry 4.0 (Lasi et al., 2014). Analogously, the term eGovernment 3.0 (eGOV3.0) is used to describe the ever-increasing use of disruptive information and communication technology (ICT), such as AI (Lachana et al., 2018). For example, AI can be used to identify needs in crisis situations (Sprenkamp et al., 2023). While eGovernment 1.0 describes the use of ICT for the realization of public services (Lachana et al., 2018), eGovernment 2.0 focuses on the ICT-enabled participation of citizens. Today, eGOV3.0 uses more advanced and data-driven technologies to solve societal problems through collected data (Lachana et al., 2018).

In order to fulfill the aspiration of AI to solve societal problems, vast amounts of data are needed (Duan et al., 2019). Intergovernmental data sharing has the potential to multiply available data. However, from a global perspective, this need for data immediately creates tension between governments’ interests and incentives. While one government might want to pursue open national data sharing, another might seek to keep their data private. National data, in this context, might refer to aggregated datasets like country-wide health statistics or economic metrics. However, intertwined with these datasets could be finer granular data points that pertain to individual citizens, such as individual health records. This tension is illustrated by the statement of Germany’s former health minister regarding the World Health Organisation’s (WHO’s) potential to place sanctions on countries that do not share their data during disease outbreaks (Wheaton and Martuscelli, 2021). While such a pandemic treaty exists, it is only actively supported by 25 countries (WHO, 2021). This is making the creation of accurate AI models regarding epidemiology complicated. Therefore, the question arises of how to access data without sharing it to solve global problems collaboratively. It can be argued that we can still build AI methods on private data. However, more data results in better-performing AI models. To advance the newly created eGOV3.0, we need to ensure that AI models operate as well as possible, as the performance of these systems has a far-reaching influence on governments and directly on citizens’ lives.

One recent approach that promises to solve these problems is federated learning (FL) (McMahan et al., 2017). The core idea of FL is that individual entities build their own AI models and share them at a centralized point. Another AI model is built that aggregates the individual models. At no point is the data of any individual entity shared. In general, FL shows that the *federated model* performs better than *individual models*, which are built solely on the data of a single entity. However, federated models generally perform worse than the *oracle model*, a model built with all available data. Nonetheless, it is

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

often impossible to build an oracle model due to the limitations of data sharing (Jordan and Mitchell, 2015), leaving the question of which technological approach would be the most fitting.

We propose to use FL to enable better intergovernmental collaborations, from which governments and citizens alike can profit. We, therefore, investigate the research gap regarding how FL can be used for international eGOV3.0 in use cases where data cannot be shared. The following research question is formulated:

RQ *How can federated learning address the problem of data sharing in intergovernmental collaboration?*

To answer this RQ, we investigate how challenges in data sharing listed by the Organisation for Economic Co-operation and Development (OECD) in OECD (2019) can be mitigated through FL. To our knowledge, there is no scientific framework that addresses intergovernmental data-sharing challenges in AI. Yet, the OECD's framework offers the benefit of being both broadly recognized and grounded in practical application. We analyze these challenges through a structured literature review, aiming to propose FL as a solution to the challenges of intergovernmental data sharing. In addition to the structured literature review, we present a showcase to provide an illustrative example of FL in a real-world context. We thus strive to not only contribute to academic discourse but also create a practical, easily comprehensible resource for policymakers. Our goal is to aid them in grasping the implications and potential applications of FL to facilitate the use of AI for intergovernmental use cases where data cannot be shared.

2. Related work and background

2.1. Intergovernmental data sharing

Data sharing is an ever-increasing factor in intergovernmental collaboration and success (Wiseman, 2020). Examples of successfully created AI applications trained on intergovernmental data include health, mobility, and the social sector (Wiseman, 2020). Yet, there are legitimate national, public, and private interests (OECD, 2019). The framework of OECD (2019) lists risks and challenges of data access and sharing and is established in research (e.g., Reimsbach-Kounatze, 2021; Yukhno, 2022; Trampusch, 2023). Moreover, the results of the OECD study are based on workshops held from October 2 to 3, 2017, in Copenhagen. These workshops were attended by industry data professionals from companies like Microsoft or Facebook, policy leaders such as state ministers, and scholars, all from various countries, ensuring a diverse and holistic view (OECD, 2019). Further, we choose to apply the framework from OECD (2019) as we intend not only to make significant contributions to the academic discourse around intergovernmental data-sharing challenges but also to provide pragmatic solutions and support to the public sector. The challenges of data sharing identified by OECD (2019) are grouped into three challenges: (1) balancing the benefits of data openness with legitimate interests, policy objectives, and risks; (2) trust and empowerment for the effective re-use of data across society; and (3) misaligned incentives, and limitations of current business models and markets. These challenges consist of various sub-challenges formulated concerning the growing importance of AI. These challenges and sub-challenges are introduced in the following. Additionally, we extended the OECD framework by incorporating insights and findings from relevant scientific articles to enhance its applicability and comprehensiveness in addressing the challenges of intergovernmental data sharing.

2.1.1. Challenge 1: Balancing the benefits of data openness with legitimate interests, policy objectives, and risks

The sub-challenge (1.1) "Security risks and confidentiality breaches" concerns the integrity and confidentiality of data and information systems, impacting organizations' assets, reputation, and competitiveness. Data breaches, particularly those involving personal data, pose significant harm to individuals' privacy and can result in economic losses, reputational damage, and consumer harm, such as identity theft. While personal data breaches are less frequent than other security incidents, their impact is growing as

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

large-scale breaches become more frequent. Examples of such data breaches are the *Cambridge Analytica* scandal (Isaak and Hanna, 2018) and the disclosure of voter records in the United States (Bennett, 2016).

Moreover, sub-challenge (1.2) "Violation of privacy and intellectual property" is described as violating contractual and socially agreed terms of data reuse poses risks to privacy, intellectual property rights, and commercial interests. These risks can undermine incentives to invest and innovate, particularly for small and medium-sized enterprises. Further, ethical considerations, such as fairness, bias, and discrimination, play a significant role. These factors limit the possibility of exchanging data.

Sub-challenge (1.3) "Difficulty of risk management approaches" presents obstacles to effective risk management practices. Despite its widespread acceptance, organizations face challenges in adopting digital risk management. Relying solely on technological solutions to create secure digital environments can hinder data access and sharing, impeding innovation. Cross-border data flows are restricted by localization requirements, hindering the free flow of information.

The increasing reliance of individuals and organizations on sub-challenge (1.4) "Cross-border data access and sharing" has become essential for global information exchange and economic activities. However, restrictions on transborder data flows can hinder the functioning of data markets and societies, limiting the benefits derived from data sharing and re-use across countries. Concerns arise from data localization requirements that confine data access and sharing within national borders. While these restrictions may be justified by concerns such as privacy, digital security, intellectual property rights, national security, and law enforcement, their proportionality to the risks involved requires careful evaluation. The lack of common approaches and rules for sharing personal and confidential data across countries impedes cross-border data access and sharing.

2.1.2. Challenge 2: Trust and empowerment for the effective re-use of data across society

The sub-challenge (2.1) "Supporting and engaging communities" presents challenges in building trust and facilitating effective data sharing. Communities comprising data users, holders, and third parties play a critical role in defining acceptable risk levels and allocating responsibilities. However, community structures and governance vary based on data openness and potential value. Partnerships between the public and private sectors are crucial, but sustaining community engagement involves costs and reconciling opposing interests.

Further sub-challenge (2.2) "Fostering data-related infrastructures and skills" presents a barrier to effective data reuse, even when data is available. There is a growing demand for data specialist skills, surpassing the available supply in the labor market. Insufficient data-related skills hinder the re-use of data, including open access initiatives. Poor levels of skills and competencies to manage, curate, and re-use data are observed in the scientific community. While in the age of cloud computing, the infrastructure cost of storing, copying, and analyzing data has shrunk, sharing intergovernmental data still involves significant costs for collecting, preparing, sharing, scaling, maintaining, and updating data (Chen and Zhang, 2014; Johnson, 2016).

Another significant barrier is the sub-challenge (2.3) "Lack of common standards for data sharing and re-use" that hinders effective data usage. Inconsistent data formats and incompatible standards impede the creation of longitudinal data sets. Longitudinal data refers to a type of data collected over a period of time from the same individuals, entities, or subjects, allowing one to examine how variables behave over time. The need for data standardization stems from the varying data quality due to inconsistency and incompleteness (Chen and Zhang, 2014; Mikhaylov et al., 2018). Without standardization, it becomes difficult to identify patterns within the data.

Sub-challenge (2.4) "Data quality" greatly influences the accuracy of data analysis and results. High-quality data are essential for effective data analysis and reuse, and a lack of it can deter stakeholders from data-sharing arrangements. Data quality's complexity arises from its dependence on the specific use of data, meaning that the quality of a data set varies. What is good quality for one application may not be for another. The distribution of data among multiple administrations with conflicting bilateral agreements is especially complicated. An example could be the cross-border transfer of data between EU countries,

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Japan, and the United States, which is not currently possible. Currently, the EU only recognizes nine non-EU countries as providing adequate protection for saving data. Japan is among them, but the United States is not (General Data Protection Regulation, 2018). Moreover, the EU can issue fines to any organization not complying with General Data Protection Regulation (2018), creating a further monetary disincentive to share data based on EU law binding to nations inside and outside the EU.

2.1.3. Challenge 3: Misaligned incentives and limitations of current business models and markets

The first sub-challenge is (3.1) "Externalities of data sharing, re-use and the misaligned incentives." Positive externalities describe that third parties may benefit more from data sharing than the data owner, who may struggle to capitalize on data reuse fully. Thus, data holders may be reluctant to share if costs exceed perceived private benefits, resulting in insufficient data access and sharing.

Moreover, the OECD states (3.2) "Limitations of current business models and data markets" is another sub-challenge hindering intergovernmental data sharing. Data markets and platforms play a crucial role in enabling data sharing, yet they face significant challenges that include opaque pricing schemes and limitations in serving social demand for data. Data value often varies greatly among market participants, causing pricing in data markets to be unclear. Lack of transparency may lead to risks of information asymmetry. Information asymmetry can arise between information-poor and information-rich countries and can result in negative consequences for each type of country (Clarkson et al., 2007). Information-poor countries are often in a weaker position to negotiate data-sharing agreements (Clarkson et al., 2007) and are thus inclined to make less advantageous concessions.In contrast, information-rich countries might have a counter-incentive to share their data to maintain their strong economic position. This reluctance to share data emerges from what is known as the "free rider" problem, where data are nonexclusive public goods, and information-rich countries have to accept the risk of information-poor countries utilizing their goods free of charge (OECD, 2019). Due to "free riding" on the goods provided between organizations, allocating public goods becomes ineffective, which is known as the collective action theory (Olson, 1965). Collective action thus results in difficulties for interorganizational cooperation. While Olson (1965) focuses on interorganizational cooperation, the theory has been expanded to problems regarding intergovernmental cooperation, for example, Aspinwall and Greenwood (2013) for cooperation within the European Union (EU) allowing "free riding" of public goods provided by sovereign nation-states.

Sub-challenge (3.3) "The risks of mandatory access to data" include potentially negative impacts on competitive dynamics and investment incentives. While such access is intended to prevent abuse of dominant market positions, it can compromise the competitiveness of entities that rely on data exclusivity. This could reduce their incentives to invest in data or enter certain markets, thus potentially lowering their economic value. Critics warn that this compulsory sharing might inadvertently create anti-competitive effects, possibly giving larger, data-intensive businesses an advantage by facilitating their access to data in various markets.

Last, the OECD defines the sub-challenge (3.4) "Uncertainty about data ownership" as the growing debate around who controls and owns data. This discussion is complicated by the various meanings attached to "data ownership" in different contexts. Legal frameworks, including Intellectual Property Rights and privacy protection laws, can influence the associated rights differently. Moreover, the overlapping nature of these legal regimes, the increasing reliance on contractual agreements, and potential changes in data ownership due to public-private partnerships further complicate the matter. These uncertainties and legal ambiguities can impede intergovernmental data-sharing efforts, as stakeholders may be hesitant to engage in data-sharing activities due to the potential legal risks and lack of clarity in determining data ownership and associated rights (Ward and Sipior, 2010).

These challenges, arising from national, public, and private interests, are especially relevant as they hinder data sharing and, thus, the creation of AI trained on data recorded in multiple countries. This, in turn, impacts the advance of eGOV3.0 and the realization of its benefits to society.

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

2.2. Federated Learning

A promising technology to help address the challenges of data sharing for the creation of AI models was proposed by McMahan et al. (2017), who, while leveraging data utility, maintained a clear separation between data owners. This solution, namely FL, relies on the distribution of models across different databases instead of the classical machine learning (ML) example, where all the data is stored in a single silo. By distributing the models, the model and the data are separated. This keeps the latter isolated and at the data owners' selected location without revealing it. Since then, applications in health (Xu et al., 2021), banking (Yang et al., 2019), or smart cities (Jiang et al., 2020) have been subject to research. FL thus provides incentives in terms of privacy, security, legal, and economic benefits for users (Yang et al., 2019). There are first approaches to frameworks on how to apply FL in eGOV3.0 use cases, which include the specification of client, server, model, and application programming interface requirements at the start of a project (Guberovic et al., 2022). Further, the accountability of FL in government to overcome legislative constraints has been researched (Balta et al., 2021). Balta et al. (2021) both the engineering side, focusing on system design, and the management side, underscoring the crucial role of trust-building among the participating entities.

Originally, McMahan et al. (2017) proposed FL while working at Google, utilising decentralized data stored on multiple edge devices for tasks like image or voice recognition. Now, the original technique is also termed "horizontal FL," as the data of each client shares the same feature space. We visualize FL incorporated into eGOV3.0 in cross-silo (Wiseman, 2020) use cases (see Figure 1). The term cross-silo refers to different data storages on a national or sub-national level, which, due to the challenges of intergovernmental data sharing, explained in Section 2.1, cannot be exchanged between nations. The algorithm is initialized with a base model sent to all clients (i.e., nation-states) by the server (i.e., an intergovernmental organization, intergovernmental collaboration project, or a trusted third party). This step is not part of the repeatedly performed steps and is, therefore, named step 0. In step 1, each client starts the training process from the base model using their own data. In step 2, the difference between the base model parameters and the client model parameters is sent to the server, but each client's data is not shared. While sharing model weights is also a form of information sharing, security techniques like *secure aggregation* and *differential privacy* keep data secure and private. The communication channel between client and server is ciphered via *secure aggregation* (Bonawitz et al., 2017) and thus made secure. *Secure aggegration* thus allows for the transfer of model parameters between parties that do not trust each other, other clients or the server are not capable of obtaining the model weights other clients send through the federated system. During step 3, the server utilizes *federated averaging* to calculate a weighted mean of all differences. The weight of each client is determined by the amount of data used to train the model. Then, in step 4, the server adds random noise to the aggregated model. By adding random noise, privacy is ensured, meaning that the prior steps cannot be reverse-engineered. This procedure is known as *differential privacy* (Agarwal et al., 2018) and is also used in standard ML algorithms. Differentially private models are defined by the tuple (ε,δ), where ε defines the impact of each individual piece of information on the results of the analysis. In other words, low ε indicates a robust system where the outcome is not represented by any particular client data. Additionally, δ regulates the likelihood of a data breach occurring. Step 4 is optional, but if chosen, the model becomes a secure federated model, allowing governments to keep their data private. Finally, in step 5, the private aggregated model is sent to the clients. Steps 1 to 5 represent one round of federated training, which is repeated until the federated model converges.

One of the critical challenges not explicitly highlighted in the aforementioned steps is the communication overhead associated with FL. As each client (in this case, nation-states) trains its own model and communicates the differences back to the central server, the volume of data transmitted can become significant, especially with large models and numerous clients (Niknam et al., 2020). This overhead can lead to increased latencies, bandwidth constraints, and potential costs. We discuss the communication overhead and strategies against it in Section 4 under challenge (2).

Moreover, while FL enhances data governance by training AI models on different servers and ensuring data privacy, it also complies with regulatory requirements. However, we would like to point out that the

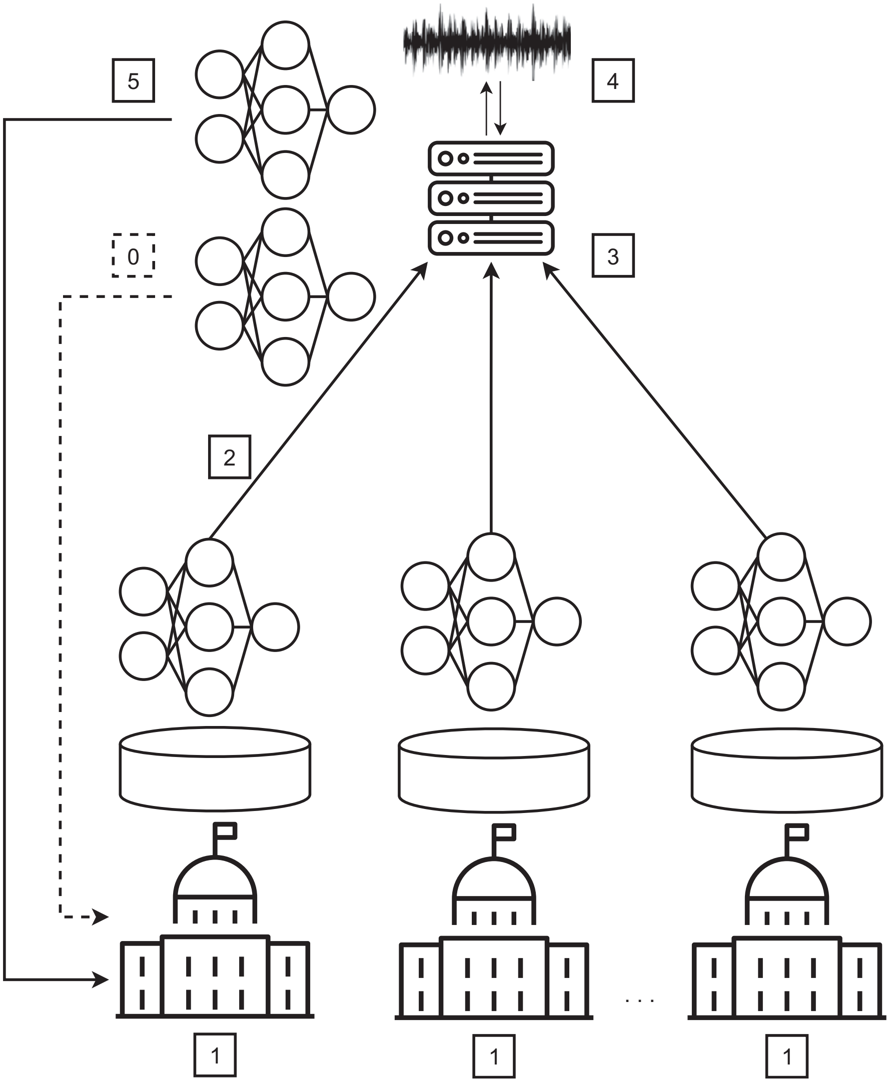

Figure 1. *Federated Learning in eGOV3.0.*

concept is safe only on the technological side. The server, whether an intergovernmental organization or a collaboration project, acts as a central coordinating entity. While it is pivotal in facilitating collaboration, care must be taken to ensure this centralized coordination does not introduce bottlenecks or become a single point of vulnerability in the system. Similarly, the client's data infrastructure is not secured by solely utilizing FL; appropriate data security measures still need to be taken by each individual server outside of the FL training.

While FL has been proposed by McMahan et al. (2017) to solve the problem of data sharing on edge devices, there have been limited attempts to adapt this technique to the public sector. It is especially unclear how the specific challenges of data sharing among governments can be solved.

3. Method

In our research design, we initially conducted a structured literature review, identifying key insights based on practical applicability and specific problems relating to intergovernmental data sharing and the use of

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

FL. Using these insights, we then scrutinize the challenges of data sharing as outlined by OECD (2019). The OECD study offers a diverse and comprehensive view of these matters. Once these challenges and their related literature have been thoroughly reviewed and analyzed, we present a practical showcase designed as a model to explore the relationship between FL and intergovernmental data sharing. The showcase serves as a tangible illustration, shedding light on how FL can be deployed as a viable solution to the challenges inherent in intergovernmental data sharing.

3.1. Literature review

In the previous section, we determined the data-sharing problems identified by OECD (2019) as major challenges. Therefore, we conducted a structured literature review to investigate how FL can solve these data-sharing problems. We utilized the challenges provided by the OECD as a single scientific framework for intergovernmental data sharing does not exist to our knowledge. The challenges of data sharing identified by OECD (2019) are grouped into three categories and various subcategories and formulated with regard to the growing importance of AI. During the literature search process, we focused on information systems and computer science literature, and we applied forward as well as backward searches (Webster and Watson, 2002). We utilized the search string `{“Federated Learning” AND “Data” AND “Challenge*”}` in the disciplines of information systems, computer, decision, manangement, and social science, yielding a total of 879 results. We searched through the SCOPUS, IEEE, and ACM databases. As pre-selection criteria, we analyzed if the given article discusses the challenges named by OECD (2019) based on the abstract; 43 articles were thus chosen for further analysis. We mapped 14 articles to the challenges of the OECD (2019), and a forward and backward search found 16 more articles, resulting in 30 articles being selected.

3.2. Implementation of the showcase

To exemplify the efficacy of FL for intergovernmental collaboration, where data cannot be shared, we provide a showcase based on the Ukrainian refugee crisis. We meticulously collected and curated a comprehensive dataset from Telegram, comprising a total of 669 distinct messages that address the various requirements and challenges faced by Ukrainian migrants residing in multiple European Union countries, including Belgium, the Czech Republic, France, Germany, Poland, Portugal, Spain, and Switzerland. Thus, we provide a tangible resource for policymakers to adapt FL in solutions constrained by intergovernmental data-sharing challenges. The data, model scripts, and results can be found on GitHub and are openly accessible (https://github.com/sprenkamp/federate_learning_data_and_policy). Our objective is to address a supervised multiclass classification challenge, with each Telegram message belonging to five distinct categories related to refugee needs: medical information, accommodation, governmental services, banking, and transportation. The data was annotated by two independent labelers, and only instances that received consistent classifications from both were retained in the final dataset. Regarding the privacy of the data, it is crucial to emphasize that we sourced our information exclusively from open Telegram groups, ensuring all collected messages were in the public domain. Despite their public accessibility, we recognized the inherent ethical considerations associated with such data. Thus, we considered the potential misuse of the given data, especially concerning the vulnerable population it represents. Given these concerns, we have made the data available solely upon request. In Table 1, we show the data distribution over different countries and classes. It is evident that there is a disparity in data distribution across countries, with countries like Germany, France, and Spain having a wealth of data, in contrast to others such as the Czech Republic, which exhibit data scarcity.

For the supervised multiclass classification problem, we fine-tune BERT (Devlin et al., 2018), a transformer-based (Vaswani et al., 2017) pre-trained language model. BERT can capture contextual information in text, enabling the accurate classification of needs into predefined categories. Specifically, we employ a multi-lingual version of BERT obtained from `Huggingface` (https://huggingface.co/bert-base-multilingual-uncased), trained in 102 languages, we are thus able to capture patterns within the

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Table 1. *Data distribution in a number of instances over countries and classes of the migrant Telegram dataset*

	Medical information	Accommodation	Governmental services	Banking	Transportation
Belgium	24	12	9	24	11
Czech Republic	7	11	6	13	4
France	23	13	15	28	24
Germany	19	32	17	30	30
Poland	9	24	9	12	7
Portugal	12	22	6	7	4
Spain	25	30	21	24	32
Switzerland	18	8	20	10	19

various languages used in the Telegram messages (e.g., Ukrainian or Russian). For modeling, we mainly utilized a combination of the `PyTorch` and the `Transformers` packages in `Python`. The dataset was partitioned into training, evaluation, and test sets, adhering to a ratio of 60% for training, 20% for evaluation, and 20% for testing. For assessing the model on the test set, we used accuracy, precision, recall, and the F1 score, which are standard metrics for text classification tasks. We calculate these metrics for the oracle, individual, federated and secure federated models. For the secure federated model, we utilize *differential privacy* as explained in section 2.2. There, following the original work of McMahan et al. (2017) and the implementation principles of Yang et al. (2021b). For the secure and non-secure FL model, we found that four rounds of communication were sufficient to achieve accurate learning representations. Notably, there are differences between the sampled countries for the secure and non-secure FL models. On the one hand, the non-secure FL model uses information from all countries in each communication round since the more data available, the more powerful the model will be. On the other hand, for the private model, we performed several simulations to find an optimal value of countries per round. To maximize the privacy of the model, we optimized this ratio to 0.5. This means that we randomly sampled half of the countries where more clients mean more performance, but also more noise is needed to hide individual contributions. As mentioned above, there is a clear trade-off between noise and accuracy: the more noise we input, the more private the model will be, but at the cost of performance. Given a sampling ratio of 0.5, we set the noise standard deviation to 1.0. As in Yang et al. (2021b), the central server adds the noise to the model weights after averaging, later, to evaluate the overall privacy of the model, we used Renyi-Differential Privacy (Mironov, 2017), which yields higher privacy than the original moment accountant from McMahan et al. (2017).

4. Federated learning for OECD challenges

We identified solutions to the sub-challenges (OECD, 2019) by conducting a structured literature review. We found solutions for eight of the 12 sub-challenges defined, while two of the sub-challenges were partially solved, and two of the sub-challenges remained unsolved (see Table 2).

4.1. Challenge 1: Balancing the benefits of data openness with legitimate interests, policy objectives, and risks

With FL, the security and confidentiality breaches in data sharing (1.1) can be avoided, as FL alleviates the need to share data. FL does so through a range of security and privacy techniques (Mothukuri et al., 2021). We want to point out two core approaches *differential privacy* (Agarwal et al., 2018) and *secure aggregation* (Bonawitz et al., 2017). While *differential privacy* secures the system from being reverse-

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Table 2. FL solutions to data sharing challenges (*OECD,* 2019)

Challenge	Sub challenge	Solved by FL	Proposed solution
(1) Balancing the benefits of data openness with legitimate interests, policy objectives and risks	(1.1) Security risks and confidentiality breaches	Solved	• Security and privacy techniques (Agarwal et al., 2018; Benmalek et al., 2022; Bonawitz et al., 2017; Cao et al., 2019; Y. Chen et al., 2017; Mothukuri et al., 2021) • Hierarchical federated learning (Abad et al., 2020)
	(1.2) Violation of privacy and intellectual property	Solved	• Workaround for contractual agreements (Li et al., 2020a, 2021)
	(1.3) Difficulty of risk management approaches	Not solved	• N/A
	(1.4) Cross–border data access and sharing	Solved	• Unnecessity of cross–border data sharing (Truong et al., 2021; D. Yang et al., 2021a, 2019)
(2) Trust and empowerment for the effective re–use of data across society	(2.1) Supporting and engaging communities	not solved	• N/A
	(2.2) Fostering data–related infrastructures and skill	Partially solved	• Communication cost (McMahan et al., 2017; Li et al., 2020b; Ozfatura et al., 2021; Shah and Lau, 2021) • FL toolkits (Ziller et al., 2021)
	(2.3) Lack of common standards for data sharing and re–use	Solved	• System heterogeneity (Mitra et al., 2021) • Vertical federated learning (Yang et al., 2019) • Transfer learning (Chen et al., 2020) • Meta learning (Fallah et al., 2020)

(Continued)

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Table 2. *Continued*

Challenge	Sub challenge	Solved by FL	Proposed solution
	(2.4) Data quality	Solved	• Data augmentation (de Luca et al., 2022) • FL on noisy data (Passerat–Palmbach et al., 2020; Tuor et al., 2021)
(3) Misaligned incentives and limitations of current business models and markets	(3.1) Externalities of data sharing, re–use, and misaligned incentives	Solved	• Incentives of FL (Kang et al., 2019; Yu et al., 2020)
	(3.2) Limitations of current business models and data markets	Solved	• FL as a business model (Yang et al., 2019; Balta et al., 2021; Manoj et al., 2022)
	(3.3) The risks of mandatory access to data	Partially solved	• N/A
	(3.4) Uncertainties about data ownership	Solved	• Data ownership is explicit (Liu et al., 2020; Shae and Tsai, 2018) • Model ownership is explicit on technical level (Liu et al., 2021)

engineered, *secure aggregation* secures the system by ciphering the communication channel between the clients and the server. Moreover, concerning the importance of personal information, hierarchical FL settings (Abad et al., 2020) enable FL to be applied on multiple levels. A country could thus give citizens or companies control over their data while still profiting from AI being trained by international projects. In the context of safeguarding FL systems, it is essential to address potential attacks on the federated system, such as poisoning attacks, inference attacks, communication attacks, and free-riding attacks (Benmalek et al., 2022). These malicious activities can disrupt the collaborative learning process by injecting corrupted or manipulated data into the system. Implementing robust security measures such as *differential privacy* together with anomaly detection mechanisms is crucial to defend against these threats and ensure the integrity and reliability of FL (Cao et al., 2019; Chen et al., 2017).

The violation of privacy and property rights (1.2) is, according to OECD (2019), based on contractual agreements. The violation of these agreements can lead to fines. Moreover, sharing data prematurely can reduce the chance of creating intellectual property. FL offers a technological pathway for entities to comply with these contractual agreements (Li et al., 2020a, 2021), thus preventing them from violating contractual clauses and being exposed to ensuing fines or the premature revelation of intellectual property.

Regarding mitigating the difficulty of applying risk management approaches (1.3), we currently see limited potential in FL to solve this challenge.

From a legal perspective, cross-border data sharing (1.4) can be complicated due to regulations like the General Data Protection Regulation (2018)or the CLOUD Act 2018. FL allows the training of AI without the need to share data across borders. Q. Yang et al. (2019) proposed training federated models between Chinese and American companies. Similarly, Yang et al. (2021a) presented an FL system using data from China, Japan, and Italy to predict SARS-CoV-2 from chest computed tomography images. However, Truong et al. (2021) point out that due to the exchange of model weights and the resulting threat of backward engineering, FL, without any security and privacy techniques, does not conform with General Data Protection Regulation (2018) in Europe. Therefore, FL can only be utilized with privacy and security preserving techniques when utilising data from multiple countries.

4.1. Challenge 2: Trust and empowerment for the effective re-use of data across society.

FL cannot help to create more engagement in open data communities (2.1) as it reduces the need to share data. However, we can see that the shared model training creates a community aspect. To our knowledge, this has not yet been researched.

We found that FL cannot solve problems related to data infrastructure or skills (2.2). Generally, the technique requires a more complicated setup than standard ML processes (T. Li et al., 2020b). FL profits from dividing the computational cost across multiple clients. Yet, the cost of communication is high in FL, as clients need to communicate continuously. This cost is practically non-existent in standard ML (McMahan et al., 2017). Reducing costs associated with FL is currently under research, and several solutions have been proposed, including model compression (Shah and Lau, 2021), parameter sparsification (Ozfatura et al., 2021), or structured and sketched updates (Li et al., 2020b). Model compression aims to reduce the size of the model by simplifying its architecture or the representation of its parameters, thus making it less resource-intensive to store and transmit. Techniques such as pruning, where unnecessary or less important connections within a neural network are eliminated, and quantization, which reduces the precision of the numerical parameters, are commonly applied (Shah and Lau, 2021). Parameter sparsification selectively transmits only the most significant model updates, ignoring minor changes that have minimal impact on the overall model performance. By setting a threshold for the importance of updates, only parameters that exhibit substantial changes are sent to the server for aggregation, thus reducing the absolute amount of processed model parameters within the network (Ozfatura et al., 2021). Structured updates impose a predefined structure on the model updates before they are transmitted, which could involve techniques like low-rank approximations that maintain the essential information while reducing the size of the data package. Sketched updates, on the other hand, use algorithms to create a compressed summary of the updates. These summaries or “sketches” provide an approximate but significantly smaller

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

representation of the updates, further reducing the communication load without severely compromising the model's accuracy (Li et al., 2020b). Further, due to the increased complexity and novelty of FL, tool kits like `TensorFlow Federated` based on McMahan et al. (2017) and `PySyft` (Ziller et al., 2021) have not seen wide adoption compared to other ML frameworks. Nonetheless, we see possibilities for less skilled countries to profit as they prefer to use federated models rather than training models themselves.

The lack of data standardization (2.3) reflects two core challenges of FL: system and statistical heterogeneity. Solutions to this issue exist already. System heterogeneity is described as different hardware used among clients, leading to a slower training process. For example, Mitra et al. (2021) propose re-using parts of the model during the training process, such as gradients of the learned network or the specification of concrete learning rates for individual clients depending on the hardware. Statistical heterogeneity refers to different data features being stored or features having non-identical distributions across clients. In this case, vertical FL (Yang et al., 2019) can be used, which allows for the training of models with different feature spaces. Moreover, it is possible to apply transfer FL, meaning that a model is retrained for a different learning task, benefiting from the knowledge of the previously learned task. Chen et al. (2020) employ this technique by first training a model for activity recognition for smartwatches, which is then transferred to the task of predicting Parkinson's disease. Other techniques to address statistical heterogeneity include meta-learning or data augmentation. Within FL, meta-learning allows individual client models to benefit from broader multi-task knowledge, enabling swift adaptation to local data variations (Fallah et al., 2020). Data augmentation within the FL can be locally applied at each client, introducing data variations like rotations, which helps harmonize the diverse data distributions across clients (de Luca et al., 2022).

The OECD notes that poor data quality (2.4) will lead to poor analytics. While FL cannot improve the data quality of clients, entities with poor data quality can profit from the federated model. Examples can be found within the medical field of genomics or mental health, where large amounts of noisy data can be found (Passerat-Palmbach et al., 2020). In this case, clients with poor data can profit from the federated model and the contribution made by other clients with better-quality data. Still, the clients with poor data quality will decrease the overall model quality. A solution to this challenge is proposed by Tuor et al. (2021). Each client evaluates their data set with a benchmark model trained on high-quality data. For bad-quality data, the model will be incapable of making a prediction, generating a high loss value. These data points will not be further utilized for training.

4.3. Challenge 3: Misaligned incentives and limitations of current business models and markets

A central problem within data sharing is misaligned incentives (3.1) between information-rich and information-poor countries. The problem of incentivizing information-rich clients to participate in FL has been well-researched (Kang et al., 2019; Yu et al., 2020). These methods typically offer a reward for participating within the federated system. Hence, an entity could earn depending on how much value was brought to the federated model.

Implementing FL could significantly reduce the need for data markets (3.2), as data can be kept by the owner. Moreover, according to the OECD, the ex-ante evaluation of the economic potential of data is challenging. However, given the previously shown incentive schemes of FL, it is possible to track participation in an FL project. Yang et al. (2019) estimate that FL will evolve into a business model where participants in an FL project can profit from the value they contribute to the model. FL thus allows participants to pursue joint business activities (Balta et al., 2021). An example of such joint business activity is given by Manoj et al. (2022), training a model for predicting the yield of crops. This model can be utilized by multiple stakeholders, for example, farmers for revenue estimates, banks and insurance for mitigating risks, and governments for setting export prices.

In a federated setting, mandatory data access (3.3) can be kept to a minimum. For example, the `PySyft` package (Ziller et al., 2021) within Python allows for viewing a limited number of samples of each client's data to optimize the federated model. Accessing all available data points is, in theory, not necessary, and due to the number of data points available, not always feasible while training AI models.

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Finally, the OECD notes a loss of data ownership (3.4) as an emerging challenge of sharing data. With FL, the ownership of a data point remains unaffected as it is not shared across multiple sources. For example, Shae and Tsai (2018) propose storing medical information on a blockchain for training federated models. Thus, the ownership of data cannot be falsified. Similarly, Liu et al. (2020) proposed a traffic flow prediction model utilizing data from government organizations, smart devices, and private persons, as well as private companies such as Uber or Didi. For each of these entities, the data ownership is unambiguous. Moreover, it is possible to verify the ownership over a trained federated model by implementing a watermarking technique. Thus, the contribution and resulting ownership of clients is recorded through the watermark (Liu et al., 2021). However, the watermarking technique solely clarifies ownership on a technical and not legal level. To our knowledge, the legal ownership of federated models remains unsolved.

5. Showcase

Based on the results found in the literature, we observe FL as a solution to the problem of data sharing in intergovernmental collaboration in a real-world practical setting. While in this showcase, we have all available data, one could easily think of a similar case where not all data is available to one central authority, such as documents from multiple government organisations. The total comparison of all four models can be viewed in Table 3 for the given error metrics.

Undeniably, the oracle model, with access to all data, exhibited superior performance. However, it also underscores a crucial drawback. Often, in real-world situations, sharing of all data is not feasible, considering data privacy and legal restrictions. Therefore, the practical utility of the oracle model is significantly limited. The models based on FL offer a more applicable alternative. Even though these models do not have access to a central data repository, their performances are better than the individual model. Both non-secure FL and Secure FL models exploit the power of collective intelligence without violating data privacy. In essence, despite the inevitable trade-off between model accuracy and privacy, FL-based models, secure and non-secure alike, significantly outperform individual models and provide a practical solution for contexts where data sharing is impractical. Further, we would like to illustrate which sub-challenges by OECD (2019) are solved through our showcase.

For challenge (1) “Balancing the benefits of data openness with legitimate interests, policy objectives and risks,” our showcase bypasses the need for raw data sharing, thus significantly reducing the risk of security and confidentiality breaches (1.1) and further the violation of privacy rights (1.2). The secure FL model, which uses differential privacy, stands out as an example of how data privacy can be ensured even while learning from collective intelligence and outperforming models trained by a single client. We thus can maintain privacy among the clients, from which end users, that is, refugees within the EU, profit.

Regarding challenge (2) “Trust and empowerment for the effective re-use of data across society,” our showcase exemplifies that entities with poor data quality or small amounts of data can profit from the federated model (2.4). In Table 4, we show the accuracy per country for the individual and FL models. We are able to observe that countries where data is scarce, for example, Czech Republic, profit most from the federated models. However, we see an overall increase in accuracy in countries when comparing FL and

Table 3. *Error metrics of the different models.*

	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Oracle	97.01	96.67	96.92	96.57
Individual	51.03	33.08	42.08	35.00
Non–secure FL	89.71	86.51	86.12	85.32
Secure FL	87.93	83.07	82.90	82.30

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Table 4. *Comparison of model accuracy per country*

	Accuracy individual (%)	Accuracy non-secure FL (%)	Accuracy secure FL (%)
Belgium	37.50	100.00	100.00
Czech Republic	20.00	100.00	100.00
France	63.63	90.91	90.91
Germany	61.53	92.31	92.31
Poland	42.85	100.00	75.00
Portugal	66.66	66.66	66.66
Spain	78.57	92.86	92.86
Switzerland	37.50	75.00	75.00

individual models, except for Portugal. We thus can predict refugee needs better by utilizing FL compared to each country by doing a similar analysis alone.

Concerning challenge (3) “Misaligned incentives, and limitations of current business models and markets,” our showcase exemplifies that FL helps to overcome issues regarding the reuse and ownership of data (3.1), (3.2). As within our training process, the data is not known to the other clients, the reuse of data can be efficiently managed while data ownership remains intact.

6. Discussion

This study aimed to address the research gap regarding how FL can be used in international eGOV3.0 use cases where data sharing is complicated or unfeasible. Since McMahan et al. (2017) first proposed FL, a vast number of publications have appeared in the field, including applications in health, banking, and smart cities. However, research in the area of eGOV3.0 is limited. While Guberovic et al. (2022) created a framework specifying component requirements for government FL projects and Balta et al. (2021) analyzed the accountability of FL in government, we analyzed how FL can solve the problem of intergovernmental data sharing. We did so by conducting a structured literature review, which served the purpose of analyzing how FL can solve challenges identified by OECD (2019). In doing so, we were able to answer the given RQ: *How can federated learning address the problem of data sharing in intergovernmental collaboration?*

First, FL can help to deal with legal (Ward and Sipior, 2010) and ethical (Pingitore et al., 2017; Isaak and Hanna, 2018) issues around data sharing. We provided evidence regarding how the sub-challenges (1.1), (1.2), (1.4) and (3.4) identified by the OECD can be solved. FL significantly reduces the need for data sharing agreements to build AI (Li et al., 2020a, 2021), which also applies to cross-border data sharing (Truong et al., 2021; Yang et al., 2019, 2021a). A further consequence is that data ownership cannot be falsified, as the data is stored at the owners’ selected location. Moreover, FL, mainly through *differential privacy* (Agarwal et al., 2018) and *secure aggregation* (Bonawitz et al., 2017), allows for secure model training and keeping data private. FL, thus, provides a trusted technology that is ideal for intergovernmental use cases.

Second, we showed how technological constraints in data sharing (Chen and Zhao, 2012, 2014; Johnson, 2016; Mikhaylov et al., 2018) can be overcome using FL, especially issues regarding the sub-challenges of standardisation of data (2.3) and data quality (2.4). Mitra et al. (2021) show methods that allow federated models to be trained in system heterogeneous settings. This is beneficial for FL in intergovernmental settings as countries, companies, and citizens can partake in FL projects without the need for special hardware. Vertical FL (Yang et al., 2019) and transfer FL (Chen et al., 2020) allow the training of AI models on non-standardized data sets, and they can even leverage data that was not recorded for the task they have been employed to solve. Intergovernmental collaboration can thus profit from data stored by all types of entities and further re-use data effectively. Additionally, data from both information-

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

poor and information-rich countries can be utilized to contribute to FL projects. Information-poor countries are not limited to their own data sources anymore and can contribute and profit from the federated model. Still, while federated learning shows potential and has been implemented on a scientific basis till now, applications of FL in an intergovernmental context in real-world scenarios are not known to us. This could be the case due to technological challenges solely being solved in the literature but not in real-world scenarios.

Third, FL can evolve into a business model (Yang et al., 2019; Balta et al., 2021), which gives entities an incentive to take part in intergovernmental projects. This mitigates disincentives in data sharing caused by "free riding" and the problem of inefficiency due to collective action (Olson, 1965). Consequently, we demonstrated how sub-challenges (3.1) and (3.2) could be solved through FL. With incentive mechanisms developed for FL (Kang et al., 2019; Yu et al., 2020), both information-rich and information-poor entities can be motivated to participate in an FL project by offering a reward. For example, an intergovernmental climate model could be possible where different stakeholders, for example, countries, companies, or single individuals, could earn money or CO2 credits based on the revenue or value that an intergovernmental FL project generates. Still, large contributors would earn the most, but less funded entities can profit fairly. We consider the incentives to partake in FL to be superior to the incentives for partaking in data sharing. When using FL, the ownership of the data (3.4) remains intact for each FL project an entity might participate in. In contrast, for standard data sharing, as soon as data is distributed, it becomes unclear who the real owner is. Therefore, "free riding" as described in the collective action theory (Olson, 1965) and the resulting inefficiency can be mitigated on a technical level in theory.

Lastly, the showcase stresses the possibilities of federated learning as a solution to problems of data sharing in intergovernmental collaborations. In our use case, we had all the data available, which was the only way to evaluate the technical performance. However, one can easily think of similar use cases where not all data is available to one central authority. Concerning refugee needs, instead of public Telegram channels, one could collect data from interactions on government sites, which governments could be reluctant to share with each other. Our showcase tackles the challenges outlined by the OECD (2019), specifically, the need to bypass raw data sharing to reduce security breaches (1.1) and privacy violations (1.2). Through the implementation of secure FL, which prioritizes privacy using techniques like differential privacy, we ensure data privacy while benefiting from collective intelligence. FL also addresses the challenge of trust and empowerment (2) by benefiting entities with limited data quality (2.4). The accuracy results per country in Table 4 indicate improved performance with federated models. Additionally, FL resolves issues related to misaligned incentives and current business models (3), overcoming obstacles concerning data reuse and ownership (3.1) and (3.2). FL enables efficient data reuse while preserving ownership, as the training process does not expose data to other clients.

However, not all problems related to data sharing can be solved through FL. Sub-challenges (2.2) and (3.3) are only partially solved, while sub-challenges (1.3) and (2.1) are not solved. There are two key problems. First, implementing FL infrastructure is more complicated, with higher communication costs. Even with a larger adoption, this will cause challenges in eGOV3.0. Second, although ownership of FL is actively researched from a technological point of view (Liu et al., 2021), we see problems on the organizational level, in which the geolocation and the controlling entity of the server aggregating the model will play a central role. The entity controlling the federated model presents a potential bottleneck as it could cut off countries in an intergovernmental collaboration project without a democratic process. Especially from a political realism perspective, it is unlikely that nations that do not trust each other will participate in a joint FL project. Future research could consider governments' willingness to participate in projects that benefit various multinational stakeholders. Within this analysis, the difference between incentives of information-rich and information-poor countries is likely to play a key role.

7. Conclusion

This study conducts a structured literature review to show how FL can function as a solution to various challenges related to intergovernmental data sharing. FL enables the training of models in a decentralized manner and can thus reduce incentive, legal, ethical, and practical challenges in intergovernmental data sharing. Nevertheless, secondary problems of a technical and organizational nature arise. The contribution of this study is threefold. First, we present a state-of-the-art AI method to overcome the problem of intergovernmental data sharing. This serves as a basis for FL research in international eGOV3.0, which we hope will influence both governments and citizens. Second, we contribute to the existing literature on FL, providing a structured review and a showcase of how FL should be utilized in eGOV3.0, focusing on the aspect of data sharing. We hope this will enhance the research output of real-life use cases in and outside the eGOV3.0 space. Third, we show a new possibility of how to mitigate inefficiency created by the collective action theory (Olson, 1965).

Our study presents two primary limitations. First, utilizing the OECD's framework, though it provides a practical and widely accepted basis for our analysis, it also confines us to a specific perspective. We estimate that challenges described by other authorities will be similar, but adding challenges from authorities of different cultural or economic origins would create an even more holistic and diverse picture. Furthermore, the showcased data set neither includes private data from various countries nor is extensive, containing only 669 instances. Moreover, our evaluation focused solely on the performance of the models without considering the communication overhead inherent in FL.

Considering the opportunities and challenges highlighted in this study, we find substantial potential for the information systems and related research communities. We suggest creating federated systems on proprietary, potentially unbalanced data from multi-governmental stakeholders. Dedicated qualitative research could be done, drawing insights from workshops or stakeholder interviews to further investigate the potential of FL in eGOV3.0. Moreover, at the organizational level, it is necessary to determine the owner of the server that creates the aggregated model. Finally, the higher infrastructure costs and skill levels of users in FL need to be considered in further research.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/dap.2024.19.

Data availability statement. To facilitate replication of our experiments, all necessary resources are readily accessible via our GitHub repository: https://github.com/sprenkamp/federate_learning_data_and_policy.

Acknowledgments. We would like to thank Linda Weigl for her insightful comments on the first draft of this work. Moreover, we would like to acknowledge that this manuscript is a continuation of our study presented at the 56th Hawaii International Conference on System Sciences (Sprenkamp et al., 2023). We would like to express our gratitude to the anonymous reviewers whose constructive feedback greatly enhanced the quality of this article. We also express our appreciation to the organizers of the 56th Hawaii International Conference on System Sciences, whose efforts provided us with a profoundly valuable and enriching experience.

Author contribution. Conceptualization: K.S., J.D.F., S.E., L.Z.; data curation: K.S., J.D.F., S.E.; formal analysis: K.S., S.E., L.Z.; funding acquisition: J.D.F., L.Z.; investigation: K.S., J.D.F., S.E., L.Z.; methodology: K.S., J.D.F., S.E., L.Z.; project administration: K.S., L.Z.; resources: K.S., J.D.F., L.Z.; software: K.S., J.D.F.; supervision: L.Z.; validation: K.S., J.D.F., S.E., L.Z.; visualization: K.S., J.D.F., S.E.; writing—original draft: K.S., J.D.F., S.E., L.Z.; writing—review & editing: K.S., J.D.F., S.E., L.Z.

Funding statement. We thank the University of Zurich and the Digital Society Initiative for (partially) funding this study under the Digitalization Initiative of the Zurich Higher Education Institutions postdoc fellowship of L.Z. Further, this work has been supported by the European Union (EU) within its Horizon 2020 program, project MDOT (Medical Device Obligations Taskforce), grant agreement 814,654, and from PayPal and the Luxembourg National Research Fund FNR (P17/IS/13342933/PayPal-FNR/Chair in DFS/Gilbert Fridgen).

Competing interest. The authors have no competing interests to declare.

Ethical standard. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

Abad MSH, **Ozfatura E**, **Gunduz D and Ercetin O** (2020) Hierarchical federated learning ACROSS heterogeneous cellular networks. In *ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain*, pp. 8866–8870.

Agarwal N, **Suresh AT**, **Yu FXX**, **Kumar S and McMahan B** (2018) cpSGD: Communication-efficient and differentially-private distributed SGD. *Advances in Neural Information Processing Systems*.

Antonio N (2022) The Public Sector Must Accelerate Digital Transformation – Or Risk Losing Sovereignty and Trust. Available at https://www.weforum.org/agenda/2022/05/the-public-sector-must-accelerate-digital-transformation-or-risk-losing-sovereignty-and-trust/ (accessed 23 May 23 2022).

Aspinwall M and Greenwood J (2013) *Collective action in the European Union: Interests and the New Politics of Associability.* London: Routledge.

Balta D, **Sellami M**, **Kuhn P**, **Schöpp U**, **Buchinger M**, **Baracaldo N**, **Anwar A**, **Ludwig H**, **Sinn M**, **Purcell M and Altakrouri B** (2021) Accountable federated machine learning in government: engineering and management insights. In *Electronic Participation: 13th IFIP WG 8.5 International Conference*, *ePart 2021, Granada, Spain*, *September 7–9, 2021, Proceedings 13*, pp. 125–138. Springer International Publishing.

Benmalek M, **Benrekia MA and Challal Y** (2022) Security of federated learning: Attacks, defensive mechanisms, and challenges. *Revue des Sciences et Technologies de l'Information-Série RIA: Revue d'Intelligence Artificielle 36*(1), 49–59.

Bennett CJ (2016) Voter databases, micro-targeting, and data protection law: Can political parties campaign in Europe as they do in North America? *International Data Privacy Law 6*(4), 261–275.

Bonawitz K, **Ivanov V**, **Kreuter B**, **Marcedone A**, **McMahan HB**, **Patel S**, **Ramage D**, **Segal A and Seth K** (2017) Practical secure aggregation for privacy-preserving machine learning. In *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1175–1191.

Cao D, **Chang S**, **Lin Z**, **Liu G and Sun D** (2019) Understanding distributed poisoning attack in federated learning. In *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 233–239. Available at https://api.semanticscholar.org/CorpusID:210992177.

Chen CP and Zhang C-Y (2014) Data-intensive applications, challenges, techniques and technologies: A survey on big data. *Information Sciences 275*(10), 314–347.

Chen D and Zhao H (2012) Data security and privacy protection issues in cloud computing. In *2012 International Conference on Computer Science and Electronics Engineering*.

Chen Y, **Qin X**, **Wang J**, **Yu C and Gao W** (2020) Fedhealth: A federated transfer learning framework for wearable healthcare. *IEEE Intelligent Systems 35*(4), 83–89.

Chen Y, **Su L and Xu J** (2017) Distributed statistical machine learning in adversarial settings: Byzantine gradient descent. *Proceedings of the ACM on Measurement and Analysis of Computing Systems 1*(2), 1–25.

Clarifying Lawful Overseas Use of Data Act (CLOUD Act). (2018) *115th Congress of the United States of America.* Available at https://www.congress.gov/bill/115th-congress/house-bill/4943/text.

Clarkson G, **Jacobsen TE and Batcheller AL** (2007) Information asymmetry and information sharing. *Government Information Quarterly*.

de Luca AB, **Zhang G**, **Chen X and Yu Y** (2022) Mitigating data heterogeneity in federated learning with data augmentation. *arXiv preprint arXiv:2206.09979*.

Devlin J, **Chang M-W**, **Lee K and Toutanova K** (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Duan Y, **Edwards JS and Dwivedi YK** (2019) Artificial intelligence for decision making in the era of big data–evolution, challenges and research agenda. *International Journal of Information Management48*, 63–71.

Eckhardt S, **Sprenkamp K**, **Zavolokina L**, **Bauer I and Schwabe G** (2022) Can artificial intelligence help used-car dealers survive in a data-driven used-car market? In *International Conference on Design Science Research in Information Systems and Technology*, pp. 115–127.

Fallah A, **Mokhtari A and Ozdaglar A** (2020) Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach. *Advances in Neural Information Processing Systems 33*, 3557–3568.

General Data Protection Regulation (2018, May 25). *European Commission.* Available at https://gdpr.eu/.

Goldsmith J (2007) Who controls the internet? illusions of a borderless world. *Strategic Direction 23*(11).

Guberovic E, **Alexopoulos C**, **Bosnić I and Čavrak I** (2022) Framework for federated learning open models in e-government applications. *Interdisciplinary Description of Complex Systems 20*, 162–178.

Isaak J and Hanna MJ (2018) User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer 51*(8), 56–59.

Jiang JC, **Kantarci B**, **Oktug S and Soyata T** (2020) Federated learning in smart city sensing: Challenges and opportunities. *Sensors 20*(21), 6230.

Johnson PA (2016) Reflecting on the success of open data: How municipal government evaluates their open data programs. . *International Journal of E-Planning Research5*(3).

Jordan MI and Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. *Science 349*(6245), 255–260.

Kang J, **Xiong Z**, **Niyato D**, **Xie S and Zhang J** (2019) Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory. *IEEE Internet of Things Journal 6*(6), 10700–10714.

Lachana Z, **Alexopoulos C**, **Loukis E and Charalabidis Y** (2018) Identifying the different generations of eGovernment: An analysis framework. In *12th Mediterranean Conference on Information Systems*.

Lasi H, **Fettke P**, **Kemper H-G**, **Feld T and Hoffmann M** (2014) Industry 4.0. *Business & Information Systems Engineering 6*, 239–242.

https://doi.org/10.1017/dap.2024.19 Published online by Cambridge University Press

Li L, **Fan Y**, **Tse M and Lin K-Y** (2020a) A review of applications in federated learning. *Computers & Industrial Engineering 149*, 106854.

Li Q, **Wen Z**, **Wu Z**, **Hu S**, **Wang N**, **Li Y**, **Liu X and He B** (2021) A survey on federated learning systems: Vision, hype and reality for data privacy and protection. *IEEE Transactions on Knowledge and Data Engineering 35*(4), 3347–3366.

Li T, **Sahu AK**, **Talwalkar A and Smith V** (2020b) Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine 37*(3), 50–60.

Liu X, **Shao S**, **Yang Y**, **Wu K**, **Yang W and Fang H** (2021) Secure federated learning model verification: A client-side backdoor triggered watermarking scheme. *IEEE International Conference on Systems, Man, and Cybernetics*, 2414–2419.

Liu Y, **James J**, **Kang J**, **Niyato D and Zhang S** (2020) Privacy-preserving traffic flow prediction: A federated learning approach. *IEEE Internet of Things Journal 7*(8), 7751–7763.

Manoj T, **Makkithaya K and Narendra V** (2022) A federated learning-based crop yield prediction for agricultural production risk management. In *2022 IEEE Delhi Section Conference* 1–7.

McMahan B, **Moore E**, **Ramage D**, **Hampson S and Arcas BA** (2017) Communication-efficient learning of deep networks from decentralized data. *Artificial Intelligence and Statistics*.

McMahan HB, **Ramage D**, **Talwar K and Zhang L** (2017) Learning differentially private recurrent language models. Preprint, arxiv:1710.06963.

Mikhaylov SJ, **Esteve M and Campion A** (2018) Artificial intelligence for the public sector: Opportunities and challenges of cross-sector collaboration. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 376*(2128), 20170357.

Mironov I (2017) Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*. https://doi.org/10.1109/csf.2017.11.

Mitra A, **Jaafar R**, **Pappas GJ and Hassani H** (2021) Linear convergence in federated learning: Tackling client heterogeneity and sparse gradients. *Advances in Neural Information Processing Systems*.

Mothukuri V, **Parizi RM**, **Pouriyeh S**, **Huang Y**, **Dehghantanha A and Srivastava G** (2021) A survey on security and privacy of federated learning. *Future Generation Computer Systems*. *115*, 619–640.

Niknam S, **Dhillon HS and Reed JH** (2020) Federated learning for wireless communications: Motivation, opportunities, and challenges. *IEEE Communications Magazine 58*(6), 46–51.

OECD (2019) Enhancing access to and sharing of data: Reconciling risks and benefits for data re-use across societies. Paris: OECD.

Olson M (1965) *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Ozfatura E, **Ozfatura K and Gündüz D** (2021) Time-correlated sparsification for communication-efficient federated learning. In *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 461–466.

Passerat-Palmbach J, **Farnan T**, **McCoy M**, **Harris JD**, **Manion ST**, **Flannery HL and Gleim B** (2020) Blockchain-orchestrated machine learning for privacy preserving federated learning in electronic health data. In *2020 IEEE International Conference on Blockchain*.

Pingitore G, **Rao V**, **Dwivedi K and Cavallaro K** (2017) To Share or Not to Share. Available at https://www2.deloitte.com/content/dam/insights/us/articles/4020_To-share-or-not-to-share/DUP_To- share-or-not-to-share.pdf.

Reimsbach-Kounatze C (2021) Enhancing access to and sharing of data: Striking the balance between openness and control over data. In *Data access, consumer interests and public welfare*, pp. 25–68. Nomos Verlagsgesellschaft mbH & Co. KG.

Shae Z and Tsai J (2018) Transform blockchain into distributed parallel computing architecture for precision medicine. In *International Conference on Distributed Computing Systems*, pp. 1290–1299.

Shah SM and Lau VK (2021) Model compression for communication efficient federated learning. *IEEE Transactions on Neural Networks and Learning Systems 34*(9), 5937–5951.

Sprenkamp K, **Delgado Fernandez J**, **Eckhardt S and Zavolokina L** (2023) Federated learning as a solution for problems related to intergovernmental data sharing. In *56th Hawaii International Conference on System Sciences*.

Sprenkamp K, **Zavolokina L**, **Angst M and Dolata M** (2023) Data-driven governance in crises: Topic modelling for the identification of refugee needs. In *Proceedings of the 24th Annual International Conference on Digital Government Research*, pp. 1–11.

Trampusch C (2023) Regulating the digital economy: Explaining heterogenous business preferences in data governance. *Journal of European Public Policy*, 1–25.

Truong N, **Sun K**, **Wang S**, **Guitton F and Guo Y** (2021) Privacy preservation in federated learning: An insightful survey from the GDPR perspective. *Computers & Security 110*, 102402.

Tuor T, **Wang S**, **Ko BJ**, **Liu C and Leung KK** (2021) Overcoming noisy and irrelevant data in federated learning. In *International Conference on Pattern Recognition*.

Vaswani A, **Shazeer N**, **Parmar N**, **Uszkoreit J**, **Jones Ł**, **Gomez AN**, **Kaiser F and Polosukhin I** (2017) Attention is all you need. *Advances in Neural Information Processing Systems 30*.

Ward BT and Sipior JC (2010) The internet jurisdiction risk of cloud computing. *Information Systems Management 27*(4), 334–339.

Webster J and Watson RT (2002) Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly 26*(2), xiii–xxiii.

Wheaton S and Martuscelli (2021) WHO, Berlin Float Sanctions If Countries Suppress Information on Pandemics. Available at https://www.politico.eu/article/who-berlin-float-sanctions-if-countries-suppress-information-on-pandemics/ (accessed 20 May 2022).

WHO (2021) Global leaders unite in urgent call for international pandemic treaty. Available at https://www.who.int/news/item/30-03-2021-global-leaders-unite-in-urgent-call-for-international-pandemic-treaty (accessed 19 May 2022).

Wiseman J (2020) Silo busting: The challenges and success factors for sharing intergovernmental data. IBM Center for The Business of Government.

Xu J, **Glicksberg BS**, **Su C**, **Walker P**, **Bian J and Wang F** (2021) Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research 5*, 1–19.

Yang D, **Xu Z**, **Li W**, **Myronenko A**, **Roth HR**, **Harmon S**, **Xu S**, **Turkbey B**, **Turkbey E**, **Wang X**, *et al.* (2021a) Federated semi-supervised learning for covid region segmentation in chest ct using multi-national data from China, Italy, Japan. *Medical Image Analysis 70*, 101992.

Yang Q, **Liu Y**, **Chen T and Tong Y** (2019) Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology 10*(2), 1–19.

Yang W, **Zhou Y**, **Hu M**, **Wu D**, **Zheng X**, **Wang JH**, **Guo S and Li C** (2021b) Gain without pain: Offsetting Dp-injected noises stealthily in cross-device federated learning. *IEEE Internet of Things Journal 9*(22), 22147–22157.

Yu H, **Liu Z**, **Liu Y**, **Chen T**, **Cong M**, **Weng X**, **Niyato D and Yang Q** (2020) A fairness-aware incentive scheme for federated learning. In *Conference on AI, Ethics, and Society*, pp. 393–399.

Yukhno A (2022) Digital transformation: Exploring big data governance in public administration. *Public Organization Review 24*, 335–349.

Ziller A, **Trask A**, **Lopardo A**, **Szymkow B**, **Wagner B**, **Bluemke E**, **Nounahon J-M**, **Passerat-Palmbach J**, **Prakash K**, **Rose N**, *et al.* (2021) Pysyft: A library for easy federated learning. In *Federated Learning Systems*. Berlin: Springer.

Cite this article: Sprenkamp K, Delgado Fernández J, Eckhardt S and Zavolokina L (2024). Overcoming intergovernmental data sharing challenges with federated learning. *Data & Policy*, 6: e27. doi:10.1017/dap.2024.19