German Data-Sovereign Government-to-Citizen Use Case

authors omitted

Abstract

Amidst data protection incidents and IT security breaches, it is crucial that web applications processed data remains secure from unauthorized access. This is true in particular in the context of citizens' data due to the possible negative implications. However, the digitization of government-to-citizens (G2C) processes is required and has a huge potential to optimize administrative workflows while also providing benefits for citizens. Our implementation showcases citizen data storage and processing in the context of public administration processes using Solid principles. We are particularly concerned with the re-sharing of collected data during the gradual filling of the citizens' data storage for later reusing the collected data in other Solid G2C applications.

Keywords

Solid, Data Souvereignity, Web, Government Data, Citizen

1. Introduction

Navigating public services is a task every citizen is confronted regularly – also in Germany. Often, public service processes, such as getting a license plate, require citizens to provide additional data, which must be obtained from other public authorities.

Offering public services digitally can facilitate the navigation of processes for citizens (cf. the German *Onlinezugangsgesetz* [1, 2] aiming to provide 575 public services digitally). Additionally, if personal or public service data is available digitally, communication between public authorities can be streamlined, reducing the effort needed on behalf of citizens to provide data as well as providing the public authorities machine-readable access to the required data, hence, reducing the manual effort for citizens and public authorities.

One of the challenges in providing digitally available public services is ensuring data protection and privacy. If personal data concerning a citizen is produced and maintained by one public office, this data cannot easily be shared with other parties, even if it is needed for an official process. This places responsibility for obtaining information from secondary public offices with the citizens, which leads to reduced efficiency since navigating service offices can be time-consuming.

A parallel technological advancement is a Solid specification, a decentralized and secure storage solution that has been under development by the Solid Community Group since 2018. Additionally, this approach is based on publicly defined schemas, open to all.

As these specifications aim to provide solutions for problems similar to those described previously, the applicability of the specification for German public authorities should be examined.

The 1st Solid Symposium Poster Session, co-located with the 2nd Solid Symposium, May 02 – 03, 2024, Leuven, Belgium

Following an agreed-upon set of standards and specifications, each office can develop its IT infrastructure at its own pace, while ensuring compatibility of produced data types.

The situation described leads to the following key research question this paper aims to address: *Is the Solid technology suitable for a naive e-government solution, or does it require additional specifications to serve as an approach for the development of e-government applications?*

In this context, we refer to solutions proposed in this work as naive, since they were developed focusing only on the citizen as a user. As a result, complicated data processing and government business logic are excluded, and only the storage, access, and submission of citizen data are covered. To answer these questions, in this work, we will formulate requirements for e-government applications based on common scenarios. Based on these requirements, we will implement a showcase scenario with multiple governmental authorities' applications. These applications will be built on the specifications developed by the Solid Community Group and use currently available reference implementations. Various concerns need to be considered in the development of e-government applications, such as interoperability of produced applications and data with other systems or usability of the final product, etc. These concerns are out of the scope of this paper, as more advanced testing and research need to be done.

2. Related Work

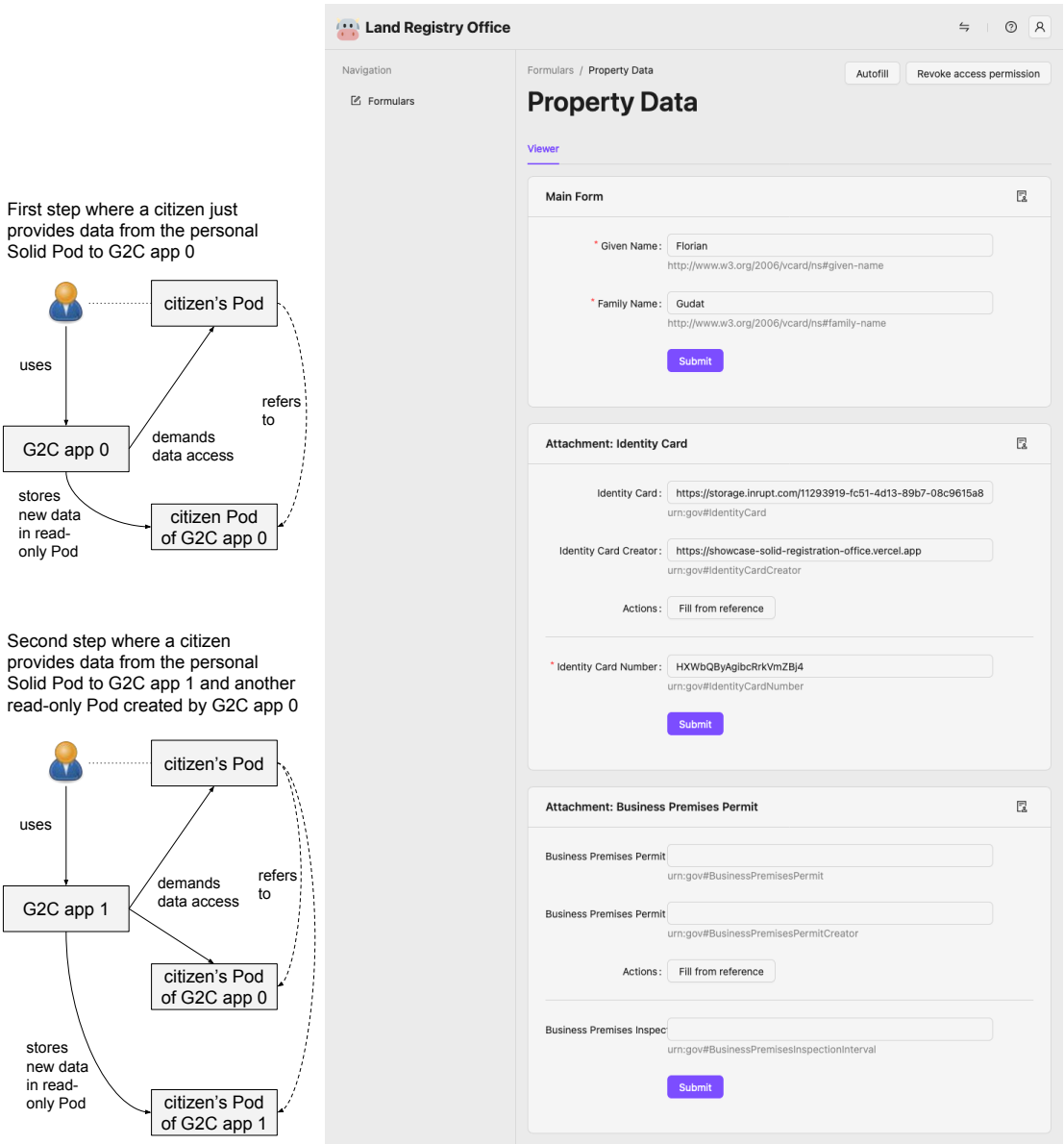
Since the Solid project started, and the first demonstration of a Social Application using the Solid specifications was published [3], there have been several attempts to apply the ideas to different domains.

Zhao et al. [4] demonstrated a way to share public transportation data through a Solid framework. Werbrouck et al. [5] showed an application to access common data shared in construction processes. They identified Solid specifications as a useful way to adapt Linked Data principles in the domain of Building Information Modeling (BIM). Henselmann et al. [6] applied the Solid specifications to a loan request use case, showing that Solid can be useful in a B2B or B2C context. Some research has been made on the use of Solid specifications in the healthcare domain, e.g., [7, 8].

In "Making Sense of Solid for Data Governance and GDPR" [9], a theoretical exploration of Solid in the context of GDPR, the authors describe a lack of features in current Solid specifications, given the compliance obligations and enforcement as envisioned by GDPR. However, Sun et al. [10] and Bailly et al. [11] investigated the development of Solid applications under the GDPR, focused on personal health data and general personal data, respectively. The latter especially explored the difficulties in UI design that arise, when combining the conceptual requirements of both interoperability and data privacy restrictions. As pointed out by Penteado et al. [12] the biggest challenge in providing government data in a Linked Data structure is still that there are no uniform processes yet, even though linked open data offers a lot of useful advantages.

3. Architecture and Implementation

The expected architecture is driven (besides the well-known GDPR [13] and other typical requirements for data-driven ecosystems, e.g., [14, 15]) by the particular constellation, appearing



(a) Big picture of architecture. (b) Screenshot of the G2C application *land-registry-office*.

Figure 1: G2C use case big picture and a G2C app screenshot

in administrative scenarios, which can be summarized as follows: Typically an administrative authority A creates new data on the demand of a citizen C and depending on C's input. As C owns all the data, the person can share it, but cannot change or delete data that was created by A. Hence, the created data needs to stay under the control of A while this data needs to integrate into the Solid infrastructure, s.t., it can be shared seamlessly with other applications. The concluding architecture requires a network of Solid pods providing C's data and one Solid

pod of the citizen. In the latter, references to C's data in A's pod are maintained, s.t., C is able to access their data, as well as share these references with other parties. Figure 1a shows the described meta-structure of our approach with two G2C apps.

In our showcase, we implemented a network of multiple applications providing (simulated) e-governance processes and corresponding data for citizens. Two distinct types of applications are required, along with a data storage system used by these applications. Via a specifically designed application, citizens can manage their personal data, i.e., preparing information manually that could be used in the governmental processes. Moreover, when used with other applications, it provides supplementary functions such as managing the data and the data access.

For e-governance processes, we propose the second type of application, the government applications, each of which fulfills the role of a single public administration service. These government applications A enable the (simulated) e-governance functionality and allow citizens to interact with government services, such as providing personal information via a form that subsequently creates new data in the corresponding governmental Solid pod (cf. Figure 1b). If C decides to share it's information with A, the G2C application scans C's pod (here: to fill the first form block) and the referred pods of other governmental applications (here: to fill the second form block), identifying and collecting the data required for the current process, leading to the automation of administrative procedures. However, missing data needs to be added manually by C (i.e., if C doesn't share his pod or the data is not available, all data needs to be entered manually), cf. the last form block.

In our implementation, we created manually an ontology to represent the governmental data. Here, concepts representing the data created by government authorities are defined, s.t., instances created for C can be reused by other government authorities via C's pod. To validate our approach, we used `vcard:given-name`, `vcard:family-name`, `vcard:locality`, and 12 data types (e.g., `gov:VehicleRegistration`, `gov:IdentityCard`) that would be created by the 12 government authorities' application we implemented for the use case. For example, the showcase's government authorities' application *land-registry-office* requires the `vcard:given-name`, `vcard:family-name`, `gov:IdentityCard`, as well as optionally `gov:BusinessPremisesPermit` and will create `gov:PropertyData` which might be used by the applications *construction-office* or *environmental-office*.

4. Conclusions and Future Work

While utilizing the Solid technologies, our approach enables a digitized, data-sovereign ecosystem of government authorities. We described the special requirements of the G2C scenario where data can typically only be shared by a citizen but not changed. Our use case shows that Solid can be used as a solution to overcome the typical problem of government environments, where inter-authority data exchange is usually poorly established. Hence, using our approach, overcoming this problem is possible while additionally providing a data-sovereign sharing option under the control of the citizens.

References

[1] Gesetz zur Verbesserung des Onlinezugangs zu Verwaltungsleistungen, ???? URL: https://www.gesetze-im-internet.de/ozg/, translation: Act to improve online access to administrative services.

[2] Dashboard digitale verwaltung, ???? URL: https://dashboard.ozg-umsetzung.de/, translation: Dashboard digital administration.

[3] E. Mansour, A. V. Sambra, S. Hawke, M. Zereba, S. Capadisli, A. Ghanem, A. Aboulnaga, T. Berners-Lee, A demonstration of the Solid platform for social web applications, in: Proceedings of the 25th International Conference Companion on World Wide Web, WWW '16 Companion, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 2016, p. 223–226. doi:10.1145/2872518.2890529.

[4] W. Zhao, B. Zhou, C. Zhang, Heterogeneous social linked data integration and sharing for public transportation, Journal of Advanced Transportation 2022 (2022) 6338365. doi:10.1155/2022/6338365.

[5] J. Werbrouck, P. Pauwels, J. Beetz, L. van Berlo, Towards a decentralised common data environment using linked building data and the Solid ecosystem, in: 36th CIB W78 Conference, 2019, pp. 113–123.

[6] D. Henselmann, K. Kolinsky, S. Schmid, D. Schraudner, A. Both, A. Harth, Solid proof of concept in an enterprise loan request use case, 2022. URL: https://publica.fraunhofer.de/handle/publica/445586.

[7] H. Chen, Ubi-care: a decentralized ubiquitous sensing healthcare system for the elderly living support, in: 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2019, pp. 543–547. doi:10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00108.

[8] X. Weng, H. Wu, Y. Pan, H. Chen, Decentralized personal cloud data model and its application in campus health information system, in: 2021 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech), 2021, pp. 879–883. doi:10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00146.

[9] H. J. Pandit, Making sense of Solid for data governance and GDPR, Information 14 (2023). URL: https://www.mdpi.com/2078-2489/14/2/114. doi:10.3390/info14020114.

[10] C. Sun, M. Gallofré Ocaña, J. van Soest, M. Dumontier, ciTIzen-centric DAta pLatform (TIDAL): Sharing distributed personal data in a privacy-preserving manner for health research, Semantic Web 14 (2023) 977–996.

[11] H. Bailly, A. Papanna, R. Brennan, Prototyping an end-user user interface for the Solid application interoperability specification under GDPR, in: The Semantic Web: 20th International Conference, ESWC 2023, Hersonissos, Crete, Greece, May 28–June 1, 2023, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2023, p. 557–573. doi:10.1007/978-3-031-33455-9_33.

[12] B. E. Penteado, J. C. Maldonado, S. Isotani, Methodologies for publishing linked open

government data on the Web: A systematic mapping and a unified process model, Semantic Web 14 (2023) 585–610. doi:10.3233/SW-222896, 3.
[13] The European Parliament and the Council of the European Union, Regulation (EU) 2016/679 (General Data Protection Regulation) GDPR, ???? URL: https://gdpr-info.eu/.
[14] V. Stafford, Zero trust architecture, NIST special publication 800 (2020) 207.
[15] M. Shore, S. Zeadally, A. Keshariya, Zero trust: The what, how, why, and when, Computer 54 (2021) 26–35.