HTWK Leipzig University of Applied Sciences

Exposé for a Master Thesis on the Topic

Solid-Based Extended Access Control and Traceability in Data-Driven Web-Based Systems

Florian Gudat

Version 43554df, 2024-08-07

Table of Contents

Version

43554df, 2024-08-07

Editor

Florian Gudat

Module

Mastermodul (C533.2 Compulsory module)
https://modulux.htwk-leipzig.de/modulux/modul/6291

Module Supervisor

Prof. Dr.-Ing. Jean-Alexander Müller

Lecturer

Herr Prof. Dr. rer. nat. Andreas Both
Herr M. Sc. Michael Schmeißer

Institute

Leipzig University of Applied Sciences

Faculty

Computer Science and Media

The [Solid Protocol] enables secure storage of data in a decentralised environment, which can then be shared with others, such as individuals, organisations, or applications via the Internet. The data is decentralised and so is the access control information. The owner of the data and access control information is responsible for managing these contents.

Chapter 1. Motivation

The Solid project, and the protocol that is part of it, is under the control of the Solid Community Group [https://www.w3.org/community/solid/], which started in 2018. As evident in their technical reports [https://solidproject.org/TR/], there are still parts of the technology that are draft and not yet finalised. The main problem here is that all this technology is quite new and not yet widely used. There are few implementations of the protocol and the technology is not yet mature. As a result, there are many open questions and issues that need to be addressed.

Esposito et al. (2023, p. 18) summarized the current state of the [Solid Protocol] when it comes to security and privacy obligations. One of his requirements is the need for tamper-proof access logs with different views depending on their mandate. Due to his research, new measures are necessary to meet security and privacy obligations.

Access monitoring in distributed systems has also been recognised as critical by the German authorities. The *XDatenschutzcockpit* project aims to increase transparency for citizens by offering a national data protection cockpit that monitors access to their data and allows them to assess what data has been exchanged between authorities, as well as the timing and purpose of such exchanges. The pilot phase was completed in September 2022, leading to the creation of the specification, which is tied to e-government services. (Diederich, 2023)

The combination of both ideas leads to the approach of applying the data protection cockpit specification to the [Solid Protocol], as the [Solid Protocol] represents a promising approach that is independent of the business logic. It is proposed to use the [Solid Protocol] as the technical basis for the data protection cockpit. This is supplemented by the findings from the *XData Protection Cockpit* specification.

Chapter 2. Objectives and research interest

Tao & Chen (2016) noted the existence of multiple protocols, with new ones arising frequently. In response, they developed a universal method for augmenting the supported protocols of a server, using a server-side application layer proxy. When a new protocol is introduced, a module specific to that protocol is included in the set of modules managed by the application layer proxy. The module will solve the specific problem of the protocol, or throw an error if unable to do so. This method is adaptable and may be utilised with the [Solid Protocol]. Further division may be accomplished by implementing dedicated modules to handle specific concerns of the data privacy cockpit, as defined by Diederich (2023). By doing this, new features can be added to the data privacy cockpit without the need to change the basic infrastructure.

In addition, Yan & Wang (2017, pp. 233-235) proposed a gateway mechanism with analysis and filtering modules such as blacklisting, whitelisting, anomaly detection and validation. This approach, especially when the rules and configurations are under the control of the data owners, could significantly improve the ability of the data privacy cockpit to detect and prevent data theft.

The aim of this thesis is to determine how to achieve logical topologies that allow the use of a data privacy cockpit through a server-side application layer proxy. Furthermore, it is necessary to identify the circumstances that allow the retrieval of information while ensuring the compatibility with the [Solid Protocol] and the general properties of the proxy (Shapiro, 1986, p. 9).

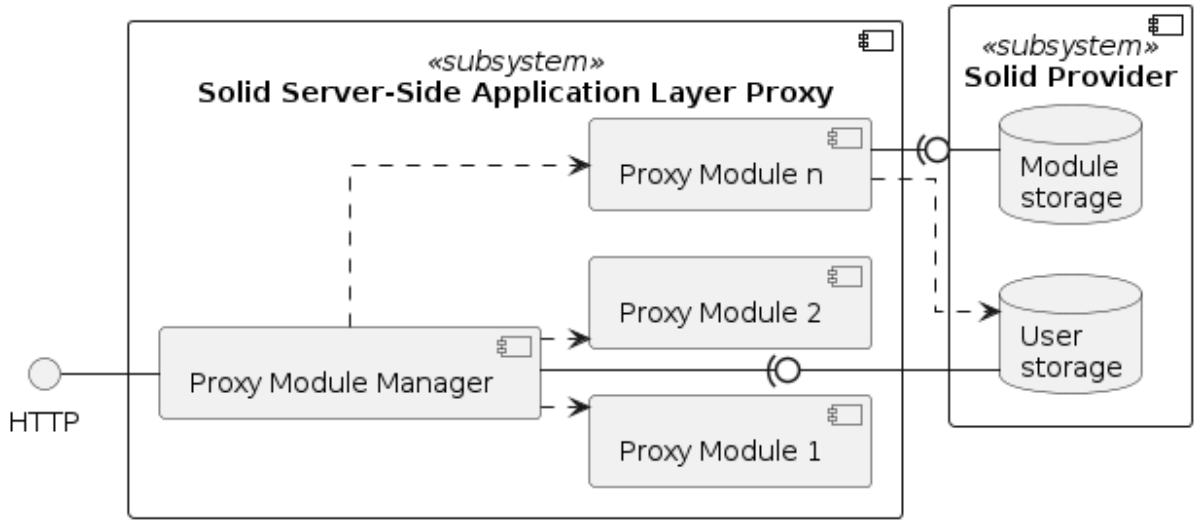

Figure 1. Architectural sketch of the logical topology

As the figure above shows, prototypes of a server-side application layer proxy and a data privacy cockpit have to be developed. The prototypes need to enhance the [Solid Protocol]'s capabilities while ensuring extensibility and decoupling from the underlying infrastructure.

Chapter 3. Research Concept

The following key research question is to be answered in this Master's thesis:

- What requirements must a server-side proxy of the application layer fulfill to extend the capabilities of the [Solid Protocol] while ensuring extensibility and decoupling from the underlying infrastructure?

Derived from this key research question, the following sub-questions are to be answered:

- What effects do different logical topologies, such as changes in access, transport, addressing, protocols and data paths, have on the availability of information within the application layer proxy?
- In which use cases can server-side application layer proxies be used, and where are the limits if conformity with the [Solid Protocol] is to be guaranteed?
- How high is the system resource load during filtering and modification by a server-side proxy of the application layer using the [Solid Protocol]?

For the analysis, a data privacy cockpit is developed as a Solid app that manages the information collected by the server-side application level proxy. It is assumed that the range of functions can be extended variably, and that the possibilities of the various logical topologies can therefore be utilised. The data privacy cockpit is just an example of a complex situation in which the procedure can be used and serves to evaluate the concept. Conclusions for similarly complex use cases can be derived from this if necessary.

Steps: As part of the Master's thesis, the following methodological steps are carried out:

1. Analysis of the requirements for a server-side application layer proxy and data privacy cockpit
2. Development of a server-side application layer proxy and data privacy cockpit
3. Evaluation of the developed solution
4. Comparison of the results with the requirements
5. Discussion of the results

This solution includes the development of an experimental prototype to identify different logical topologies and exclude them if necessary. The prototype will also be used to test the limits of the use cases and evaluate the system load.

Chapter 4. Preliminary Outline

1. Introduction
    - Requirements
    - Research
2. Terminology
    - Solid Protocol
    - Proxy
    - Data Privacy Cockpit
3. Application Design
    - Logical Topology (Composition)
    - Information Retrieval
4. Integration
    - Logical Boundaries
    - Application Programming Interfaces
5. Quality Model
    - Characteristics
    - Indicators
6. Methodology
    - Laboratory Prototype
    - Quality Management
7. Result Presentation
    - Boundaries of the Composition
    - Possible Uses of the Concept
    - Resource Requirements
8. Discussion
9. Conclusion

Chapter 5. Roadmap

Duration: Six months (2023-12-05 to 2024-06-05)

Table 1. Timetable for writing the Master's thesis

Until 18.12.	Literature research
Until 01.01.	Thematic introduction and hypotheses
Until 14.02.	Draft of introduction and main body
Until 27.03.	Creation and evaluation of prototypes
Until 23.04.	Finalisation of introduction, main body and conclusion
Until 20.05.	Revision and correction
Until 29.05.	Layout and title page
Until 01.06.	Print
Until 05.06.	Submission

Bibliography

Diederich, G. (2023). *XDatenschutzcockpit Teil 1: Technologieunabhängige Spezifikation*. https://www.xrepository.de/details/urn:xoev-de:kosit:standard:xdatenschutzcockpit

Esposito, C., Horne, R., Robaldo, L., Buelens, B., & Goesaert, E. (2023). Assessing the Solid Protocol in Relation to Security and Privacy Obligations. *Information*, *14*(7), 411. https://doi.org/10.3390/info14070411

Shapiro, M. (1986). *Structure and Encapsulation in Distributed Systems: the Proxy Principle*. 198–204.

Tao, Y., & Chen, G. (2016). An Extensible Universal Reverse Proxy Architecture. *2016 International Conference on Network and Information Systems for Computers (ICNISC)*, 8–11. https://doi.org/10.1109/ICNISC.2016.012

Yan, F., & Wang, Y. (2017). A Security Web Gateway Based on HTTP Reverse Proxy. *DEStech Transactions on Engineering and Technology Research*, *iceta*. https://doi.org/10.12783/dtetr/iceta2016/7003