SEAL Project: User-Centric Application of Linked Digital Identity for Students and Citizens

Francisco José Aragó-Monzonís
Universitat Jaume I
Castellón, Spain
e-mail: farago@uji.es

Laura Domínguez-García
Universidad de Málaga
Málaga, Spain
e-mail: lauradomg@uma.es

Victoriano Giralt
Universidad de Málaga
Málaga, Spain
e-mail: victoriano@uma.es

Basurte-Durán, Alberto
Universidad de Málaga
Málaga, Spain
e-mail: abasurte@uma.es

Raúl Ocaña
Universidad de Málaga
Málaga, Spain
e-mail: raulocana14@uma.es

***Abstract*— The SEAL (Student and Citizen Identity Linked) project targets two important aspects of identity management: give the users control over their own data, and facilitate the reconciliation of the multiple electronic identities citizens have across public and private institutions. Through a modular, extensible, and scalable design, the SEAL service empowers the citizen to build his own persistent and unique identity, while still having control over the anonymity and traceability. This citizen empowerment is due to the exploration of the Self-Sovereign Identity concept, which reduces data exposition and critical dependencies, building a bridge between the federated data world and the Self-Sovereign Identity new horizons. In this paper we describe the paradigm that SEAL proposes for dealing with identity reconciliation issues and its contribution to secure digital identities, and how users can control all of this through the SEAL service dashboard.**

Keywords-identity reconciliation; self-sovereign identity; identity federation; know your customer; eIDAS.

I. INTRODUCTION

SEAL has been born amidst a complex situation. The rapid growth in online service usage and needs in the last decades greatly exceeded the organisational capacity of the regulatory bodies, producing an environment where users own personal data in a myriad of different and non-interoperable sources. This situation has been tackled partially through different aspects, like for example the authentication. In the academia sector, the eduGAIN (Education Global Authentication Infrastructure) network [1] facilitated an effective sharing of online services and resources among higher education and research institutions using SAML2 (Security Assertion Markup Language v2) standard [2], and later, the European Commission has taken a similar approach to produce a framework for cross-border acceptance of state-issued electronic identities [3]. However, this tendency is not limited to authentication: EMREX project [4] created a framework for the trusted exchange of academic records, to support students moving abroad in getting credit recognition. The main common factor behind these initiatives is a user-centric approach, in accordance with the evolution of the data protection legislation that aims at giving the users full sovereignty over the personal data.

The increasing demand for integrated online user services requires trusted, effective data interoperability. At the centre of this lies a key but usually ignored issue: how to determine if two sets of data identified by non-matching identifiers belong to the same individual. The usual solution to this lies in trusting the users' statement of ownership. However, on many high-profile use cases (especially on official administrative procedures), this is not an option, as the users can benefit from counterfeiting data. This would not be an option in the context of trusted infrastructures, but the blind spot of non-matching identifiers in data sources can open the door to borrowing data from another citizen. Ensuring the ontological relationship between two sets of data and an individual requires a process of comparing contextual data of the subject of both sets (i.e., comparing name, surname, date of birth, etc.). This process is costly and probabilistic, requiring human interaction in many edge-cases, especially if the trust standards are high. This is the reason for the growth of specialised Know-Your-Customer (KYC) companies that provide this identity reconciliation service. However, this is still a chaotic and non-standard field.

SEAL tries to build a key element to formalise the management of data and their relationships, by coordinating the different involved actors to let the users be in the centre of it, bringing the data under a virtually unified and persistent identity, while still guaranteeing all his rights and giving full control over the traceability and anonymity. The SEAL service develops a modular and extensible framework that allows plugging in to sources of identity and data, KYC providers and data consumers, to let the users collect her data sets, establish trusted links among them, and deliver them to the consuming parties, namely university user services.

This work is organised in two main sections: SEAL innovation as digital identity, which shows SEAL architecture, and user-centric management for students, where the user-centric approach of the SEAL project is detailed and the users' experience interaction -web and mobile- is described.

 ISBN: 978-1-61208-760-3

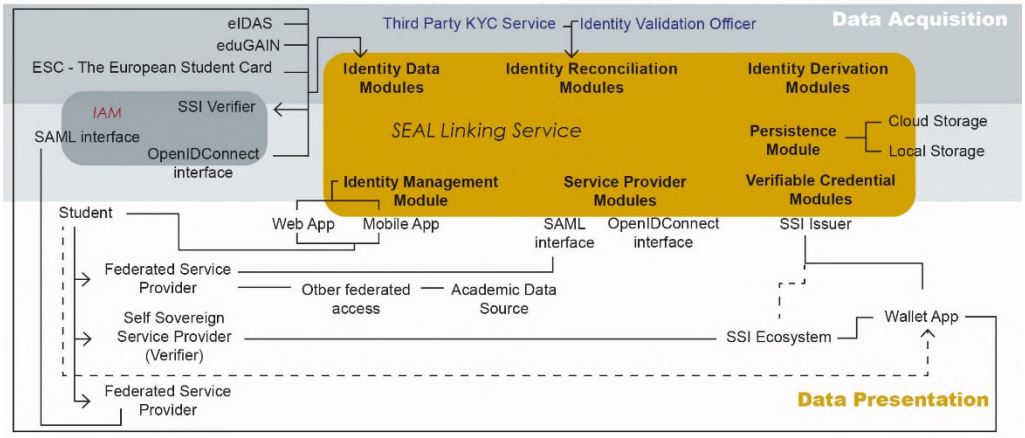

Figure 1. SEAL service architecture structure layout.

II. SEAL INNOVATION AS DIGITAL IDENTITY

SEAL proposes an innovative framework aimed at creating a standardized and holistic way of dealing with an often disregarded issue: identity reconciliation. Only in the latest years, the growth in the demand for online services, has brought into light this missing piece, that hinders the establishment of online services and procedures(especially in the public sector and in banking) that show a big potential demand but also need to establish with a high level of trust that two identities refer to the same individual citizen, without having to trust the citizen for that. This has brought to the appearance of KYC companies that develop complex assurance procedures to establish this fact, and sell their services to interested companies, and also the approval of specific regulations, like eIDAS (Electronic Identification Authentication and trust Services) [5] regulation, that settle the legal grounds for the public acceptance of said methods.

The state of the art, shows independent KYC providers dealing with independent consumer entities, doing specific and no-portable integrations. The SEAL project has recognised this gap and proposes a structured, common and standardised approach to create a common ground that brings together all the actors in this scenario: KYC providers, data providers, data consumers, and builds it around the user, to ensure the protection of the users' rights regarding data management. And it tries to do so in a manner to open the way for standardization of procedures, and minimisation of efforts for the implied stakeholders, coordinating the actions of all of them in the most efficient way.

A. Architecture

SEAL is not a new data transport infrastructure, but a support framework. It minimises the reconciliation effort for data consumers by centralising this management on the SEAL service and allowing the reuse of this effort by storing the trusted links on secure storage areas. SEAL is divided in three layers, with information flowing from top to down, as shown in Figure 1. First, we find a data acquisition layer, where we find three specific interfaces that allow data into SEAL: identity data access, identity reconciliation and identity derivation. Of these interfaces, multiple pluggable implementations can be provided to support almost any source of data or KYC provider. Out of the box, SEAL provides integration with eduGAIN and with eIDAS, and internal automated and manual KYC support procedures.

On the middle layer, we find the management and storage interfaces. SEAL allows any API (Application Programming Interface) compliant client to allow users to connect and interact, to retrieve data and establish links. Two clients are provided: a mobile and a browser-based client. The storage can be performed on a series of locations under the control of the users: mobile storage, local storage on the computer or, cloud storage services. Stored data is encrypted, so only the users can decrypt it and decide when it is decrypted and fed into SEAL, but also the data is signed to ensure the users do not tamper with it, breaking the chain of trust: the user owns the data, however, SEAL does not trust the users.

The lower layer is the data presentation layer. There, any data consumer can access the data users want to deliver through multiple ways, divided into two main blocks: federated access and self-sovereign access. On the federated access, the relaying party sends a request to SEAL using a standard delegation protocol: SAML2 and OIDC (Open ID Connect) [6], having this last one more fitted with the concept of user-centric federated identity management [7]. Then, the users are requested to access their storage to retrieve the required data and send it back to the relaying party. On the self-sovereign access, SEAL acts as an issuer of verifiable claims. These claims are stored on a user wallet (similar to the SEAL storage, but designed as an independent agent) that the users can carry around and deliver to data consumers, that will verify the integrity and trust on those claims through validation data published on a distributed ledger. This access model skips the SEAL service once the data has been issued, removing critical central points in the processing, which is the base of the self-sovereign concept.

III. USER-CENTRIC APPLICATION MANAGEMENT FOR STUDENTS

A user-centric approach allows for general-purpose infrastructures supporting a wide array of use cases, because

the users interact with the components. They can be properly informed and can request their consent over any operation that needs to be performed on their data, asked to discover the sources, and is able to follow and cut the process at any point.

This concept has evolved into the self-sovereignty of the data, which advocates for the users to be the keeper and have an effective control over their data, not just a formal one. The concept has been gaining strength, especially in the context of the increasing risks and costs for private companies to secure the sensitive user data being kept in their systems: if the users keep their own data and offer it to be processed only when needed, the window of risk is greatly reduced, and companies become less appealing targets. SEAL linked identity offers key advantages regarding privacy and security of High Education Institutions (HEI) students in their operations and interactions as users of educational, institutional and private services. In this regard, there are some potential features to be addressed not only from the student perspective, but also from the whole academic community. The SEAL project is an example of how compatible data of interest can be linked as other projects previously did, like Europeana [8] focuses on electronic resources data aggregation, to fulfil citizens needs providing a secure user-controlled digital identity framework.

The European Blockchain Services Infrastructure (EBSI) specifies four use cases for 2019: Notarisation, Diplomas, European Self-Sovereign Identity and Trusted Data Sharing [9]. The SEAL service covers all these cases and can contribute to secure the process, resulting in a reduction of the verification costs and an improvement of the authenticity trust between organizations [10]. The potential of the SEAL service lies in its power to create and utilise the user-centric service features of digital identity to improve the user experience of accessing restricted personal digital areas among institutions or e-Government authorities. Facilitating the authentication process improves the user experience by avoiding nowadays long administrative processes. This has been addressed by A. Crespo et al. [11] as the offering of cross-border services allows access to new services with credentials that the citizen already possesses and without the need for issuing local ones.

A. Dashboard application and user interaction

As described by X. Liu et al. [12] the Web has undergone toward a highly user-centric environment where millions of users can participate and collaborate for their own interests and benefits. Thus, the SEAL dashboard allows users to lead their actions. The SEAL service can be used from different devices and two service applications have been designed: a web dashboard and a native mobile application compatible with Android and iOS (iPhone Operating System) systems.

1) Access methods and data storing: Two access options are offered: centralised and decentralised. Both of them fulfil the security requirements established for the SEAL service and the same privileges are granted. The centralised Personal Data Store (PDS) access requires the user's password. Nevertheless, as personal data is a complex issue, not all the attributes are available to users as linked ones [13]. This access method offers two options for storing the data file: local and cloud, both of them require the user's password, and when locally stored, the file is saved in the user device's system.

The decentralised Self-Sovereing Identitty (SSI) access method uses a Distributed Ledger Technology (DLT) to assure and verify the identity of a person, together with the integrity of the personal data. As suggested by A. Benzekri et al, this could avoid uploading official documents during the registration process, which is tedious and degrades the users' experience, especially with mobile [14]. This access method requires an external mechanism to connect, read and generate data over the DLT infrastructure, and it behaves as an intermediary between the SEAL service and the ledger: the uPort [15] API. Since this API needs a local wallet [16] to fulfil the users' credentials, which in turn authorise the operations over the DLT, the SSI access method can be seen as a conjunction of the dashboard, the uPort API and the DLT infrastructure.

2) SEAL Dashboard functionalities: Linked digital identities can be defined as linked data capable of storing data and allowing a service to tie linked resources into a worldwide network [17]. The SEAL service uses this new technology to retrieve different identities, link them and expand users' feature options while strengthening the authentication process; as this is the major advantage of Linked Data technology [18].

The soul of the service in terms of identity collection is the Identity Reconciliation functionality, which allows the users to link two identities (manually or automatically) and create a new one. In order to assure the users' privacy, the identity reconciliation only takes place when the users select a suitable pair of identities and start the process. The users can check its progress thanks to the Reconciliation Status function. Besides, the SEAL service offers the possibility of loading a new identity and including it among those from the Identity Reconciliation functionality. The Retrieve Identity Data function performs the data collection and its storage in the session. Thus, the users can display the loaded identities data and store the current session. About the storage, the Configuration Data Store functionality presents the two options: local and cloud; the selected one will remain for future access.

Concerning the DLT operations, the Manage Verifiable Claims functionality allows the users to create, retrieve and store a Verifiable Credential (VC), which is a set of tamper-resistant claims made by an issuer (the SEAL service in this case) where each claim asserts a set of properties about a subject [19]. The last feature of the Identity Manager interface is the Derive Identifier functionality. It can generate a new identifier from a current stored identity.

B. Mobile – Desktop interoperability

In today's world, smartphones have become a central technology. For this reason, the SEAL service is designed to be accessible from a mobile perspective. The SEAL App provides all the web service functions to cover the users' needs when using their digital identity. This is possible as

the digital identity is linked to the web app via the SEAL servers. Besides, the SEAL App can be combined with a personal wallet for decentralised authorisation and identification process. Thus, the verification and certification of personal data can be done on demand basis.

IV. CONCLUSIONS

The SEAL project addresses security and assurance, the two critical issues of cloud computing in public sectors as described by D. Shin [20]. SEAL-linked digital identity can be used as a multiple identity detector becoming a keystone in the EU ecosystem. Its platform allows both authorities and users to obtain limitless digital resources from authentication and identification mechanisms. Furthermore, features and functionalities of the SEAL service are designed from a user-centric approach, focusing on the final user' benefits and the compliance of the service interoperability with the EU information systems.

The SEAL service offers an easy-to-use user experience, ending up in a better users' interaction of their digital linked identities, giving the users the full control of their single lifetime identity and data across borders - in a trusted and secured manner. SEAL, more than competing with existing infrastructures, tries to fill a gap in the management of identities: the reconciliation of identities. Moreover, it tries to do so in a manner to open the way for standardization of procedures and minimisation of efforts for the implied stakeholders, coordinating the actions of all of them in the most efficient way.

REFERENCES

[1] J. Howlett, V. Nordh, and W. Singer, "Deliverable DS3. 3.1: eduGAIN service definition and policy Initial Draft," GÉANT. Deliverable DS3.3.1, May 2010, URL: https://geant3.archive.geant.org/Media_Centre/Media_Library/Media%20Library/GN3-10-081-DS3_3_1_eduGAIN_service_definition_and_policy.pdf accesed: 2020-11-05.

[2] S. Cantor et al., "Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0," Oasis Standard, March 2005, URL: https://docs.oasis-open.org/security/saml/v2.0/saml-core-2.0-os.pdf accessed: 2020-11-05.

[3] European Commission, "eID Documentation," URL: https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/What+is+eID accessed: 2020-11-05.

[4] T. Fridell, "EMREX The User-centered Solution for Electronic Transfer of Student Data," URL: https://emrex.eu/wp-content/uploads/2020/09/Emrex-EUNIS2019-16-9.pdf accessed: 2020-11-05.

[5] European Commision, "Trust Services and Electronic identification (eID)," URL: https://ec.europa.eu/digital-single-market/en/trust-services-and-eid accessed: 2020-11-05.

[6] R. Sakimura, N. Bradley, J. Jones, M. De Medeiros, and C. Mortimore, "Openid connect core 1.0.," The OpenID Foundation S3, February 2014, URL: https://openid.net/specs/openid-connect-core-1_0-final.html accessed: 2020-11-05.

[7] R. Laborde et al., "A User-Centric Identity Management Framework based on the W3C Verifiable Credentials and the FIDO Universal Authentication Framework," IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), January 2020, doi: 10.1109/CCNC46108.2020.9045440. URL: https://ieeexplore.ieee.org/abstract/document/9045440 accessed: 2020-11-04.

[8] A. Isaac, R. Clavphan, and B.Haslhofer, "Europeana: Moving to Linked Open Data," Information Standards Quarterly, vol. 24, Jun. 2012, pp. 34–40, doi: 10.3789/isqv24n2-3.2012.06 URL: https://www.niso.org/niso-io/2012/06/europeana accessed: 2020-11-04.

[9] European Commission, "European countries join Blockchain Partnership," April 2018, URL: https://ec.europa.eu/digital-single-market/en/news/european-countries-join-blockchain-partnership accessed: 2020-11-04.

[10] European Commission, "Introducing the European Blockchain Services Infrastructure (EBSI)," URL: https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/EBSI accessed: 2020-11-05.

[11] A. Crespo, N. Rodríguez, F.J. Aragó, and V. Andreu, "Feasibility Study on Cross-border Use of eID and Authentication Services (eIDAS compliant) to support Student Mobility and Access to Student Services in Europe," doi: 10.2759/735197. URL: https://op.europa.eu/en/publication-detail/-/publication/c0bc89a9-437b-11e8-a9f4-01aa75ed71a1/language-en/format-PDF/source-69424735 accessed: 2020-11-05.

[12] X. Liu, G. Huang, and H. Mei, "Discovering Homogeneous Web Service Community in the User-Centric Web Environment," IEEE Transactions on Services Computing, vol. 2, nº. 2, 2009, pp. 167–181, doi:10.1109/TSC.2009.11. URL: https://ieeexplore.ieee.org/abstract/document/4912192 accessed: 2020-11-05.

[13] H. L. Moulaison and A. J. Million, "The disruptive qualities of linked data in the library environment: Analysis and recommendations," Cataloging & Classification Quarterly, vol. 52, nº. 4, 2014, pp. 367–387, doi: 10.1080/01639374.2014.880981 URL: https://www.tandfonline.com/doi/abs/10.1080/01639374.2014.880981 accessed: 2020-11-05.

[14] A. Benzekri et al., "Know Your Customer: Opening a new bank account online using UAAF," IEEE 17th Annual Consumer Communications & Networking Conference (CCNC), January 2020, doi:10.1109/CCNC46108.2020.9045148 URL: https://ieeexplore.ieee.org/document/9045148 accessed: 2020-11-05.

[15] uPort Mobile, "uPort Project," Github page. URL: https://github.com/uport-project/uport-mobile accesed: 2020-11-05.

[16] O. Johnston-Watt, "DLT is your passport to work," URL: https://medium.com/blockchaintp/dlt-is-your-passport-to-work-e0183457b54b accessed: 2020-11-04.

[17] E. Miller, U. Ogbuji, V. Mueller, and K. MacDougall, "Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services," Washington DC: Library of Congress, 2012 URL: https://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf accessed: 2020-11-04.

[18] T. Baker et al., "Library linked data incubator group final report," W3C Incubator Group Report, vol. 25, 2011 URL: https://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/ accessed: 2020-11-05.

[19] D.W. Chadwick et al., "Improved Identity Management with Verifiable Credentials and FIDO," IEEE Communications Standards Magazine, December 2019, doi:10.1109/MCOMSTD.001.1900020 URL: https://ieeexplore.ieee.org/document/9031543 accessed: 2020-11-04.

[20] D. Shin, "User centric cloud service model in public sectors: Policy implications of cloud services," doi: 10.1016/j.giq.2012.06.012, URL: https://www.sciencedirect.com/science/article/pii/S0740624X13000099 accessed: 2020-11-05.