Improving the Scalability of Identity Federations through Level of Assurance Management Automation

Michael Grabatin, Wolfgang Hommel, Stefan Metzger, Daniela Pöhν

Abstract:
Access to remote IT services through identity federations (IFs) has solid technical foundations such as the Security Assertion Markup Language (SAML). However, reliable delegated user authentication and authorization also pose organizational challenges regarding the quality management of user data. Level of Assurance (LoA) concepts have been adapted and applied to IFs, but their inhomogeneous proliferation bears the risk of aggravating instead of simplifying the manual work steps required for providing IT services for multiple or dynamically set up IFs. This paper presents a novel LoA management approach that has been designed for a high degree of automation and gives an outlook to its application based on the GÉANT-TrustBroker toolchain.

1 Motivation

Identity federations, such as the German DFN-AAI [DFa], are a well-established key technology to enable users from one organization (the identity provider, IDP) to access IT services operated by another organization (the service provider, SP) without the need to manually set up user accounts on the SP side. By using standardized protocols like the Security Assertion Markup Language (SAML, [Ca05]), SPs can delegate the authentication step of, e.g., the login to a web-based application to the user’s IDP and also request certain user profile data, referred to as user attributes, from the IDP. For example, DFN-AAI can be used by software vendors like Microsoft to make commercial software available to students of German higher education institutions (HEIs) free of charge while ensuring that the downloading user indeed belongs to a certain eligible university and is enrolled as a student there [Re].

SAML and similar protocols have initially been designed with pairs or small sets of IDPs and SPs in mind. They provide the technical basis for inter-organizational authentication and authorization (AuthNZ), but assume that additional organizational or contractual measures are in place to control whose IDP’s users are allowed to access whose SP’s services. However, over the past decade, DFN-AAI and other federations have attracted so many member organizations that, on a technical level and in theory, the users of hundreds of IDPs could use the services of hundreds of SPs. Inter-federations, such as eduGAIN [GÉ], which meanwhile unites 40 national federations including DFN-AAI, and dynamic federation enablers like GÉANT-TrustBroker [PMH14] further contribute to the international and cross-industry-sector use of identity federations with a significant practical impact.

1 Universität der Bundeswehr München, 85577 Neubiberg, [michael.grabatin,wolfgang.hommel]@unibw.de
2 Leibniz Supercomputing Centre, 85748 Garching n. Munich, [metzger,poehn]@lrz.de
With the number of federation partners rising, it becomes more and more complex and tedious to manually configure all the individual IDP ↔ SP relationships. Instead, an SP might want, for example, to accept users from any IDP as long as there is a guarantee that the IDP is a state-approved HEI and the users’ student status is reliably removed whenever a student is disenrolled. Such conditions on selected data quality management aspects have led to the concept of Levels of Assurance (LoA); for example, DFN-AAI distinguishes between the three LoAs (Verlässlichkeitsklassen) Test, Basic, and Advanced [DFb].

Unfortunately, technical support for different LoAs is quite limited in current identity federation standards and products, which basically leads to the situation that DFN-AAI is not one federation, but three – one federation for each LoA. Furthermore, each federation operator can define its own data quality criteria and LoAs, so SPs that accept IDPs from different federations cannot assume that, e.g., IDPs from Germany’s DFN-AAI-Advanced, Switzerland’s SWITCH-AAI, and Australia’s AAF adhere to the same data quality standards. For example, unlike the three German DFN-AAI federations, Australia’s AAF uses four LoAs that are not available as separate federations but as user attribute, i.e., similar to the user’s email address in the mail attribute, her level of identity assurance is available as eduPersonAssurance attribute.

Managing LoAs therefore is crucial for the efficient and scalable use of large federations and inter-federations from an SP perspective as well as key to adequate service access from the IDPs’ point of view. This paper presents a novel LoA management approach that keeps the currently established, diverse solution attempts in mind and leverages common ground with the goal of a high degree of automation in operation; its viability is discussed based on the GÉANT-TrustBroker IDP ↔ SP metadata exchange service. The remainder of this paper is structured as follows: Section 2 compares selected current real-world deployments and summarizes related approaches by research as well as standardization bodies. Section 3 then presents the novel LoA management concept along with its data model and automated comparison workflow. Finally, Section 4 gives an outlook to its application within GÉANT-TrustBroker.

2 State of the Art and Requirements regarding Levels of Assurance

An LoA, such as the DFN-AAI Advanced mentioned above, basically is a set of \{LoA aspect, value\}-pairs. While there are no restrictions regarding which LoA aspects are used to compile an LoA, the following LoA aspects are often used in practice:

**Identification:** How has the user been identified by her IDP? For example, manual verification of a government-issued photo identification document can be considered as more trustworthy than web-based self-registration.

**Data Management:** How up-to-date and precise is the user data? For example, are accounts of retired employees removed immediately or only once per quarter?

**Authentication:** How are users authenticated by their IDP during login, e.g., using only passwords or based on multi-factor authentication methods such as smartcards?
**Assertion Representation:** How is user data transported from to IDP to the SP considering risks such as unintended information leaking and spoofing? For example, are TLS-encrypted connections used and is the data digitally signed by the IDP?

**Accountability:** Which state-of-the-art technical security measures does the IDP have in place to prevent, e.g., unauthorized user data modification?

**Organizational Management:** Which organizational security management procedures are applied on the IDP side? For example, did an independent third party perform a security audit?

For a specific LoA, a specific value needs to be assigned to each of the chosen LoA aspects. The available values and their semantics, i.e., their precise and binding meaning, is typically documented in human-readable prose text approved by legal experts. To enable automated machine-processing, ordered identifier sets are used: For example, in the context of the LoA aspect data management, the values 0, 1, and 2 could be used to distinguish between ”no guarantees”, ”data is no older than 3 months”, and ”data is no older than 14 days”. Obviously, an IDP fulfilling value 2 of this LoA aspect would also be suitable for SPs only requiring value 1. As a consequence, a flexible solution must support arbitrary LoA aspects with arbitrary values; as shown below, we require values to be ordered, i.e., they must be comparable and a higher value must implicitly also fulfill the requirements of lower values.

Several federations operated by national research and education networks as well as standardization bodies have specified LoAs, e.g., NIST in SP 800-63-2 [Bu13], ISO/IEC 29115:2013 [IS13], the electronic identification and trust services (eIDAS) [Eu14] of the European Commission, and the Kantara Identity Assurance Framework (IAF) [Ka]. Typically, 2–4 values are used per LoA aspect and an IDP is assigned to the overall LoA corresponding to the lowest value found in any of its LoA aspects. For example, if a German IDP fulfills all criteria for DFN-AAI Advanced except one, it is assigned to the Test or Basic LoA depending on the value of this one criterion. Although this is a simple and effective assessment procedure, it does not flexibly take into account that SPs’ data quality requirements can be more fine-grained than ”one of three LoA levels”.

Given the rich semantics of the LoA aspects’ values and the lack of a global overall standard, the LoAs used by different federations must initially be compared manually in a LoA aspect-by-LoA aspect manner, determining the semantic equivalence of the available values. For example, regarding the LoA aspect data management, InCommon’s Bronze LoA fulfills the requirements of DFN-AAI Advanced, but no DFN-AAI LoA fulfills InCommon’s Silver LoA requirements; similarly DFN-AAI Basic IDPs are not suitable to use services operated by InCommon Bronze SPs. The results of such manual comparisons lead to a simple mapping of LoA aspect value equivalence, which can later be used to automatically compare LoAs of arbitrary inter-federation IDP ↔ SP pairs.

Especially when there is no individual contract between an SP and an IDP, LoA assignment procedures become relevant in practice. For example, using any of the DFN-AAI LoAs is tied to a legally binding, written contract between the IDP organization and DFN as federation operator, which includes a self-declaration of conformity. Other national research
federations, such as Haka in Finland and SURFnet in the Netherlands, use check-list-style maturity self assessments that must be passed, whereas industry federations, e. g., in the automotive supply chain, typically require formal third-party audits and certification processes.

The primary technical challenge related to this trustworthy and authentic assignment of LoAs to IDPs is about how the IDP’s LoA is communicated to the SP and how it can be verified by the SP; as mentioned in the previous section, there are, in principle, only two approaches that can optionally be combined:

1. LoAs can be included in the federation metadata files: Federation operators provide lists of federation members including metadata such as their communication endpoints, server X.509v3 certificates, and human contact information:
   - The SAML Identity Assurance Profile [Kl10] allows to include multiple URI-styled references to LoA specifications that are fulfilled by an entity, i. e., only globally unique identifiers of LoAs that are described somewhere else are included.
   - Vectors of Trust (VoT) [RJ15], which is intended to become an IETF RFC standard, describes how an LoA’s \{LoA aspect, value\}-pairs can be represented as a simple URN-style text string. Therefore, all LoA information is included and no external documents need to be fetched, but the SP still needs to know how to interpret the LoA aspects and their values.

   When LoA information is included in official federation metadata this way, SPs trusting the federation operator can rely on the correctness of the provided IDP LoAs.

2. LoA information can be included in each communication between an IDP and an SP about a specific user who currently logs into the service:
   - The eduPersonAssurance attribute is used in several research and education federations; its expressiveness is approximately equivalent to the SAML Identity Assurance Profile mentioned above.
   - SAML responses can include the so-called Authentication Context Class statement, which, however, by intention primarily covers the LoA aspect authentication.
   - The Vectors of Trust LoA representation could also be sent as a user attribute, although no real-world deployment of this approach is currently known to the authors.

   Unless any of this information is digitally signed by a trusted third party, the SP needs to trust the IDP regarding the correctness of the provided LoA data.

Although both approaches apparently provide similar functionality, transporting LoA information in the communication between IDP and SP, i. e., outside the federation metadata,
is highly important in practice for IDPs that have different groups of users with different LoAs (see also [TM11]). For example, a university IDP may only achieve LoA 1 for all of its students due to data up-to-dateness issues in the student administration system, but achieve the better LoA 2 for its staff and faculty due to better currentness of data in its human resources management system. LoA information therefore needs to be sent to the SP on a per-user or even per-user-attribute basis to ensure that all eligible users can access the service. Similarly, dynamic IDP/SP metadata exchange solutions, such as the Metadata Query Protocol [Yo15] and GÉANT-TrustBroker, which bypass the organizational overhead of running formal federations, motivate solutions that are not based on federation-wide, federation-operator-signed metadata.

It is furthermore important that currently only the LoAs achieved by IDPs are explicitly stated in federation metadata or SAML-based communication. SPs’ LoA-specific requirements can, for example, in the case of DFN-AAI be derived implicitly from an SP’s federation membership, e.g., DFN-AAI Advanced but not Basic; however, there are no other ways to inform IDPs about SP requirements yet. This can lead to situations in which IDPs send detailed user profiles to SPs although the users are not allowed to access the services; this issue clearly should be addressed when the IDP ↔ SP communication is established.

3 Managing Levels of Assurance with a High Degree of Automation

The previous section has shown that, as of today, different LoA schemes have been proposed and are in real-world use but lack interoperability and apply different methods to transport LoA aspects and values between IDPs and SPs. This paper proposes a technical solution to the LoA comparison core issue: How can be determined whether a specific IDP provides a sufficient LoA for a specific SP in an automated, efficient manner? The proposed solution consists of three contributions that build on top of each other:

1. The SAML Identity Assurance Profile [Kl10] is combined with IETF’s Vectors of Trust [RJ15] to create an LoA data model that is used to encode LoA aspects and their values as simple text string encodings.
2. A specification defines how SPs announce their LoA requirements and IDPs indicate their LoA guarantees as a part of their SAML metadata or individual SAML messages; existing deployments and good practices are accounted for in this step.
3. A workflow describes how software tools can automate the comparison between SP LoA requirements and IDP LoA guarantees either via a trusted third party, such as the GÉANT-TrustBroker service, or locally at each IDP or SP.

The proposed solution works on both, per-IDP and per-user, levels to achieve the flexibility none of the previous approaches alone enables; optionally, it can be applied on a per-user-attribute level and supports digital signatures by trusted third parties for LoA guarantee validation. The following subsections detail each of the three steps outlined above, while Section 4 then gives an outlook to the GÉANT-TrustBroker-based implementation.
3.1 LoA data model to encode LoA aspects and values

On the one hand, the SAML Identity Assurance Profile (SAML-IAP) uses URIs, such as http://foo.example.com/assurance/loa1, to reference unique identifiers of LoA specifications that are documented somewhere else (see [KI10, section 2.3]). On the other hand, IETF’s Vectors of Trust (VoT) uses text strings, such as urn:ietf:param:[TBD]:P1.Cc.A3, to encode LoA aspect identifiers (P, C, and A) along with their values (1, c, and 3) (see [RJ15, sections 4.1 and 4.3]).

The proposed solution basically combines both approaches by creating a SAML-IAP-compliant URI that uses a fixed base LoA identifier, e.g., https://loa.geant.net/gntb, and appends an LoA identifier as well as VoT’s wire representation string as parameters loa and vot; the above example therefore results in


Given the semantical overlap of both parameters, loa and vot, the different semantics of LoAs stated by IDPs and SPs, and the practical necessity to support multiple LoAs in parallel, the semantics of the URI parts exemplified above are defined as follows in accordance with IETF RFC 3986:

- The URI scheme and hier-part (see RFC 3986, section 3), i.e., https://loa.geant.net/gntb indicate the implementation variant of this approach to distinguish it from traditionally used SAML-IAP implementations.
- To state multiple LoAs that are to be used in parallel, multiple URIs must be created and used just like other multi-value attributes in SAML metadata or SAML messages.
- At least one of the two parameters loa and vot must be present. If both of them are used, the values of the LoA aspects encoded in the vot parameter must be the same as or higher than the values derived from the LoA referenced by the loa parameter and override them. This means that the vot parameter can be used to specify either additional LoA aspects that are not covered by the referenced LoA or the over-fulfillment of the referenced LoA, but it must not be misused to express shortcomings in fulfilling the referenced LoA.
- IDPs specify their LoA guarantees, while SPs specify their LoA requirements: When comparing a pair of requirements and guarantees, the requirements are fulfilled if and only if the value of each LoA aspect in the IDP LoA matches or exceeds the value of the same LoA aspect in the SP LoA. LoA aspects that are provided by the IDP LoA but not covered by the SP LoA are discarded. However, LoA aspects specified by the SP LoA but not by the IDP LoA indicate non-conformity, i.e., the requirements are not fulfilled.
- On a per-IDP level, i.e. for use in SAML metadata, IDPs can create multiple URIs with the same loa parameter value but different vot parameter values. This indicates that the IDP has different user groups with non-uniform LoA guarantees that need to be evaluated on a per-user (or per-user-attribute) level.
Based on the discussion of human-readable prose text LoA documents in Section 2, in order to enable automated tool-supported comparisons, the values of the \( \text{loa} \) parameter must either refer to a public document containing a VoT string or the tool needs to have an internal database (i.e., mapping tables) in which the referenced LoA’s aspects and their values are defined, e.g., based on manual creation by a human interpreting the LoA document. The initial creation of these mapping tables cannot be automated unless, for example, ontology-based approaches are used that are outside the scope of this paper.

In order to restrict an LoA to selected user attributes, an optional third parameter named \( \text{attributes} \) can be used. For example, if an SP requires that an IDP provides a high value for the \( \text{LoA aspect} \) data management regarding the user attribute \( \text{mail} \), i.e., the user’s email address(es), but is willing to accept potentially outdated user telephone numbers, i.e., user attributes \( \text{telephoneNumber} \) and \( \text{mobile} \), the relevant parts of the LoA URI would look like:

\[
\ldots\&\text{vot=D2}\&\text{attributes=mail}
\ldots\&\text{vot=D0}\&\text{attributes=telephoneNumber, mobile}
\]

The comma-separated user attributes can be specified either by their human-readable name as in the example above, i.e., more formally, their SAML2 Attribute FriendlyName, or by their globally unique OID, e.g., \( 0.9.2342.19200300.100.1.3 \) for \( \text{mail} \).

### 3.2 Communicating encoded LoA requirements and guarantees via SAML

LoA URIs as specified above need to be embeddable into both SAML metadata and messages. Given the existing standards and practices, the following methods are to be used:

- Based on SAML-IAP, IDPs and SPs can include an arbitrary number of LoA URIs as \( \text{assurance-certification} \) entity attributes in their SAML metadata. This is already in real-world use for IDPs and can be applied to SPs without requiring any changes to the SAML metadata XML schema.
- IDPs that deliver per-user LoA information can transmit LoA URIs via a dedicated, multi-value user attribute. For example, the attribute \( \text{eduPersonAssurance} \) supports URI-style values and is therefore well suitable for this task; while it is the de-facto standard for research and education IDPs, other lines of business will likely choose or create a different user attribute with the same syntax that better fits their specific user data schema.
- SPs currently lack a proper, standardized way of sending full LoA requirements within their SAML requests. SAML-IAP specifies the use of LoA URIs as \( \text{RequestedAuthnContext} \) element in SAML requests, but this element’s original intention was limited to the single \( \text{LoA aspect} \) authentication. LoA URIs as discussed above are compatible regarding their syntax and can be used, but a more generic approach as part of future SAML-IAP versions would be a better long-term solution.
It is worth noting that the only standardized way of providing LoA requirements and guarantees verifiable via a trusted third party (TTP) is by having federation-wide SAML metadata files signed digitally, which is typically done by the federation operator or an entity metadata aggregation delegate. Facilitating per-user TTP signatures is technically possible in principle, but as of today not a realistic option for real-world deployments due to the related organizational overhead; a thorough analysis will be part of our future work.

3.3 Automatically comparing LoA requirements and guarantees

To determine whether an IDP’s LoA guarantees positively match an SP’s LoA requirements, both parties’ LoA URIs must be evaluated and compared to each other. Although this comparison is quite simple in principle, it is non-trivial due to the following four marginal conditions:

1. As both IDP and SP can have multiple LoA URIs, each IDP LoA URI must be compared with each SP LoA URI. The SP’s requirements are fulfilled if there is at least one (IDP LoA, SP LOA)-pair in which the IDP LoA guarantees fulfill the SP LoA requirements.
2. If an LoA URI contains both parameters \texttt{loa} and \texttt{vot}, their \texttt{effective (LoA aspect, value)-pairs} must first be determined by combining both parameters: As specified in Section 3.1, parameter \texttt{loa} determines the basic set of \texttt{(LoA aspect, value)-pairs}, which is then extended or adjusted according to the \texttt{(LoA aspect, value)-pairs} encoded in the given VoT.
3. If both parties use different sets of \texttt{LoA aspects}, e.g., due to applying LoAs from different federations, mapping tables must be applied if available.

In pseudocode, the comparison therefore can be performed as follows:

```
typedef enum {NOT_FULFILLED, FULFILLED} comparison_result;
comparison_result compare_LOA URIs(sp_loa_uris, idp_loa_uris) {
  foreach sp_loa_uri in sp_loa_uris do {
    effective_sp_loa := parse_LoA_URI(sp_loa_uri);
    foreach idp_loa_uri in idp_loa_uris do {
      effective_idp_loa := parse_LoA_URI(idp_loa_uri);
      foreach sp_loa_aspect in effective_sp_loa do {
        if (idp_loa_aspect of same type exists
            or can be derived via mapping table) {
          if (idp_loa_aspect.value < sp_loa_aspect.value) {
            // IDP guarantee does not fulfill SP requirement
            continue with next idp_loa_uri;
          }
          else { // LoA aspect requested by SP is not known to IDP
            continue with next idp_loa_uri;
          }
        }
        return FULFILLED; // All relevant LoA aspects had suitable values
      }
    }
  }
  return NOT_FULFILLED;
}
```
This workflow applies to metadata-based as well as per-request-based LoA URI comparisons. Additionally, if the `attributes` parameter is used in LoA URIs, the comparison must be repeated for each respective set of user attributes. Various optimizations could be made in implementations based on eventually available additional information, such as comparing the best IDP LoA guarantees with the lowest SP LoA requirements first, but as the workflow is neither computationally intensive nor time-critical in today’s real-world deployments, optimizations are out of scope of this paper.

4 Outlook: Application to GÉANT-TrustBroker

In practice, automatically comparing LoA requirements and guarantees is important in the following use cases:

- A specific IDP ↔ SP pair has no other trust-building measures in place, such as membership in a common federation or a written contract.
- LoA requirements or guarantees have changed since they were checked last time.
- IDPs have different user groups with inhomogeneous LoA guarantees.

Unlike traditional national or community-specific identity federations, the TrustBroker service, which is currently being developed by the pan-European research and education network GÉANT, enables the dynamic, user-triggered, on-demand setup of virtual identity federations [PMH14]. As this implies that there are neither central, trusted federation operators nor written contracts in place, automated LoA comparisons take up a key position regarding data quality assurance.

The GÉANT-TrustBroker toolchain consists of a centrally operated SAML metadata repository, in which individual IDP and SP metadata is stored, and plug-ins for IDP and SP software products, such as Shibboleth and SimpleSAMLphp, which retrieve the opposite party’s metadata on demand and integrate it into the local software configuration. Without automated LoA management, only primitive white and black lists of domain or organization names were usable to specify or restrict memberships in the created dynamic virtual federations. By implementing the approach described in Section 3, the central GÉANT-TrustBroker service can verify – to the extent of LoA information embedded in the SAML metadata – whether an IDP’s data quality is sufficient for a chosen SP and can therefore prevent the unneeded exchange of SAML metadata between IDPs and SPs with unfulfilled LoA requirements, which in turn prevents the transfer of users’ personal data without service in return.

Similarly, the GÉANT-TrustBroker plug-ins operated locally by IDPs and SPs can apply the LoA comparison whenever the SAML metadata of either party has changed or per-user LoA requirements and guarantees need to be handled. This also prevents unneeded
personal data leakage and assists the IDP and SP operators by precisely indicating which LoA requirements or guarantees need to be addressed to enable successful interoperation.

The required extensions to the GÉANT-TrustBroker tools will be implemented as part of GÉANT’s EC-funded GN4 project with pilot operations as part of the AARC project.

References


[Ca05] Cantor, Scott; Kemp, John; Philpott, Rob; Maler, Eve (Eds.): Security Assertion Markup Language v2.0. OASIS Security Services Technical Committee Standard, 2005.


