Project Géant-TrustBroker – dynamic identity management across federation borders

Daniela Pöhn, Stefan Metzger, Wolfgang Hommel
Leibniz Supercomputing Centre, Bavarian Academy of Sciences and Humanities
Boltzmannstraße 1, D-85748 Garching n. Munich, Germany
email: {poehn,metzger,hommel}@lrz.de

Paper type
Research paper

Abstract
We present the design and concept for a new service enabling user authentication and authorization management in the research and education environment, called Géant-TrustBroker. Géant-TrustBroker complements eduGAIN, an inter-federation established on top of the national higher education federations in more than 20 countries worldwide by the pan-European research and education network GÉANT, which connects entity subsets of the member federations. Motivated by real-world limitations of eduGAIN, Géant-TrustBroker pairs IdPs and SP in federations, inter-federations and bi-directional set-ups, while reducing the manual workload for the participating organisations. Furthermore, the efforts of converting user information attributes to the format of a service provider is reduced by a conversion rule repository. We contrast Géant-TrustBroker with other state-of-the-art approaches and give an overview of its core workflow and the internal technical architecture.

Keywords
Federated Identity Management, Inter-Federation, Shibboleth, SAML, service provisioning.

1. Introduction
Nowadays many users, e.g., students, employees, collaboration partners or guests, especially from a university or research community, need access to arbitrary ICT services, e.g., email or web collaboration services. This requires often unique identifiers provided along with required information about the current user, i.e., attributes. Within the scope of one organisation these information are typically provided by a centralized Identity & Access Management (I&AM) solution, often based on LDAP or relational databases. Besides a simplified user provisioning and single sign-on function, such an approach helps to divest service access and capabilities much quicker and reduces the risks through misuse of still enabled accounts of former members. However, if an ICT service is provided by another university or research institute, for example, in the context of a cross-organisational research project, this comes with huge challenges to the identity management staff. Mainly two approaches are current pragmatic solutions: Ad-hoc creation of accounts for all users at the external provider’s site, which need access to this service, does not scale well especially in case of larger projects and many project partners. Furthermore, this results in doubled account information, decreases the quality of data as well as increases the complexity for administrators and usually compromises user’s convenience. A much more scalable approach is federated identity management (FIM), where user information are exchanged on demand. These authentication and authorisation infrastructures (AAI) are based on either solution:

- OpenID’s accept-all-comers concept without any formal trust.
- SAML’s rigid bilateral trust model with Identity Providers (IDPs) and Service Providers (SPs).
While many federations in various industrial sectors consist of only a few members, most national research and education networks (NRENs) operate large infrastructures based on SAML with different organisations providing services for users, i.e., SPs, and home organisations, i.e., IDPs. Geographic and industrial sector specific borders for federations is not a requirement of FIM technology itself, though they have become reality due to the historic evolution and growth of FIM’s use in both industry and higher education institutions. For example, DFN-AAI provides an authentication and authorization infrastructure for German universities and research institutes. Most conceptual, technical and organisational issues result from two main characteristics of today’s federation solutions:

- Contracts: membership in a federation usually requires contract signing, e.g., either with all the other federation members in an ad-hoc federation or with a central AAI operator, which can be either a large company or an independent entity as it can be seen in identity networks. IDPs are liable under a contract to, e.g., provide high quality user data to avoid SP misuse based on fake accounts, while the SPs must commit themselves to obey privacy and data protection principles.
- Technology: all members of a federation must use common technology, e.g., SAML.

The next evolutionary step lies in providing services across federation boundaries or to collaborate with external institutions, which are not member of a federation. For the former a growing number of NREN-operated AAI already joined eduGAIN, an inter-federation for cooperation between organisations in different NRENs. Further NRENs, e.g., in Australia, Japan, and Estonia, are eager to opt-in. Analogous to a national federation, a common technological basis must be established to participate in such an inter-federation. In more detail this requires

- Exchanging metadata: IDPs and SPs have to know each other’s communication endpoints. Therefore, all needed information like URLs and server certificates are exchanged via signed, aggregated metadata within the involved federations, but also via a dedicated Metadata Distribution Service (MDS).
- Using a common schema: IDPs and SPs need a common syntax and semantics of user attributes, i.e., a common set of attributes, called schema, must be introduced.

The constant growth of eduGAIN and the arising increased complexity of a nearly fully meshed inter-federation environment results in serious drawbacks. For example, the aggregated XML-based metadata, which currently has the amount of around 30.000 lines of XML code for eduGAIN, becomes cumbersome to process and slows down the servers. Also, take note that involved SPs usually do not need metadata information from other SPs. On the other hand, eduGAIN only standardizes the least common denominator regarding attributes, which practically means that certain information about users may not be available and thus prevents or limits service usage. To deal with such an unsatisfactory situation, research communities like DEIS and PRACE founded their own federations maintaining dedicated IDPs in order to provide all the needed user attributes in a project-specific schema. Therefore, users typically have to manage several identities for their work, which contradicts the single sign-on paradigm and benefits of FIM. In addition, users from external AAI, e.g., employees of research project partners from industry branches, will not be able to get access to an inter-federation service, since the external AAI is not part of the aggregated metadata. Most recently, the setup workflows require a lot of manual administrators’ working steps to configure and filter the attribute release. If attributes vary significantly between the schemas, also manual, SP-specific attribute conversion could be required. Considering the SP side there is also manual work necessary, if its service is accessed by at least one user from another federation. From the user’s perspective, they cannot use new SPs before all these manual setup tasks have been finished successfully and as a consequence often lose interest in the service due to long waiting times and the lack of user-friendlyness.

To overcome and solve all these issues, a complementary approach to eduGAIN has been initiated within GÉANT as part of the GN3plus Open Call projects in order to facilitate inter-federation not only based on a minimum schema. Géant-TrustBroker’s (GNTB) aim of enabling quick technical integration of services from other federations by maximizing the degree of automation will support researchers’ access to external services and increases their efficiency. GNTB automates the technical on demand metadata exchange between currently used SPs and appropriate IDPs across federation borders.

In Section 2, we describe related work, which was also done within the GÉANT community as well as in another federation and inter-federation research projects. Afterwards we contrast these approaches with Géant-TrustBroker in Section 3. Furthermore, we describe the core workflow to establish the technical trust relationship between an SP and previously unknown IDP as well as some variants of this workflow. In Section 4, we put a finer point to the data model and the functions provided by an API used by GNTB, while we give an outlook to our future work in Section 5.
2. Related Work

Many NRENs provide a service for web-based management of SP and IDP metadata. One advanced example is the Resource Registry (RR) of the Switch federation SWITCHaai: RR provides a web-interface for IDPs and SPs to register their metadata. It allows SPs to specify which of the federation schema attributes they actually use, a seemingly very basic information that is, however, not yet provided in most other research federations. In return, IDPs have the option to describe all attributes they actually offer. Despite the wide range of provided functionalities, the web-based metadata management itself requires manual configuration and results in waiting time for the administrator to receive the building blocks for a basic attribute filter, which he can adapt for the IDP-specific configuration.

An early approach of simplifying the discovery of other entities is Dynamic SAML. The metadata consumer validates the signature using a root certificate and establishes the trust, though trust continues to lie in pre-established contractual arrangements. Despite the dynamic character, the entities have to manually convert user information, which are exchanged, or use a data schema that is the common denominator.

The IETF Draft for Metadata Query Protocol by Young proposes to retrieve metadata from entities using HTTP GET requests. Therefore, it solves the problem of huge aggregated metadata files, but otherwise has the same drawbacks as Dynamic SAML: manual work and the initial trust establishment.

The Metadata Query Protocol is part of the Metadata Exchange Protocol (MDX), where entities pick a registrar for their metadata and receive attributes from partner entities from one or more aggregators. In analogy to the DNS protocol, aggregators and registrars are linked in order to exchange metadata with each other.

Similar to MDX, the Public Endpoint Entities Registry (PEER) project implemented a public endpoint entities registry supporting both SAML and non-SAML protocols. Though PEER moves from a huge metadata aggregator to a central system, where administrators can register their domain, many manual steps are still needed, e.g., downloading and embedding an SP’s metadata file.

Dynamic Identity Management and Discovery System (DIMDS) has the purpose to achieve minor user involvement in the identity management by creating DIMDS accounts. All user attributes are stored unencrypted in the central system, which can affect the privacy of users. Furthermore, DIMDS does not distinguish between IDPs and SPs, though not all IDPs can act as SPs and vice versa, which also appears in Federated Attribute Management and Trust Negotiation (FAMTN), where internal users are supposed to perform negotiations by exploiting their single sign-on (SSO) ID without repeating identity verifications, though the SSO ID can be misused for attacks. It might appear that a provider needs less or more attributes, leading to violations of data minimization or further negotiations between providers.

IdMRep shifts from pre-configured cooperation to dynamic trust establishment by a distributed reputation-based mechanism based on local Dynamic Trust Lists (DTLs) and external reputation data. DTLs can receive recommendations from other entities, as long as an entity is not a new member of a federation or inter-federation. This results in another bottleneck, because of the amount of data processing required for all external and internal trust information, especially in an inter-federation use case. Furthermore, the problem of different attributes, syntax, and semantics is not considered.

In contrast, the proposed solution of the Credential Conversion Service for eduGAIN (eCCS) focuses on the conversion of credentials. eCCS makes use of a special credential conversion service, which translates source credentials into target credentials, based on attributes from the Schema for Academia (SCHAC) and eduPerson schema. Though conversion rules within the inter-federation eduGAIN are concurrently written manually, the proposal concentrates on the two schemas SCHAC and eduPerson.

All approaches concentrate on one particular part of the problem, but not the situation as a whole. Furthermore, they have in common the manual steps needed to set up the exchange of metadata respectively the limitation of the minimum schema.

3. Technical Solution

Put simply, Géant-TrustBroker is an on-demand repository for SPs’ and IDPs’ metadata and attribute conversion rules, which simplifies the discovery of other entities and automates the establishment of technical trust. It enables the exchange of user information across federation borders by retrieving the metadata on-demand and automating the therefore needed technical configuration steps on the IDP. As a consequence, it drops down the scalability issue of metadata release of other often seen approaches. At the same time, it significantly reduces the workload of manual steps, which are concurrently needed for establishing the trust between two provider entities.

The exchange of attribute conversion rules via GNTB data conversion rule repository enables more complex and project specific data schemas, while IDPs can re-use available data conversion rules if they fit. Therefore, not only the minimalistic schema of eduGAIN is supported, but also community- or project-internal schemas, which
provide further user information. GNTB’s design have a focus on SAML, which is the predominated standard in the R&E environment, but it could be extended for the use of other protocols, like OpenID very easily. Mainly we distinguish between two workflow types at GNTB:

- **Core workflow:** establishes the technical trust relationship between an SP, which one user wants to access, and an appropriate IDP. In a break from tradition the users themselves triggers the exchange of metadata.
- **Various management workflows:** allow SPs and IDPs to register, update, and delete their metadata as well as attribute conversion rules. To simplify the design of the core workflow, the metadata registration step has to be finished before the GNTB core service can be used. However, metadata registration could also be integrated in the core workflow and triggered by the users in future versions. Before an entity is registered, GNTB verifies the ownership of the domain and the validation of the metadata. The metadata management workflows also supports regular updates. Because of the limited validity period of the included certificate information. The rule management workflows allows the update and deletion of rules. The IDP administrators, which create the conversion rules, are permitted to update or delete only their own rules.

Currently we develop a prototype for Shibboleth, the de-facto standard in R&E, which extends IDPs and SPs with GNTB specific modules in order to enable both workflows.

### 3.1 Core Workflow

The design of the core workflow, as shown in Figure 1, is close to the practically implemented regular SAML workflows in order to seamlessly integrate GNTB in current implementations and federations.

1. Assume the user Alice from IDP I in federation F1 wants to use a web-based application service, e.g., a project wiki, from SP S in another federation, F2, which is not in eduGAIN.
2. The authentication form at S presents Alice, as often seen in usual FIM scenarios, a static list of already trusted IDPs. In a more dynamic way, an IDP Discovery Service as provided by the SAML-based Shibboleth could be used. Since I and S have no bilateral relationship established yet, Alice cannot choose I from this list directly, but because S is registered at GNTB, Alice can trigger the GNTB core workflow, e.g., via a button on the SP website.
3. Alice is redirected to the GNTB website by standard SAML mechanisms. From the list of registered IDPs, Alice picks her IDP I.
4. GNTB has to pass the IDP information back to S to have S determine whether the user-chosen IDP I is acceptable. In the case I is blacklisted, e.g., provider facebook.example.com is not acceptable for a specific service, the core workflow will be aborted.
5. Afterwards S sends the initial SAML user authentication request to GNTB, which temporarily stores it, rather than to the chosen IDP. This intermediate step is necessary to have GNTB authenticate Alice via the IDP in order to prevent malicious users to add arbitrary IDPs’ metadata to any SP and vice versa. This marks a difference to the normal SAML Discovery Service workflow.
6. In the next step GNTB redirects Alice automatically to her chosen IDP I for authentication. During this step, GNTB acts like an SP towards I, which also extends the normal SAML workflow.
7. Assuming that also S is acceptable from I’s point of view and Alice has been authenticated successfully, I fetches the metadata for S from GNTB.
8. I updates its metadata configuration automatically, reducing the former manual workload based on S’s metadata.
9. Because S and I are not members of the same federation, their schema differs. Therefore, appropriate attribute conversion is required. I checks whether suitable rules are available at GNTB. Based on such rules, I’s local attribute resolver configuration is updated automatically. This is the basis for the creation of appropriate attribute filters, i.e., definitions, which user attributes I will send to S on request.
10. I can immediately send a SAML authentication assertion and Alice’s browser is redirected back to S.
11. Because SAML assertions usually have to be signed by the sending entity, S requires and fetches I’s metadata from GNTB, which includes the public key(s) in order to verify the signature. For processing further authentication requests, S stores I’s metadata to its local configuration.
12. In the last step, S requests a SAML attribute assertion providing detailed, but filtered user information. This could also be part of the SAML authentication assertion sent in the previous step. If the provided attributes, e.g., membership of a certain faculty, meet the requirements, Alice gets access to the web service of S.

After the initial technical metadata exchange between the SP and an IDP, GNTB is not involved anymore and therefore does not interfere with existing entity configuration using other add-ons. This enhances also privacy
protection, because GNTB does not know, whether and how often Alice or another user from her IDP will use a specific service of a certain SP. However, GNTB supports retrieving updated metadata automatically. This can be done by provided pull or push mechanisms. In pull mode, IDPs regularly check if updated metadata is available at the GNTB repository and fetch this file or part of updated information automatically. GNTB also will provide a push mechanism to transfer updated metadata automatically to these providers, which holds a common trust relationship.

Figure 1: GNTB Core Workflow
3.2 Variants of the Core Workflow

Additionally to the GNTB core workflow, there are several variations to handle special conditions, e.g.:

1. IDP and SP are members of the same federation.
2. IDP is connected to subordinate Attribute Authorities (AA).
3. SP honours the GÉANT Code of Conduct (CoC).

If both the SP and the IDP are members of the same federation (1), e.g., SP and IDP are members of DFN-AAI, the core workflow is simplified, as shown in Figure 2. This is the case because no attribute conversion will be necessary and the required metadata will already be available at both the SP and the IDP. However, the IDP may not yet have been manually configured to send all required user attributes to the SP, so at least the IDP administrators could benefit from the automation that can be achieved using the GNTB workflow. Furthermore, the first creation of a federation could be simplified by collecting and retrieving metadata of other entities.

![Figure 2: Intra-federation use case (1)](image)

A growing number of SPs require information about users, which the IDP cannot provide without querying a third party, i.e., AA (2; see Figure 3). For example, the use of high performance computing resources via Grid middleware has been FIM-enabled, but Grid SPs typically require user attributes, e.g., Grid user certificate distinguished names (DNs), which most universities do not store in their central I&AM system. Therefore, the university’s IDP cannot provide this information. However, the IDP can retrieve these attributes from an AA. In this case, attribute conversion is necessary if the AA does not use the same data schema as the SP.

![Figure 3: AA use case (2)](image)
GÉANT recommends the use of the Code of Conduct (3) in order to support the trust establishment. The CoCo is a set of privacy and data protection obligations that are closely related to European data protection acts. Its basic idea is that SPs can signal that they honour the CoCo. IDPs then can be configured to send personal user data to the SP without the formal requirement of written contracts that govern data protection measures. This variation covers the use case that an IDP wants to check whether the SP honours the CoCo, for example, to either reject SPs that do not use CoCo or at least require a manual approval step.

3.3 Requirements Analysis

After specification of the workflows, the implementation requirements have to be identified. These are divided into the components SP, IDP, GNTB. An account choosing functionality is defined based on the different workflows and their variants, as described above. GNTB itself supports the metadata and attribute information management workflows. In more detail, this is related to the notification about successful or failed management tasks, i.e., registering, uploading, deleting or updating. Therefore, GNTB needs, e.g., repositories for metadata information and conversion rules, as shown in Table 1. These repositories are described in the following section. Put simply only authenticated and authorized people should have access to GNTB management interface, so we need a preceding user registration, user management functions and authentication mechanisms. The software plugins for the SP and IDP software have to support the core workflow, e.g., fetching metadata and integration in the local configuration and automatic restarting server processes to activate the new settings.

<table>
<thead>
<tr>
<th>ID</th>
<th>Short description</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNTB-REQ1</td>
<td>Metadata / attribute information repository</td>
<td>GNTB needs a database/repository to store metadata/attribute information of SPs and IDPs</td>
</tr>
<tr>
<td>GNTB-REQ2</td>
<td>Conversion rules repository</td>
<td>GNTB needs a database/repository to store conversion rules of IDPs</td>
</tr>
<tr>
<td>GNTB-REQ3</td>
<td>Management tasks should only allowed by authenticated/authorized IDP/SP administrators</td>
<td>Authentication / Authorization, maybe also previous registration of IDP/SP administrators</td>
</tr>
<tr>
<td>GNTB-REQ4</td>
<td>Mechanism to push updated metadata information of SPs</td>
<td>Which IDP needs an update of this metadata?</td>
</tr>
<tr>
<td>GNTB-REQ5</td>
<td>Triggering the IDP to fetch metadata/attribute information of requested SP</td>
<td>Prerequisite: Which component discovers user’s IDP?</td>
</tr>
<tr>
<td>GNTB-REQ6</td>
<td>Cleaning-up procedures to remove (very) old, long-time not used SP’s data</td>
<td>GNTB deletes old, long-time not used data</td>
</tr>
</tbody>
</table>

Table 1: Requirement categories for GNTB

4. Data model and API

The Géant-TrustBroker service is the central part of our approach and important for establishing technical trust between two entities and reusing conversion rules. It consists of repositories for the metadata and the conversion rules, but also of a data model, which stores all related data. The data model includes a multi-federation namespace that is the basis for registering the list of user attributes required for using the service, while a data access layer facilitates the registration of entities, users, or uploading conversion rules.

4.1 Data model

Metadata enables exchanging information about the communication endpoints and ensures the authenticity of the sending entity. Therefore, this kind of information needs to be stored centrally at GNTB. A technical implementation imposes particular requirements on an interface for up- and downloading signed metadata files, the possibility to extract information from these files, and some kind of version control. We investigated other resource registry solutions for federations different approaches exist, e.g., the

- SWITCHaai Resource Registry tool, which makes use of a relational database management system;
- DFN-AAI, which stores the metadata files directly in the file system using PHP and XSLT processing
The database schema consists of the following information management is implemented.

- PEER project, which integrated a version control system.
- Alternatively a high-performance XML database, e.g., eXist, would be an option. As GNTB should provide its service to different federations and communities, the metadata content varies. This makes the sole usage of a relational or XML database very complicated for reproducing metadata in a simple, efficient way. In this case, we had to extract information from the signed metadata file, like entity URLs, certificate and key information as well as technical or support contact persons. If an IDP requests the metadata, GNTB has to compose a new metadata file on behalf of the appropriate provider entity and sign it before transfer. This violates the authenticity requirement of the original sending party.

On the other hand, a version control system adds additional value to the service, because, e.g., more than one metadata file could be uploaded and provides, in the case of renewing a certificate, the opportunity to activate the new one shortly before the old will become invalid. In addition, changes to the metadata file could be recognized very easily to inform IDP administrators about an update.

Since GNTB needs further information about an entity, e.g., to which organization a provider entity belongs to, in which repository container its metadata file is stored, or to record its current status. If an IDP requests the metadata, GNTB has to compose a new metadata file on behalf of the appropriate provider entity and sign it before transfer. This violates the authenticity requirement of the original sending party.

The database schema consists of the following information, as described in Figure 4:

- Organizations: Each organization consists of one or more IDP or SP provider entities.
- Providers: Besides its unique name (entity ID), entity type (i.e., IDP, SP) and its current status (e.g., valid, invalid or deactivated), information about the last attribute change and the location of the metadata file are stored.
- Users: Information about authorized users, like a unique username, hashed password, given name, surname, and email address of the contact person and their technical role. In SAML-based environments, three predefined roles exist. An administrative, support, and technical role. Some providers assign each role to one specific person or team, others have only one contact person for the entity or the whole organization.
- Conversion rules: This table contains metadata about the attribute conversion rules, like description, its owner (e.g., the IDP, which uploads the rule), timestamp of the last change, its current status, and the location of the rule file in the versioning system.
- Groups: Information about communities and federations, which can be the target or source group for conversion rules. Target means that the conversion rule can be re-used by one or several groups of IDPs or, respectively, one single IDP. Source is the opposite: the conversion rule was written for the needs of one specific SP or group of SPs, which all require identical attributes. One entity could be member of several groups, e.g., one federation, one inter-federation, and several projects.
- Some relationships between
  - IDPs and SPs
  - IDPs and conversion rules
  - Rules and groups.

The relationship between an IDP and an SP indicates a successful technical trust establishment using the GNTB core workflow described above and has to be stored at GNTBs database. This is an important information in the case of updating metadata to inform all administrators and to pull or push updated files to the provider entities.

Analogous the information about the relationship between an IDP and conversion rule enable GNTB to notify the responsive contact persons about rule changes. The last table contains information about a specific rule and by which target (single entity or group of entities) it is re-used. Thus, we have to store only one copy of a conversion rule in the repository and not one for each IDP-to-SP relationship.
4.2 Data access layer

The database tables are filled up by different API functions provided by GNTB data access layer (GNTB API). These functions can be split into three main categories: account handling, provider entity handling, and conversion rule handling.
4.2.1 Account Handling
As GNTB requires authentication of users as a precondition before any metadata or rule related configuration is possible, to prevent successful malicious intent of the user, some kind of user management is essential. For Géant-TrustBroker a basic access authentication (HTTP authentication) based on username and password, as often seen in FIM environments, is acceptable from an information security perspective. However, to avoid plain-text personal user credentials in IDP- or SP-located scripts, a certificate based authentication method will be supported in future versions. User management provides the required functions to support the generally known lifecycle phases (e.g., creation, update, and deletion) of a GNTB account. API functions can be used to add, update, or delete the username, password (i.e., its hash value) and the optionally stored givenName, surname and emailAddress entries of the entity administrator to the database.

4.2.2 Provider Entity Handling
At the first time contact between an entity and GNTB, metadata information is usually not yet registered except if it has been automatically taken over from a federation or inter-federation, e.g., by a central Metadata Distribution Service as it could be an option in the inter-federation eduGAIN. The registration procedure is possible either URL- or file-based. In the first case, an entity is registered by providing its metadataURL to fetch a metadata XML file from there; otherwise, an entity provider uploads this file manually. In either case, the uploaded metadata is validated by specific XML Schema Definition (XSD) files, which validate the structure of the XML file. The ownership can be confirmed by HTTP validation, i.e., creating a resource in the root of the HTTP service for the domain with the name of a random parameter string given by GNTB, certificate validation of the uploaded metadata, or by simply verifying educational domain names by email. As described in the core workflow above, a SP typically checks its trust relationship to the IDP chosen by the user. The API provides an appropriate

gntb Ent_CheckTrustToIdp(Entity_ID[SP], Entity_ID[IDP])

function for this purpose. If an IDP is considered acceptable, the core workflow continues to establish the technical trust by invoking the API function

gntb Ent_EstabTrust(Entity_ID[SP], Entity_ID[IDP])

which inserts an entry in the relationship table mentioned in the previous section. The administrators can set up further options, like sending notification if metadata has been changed or about a near certificate expiration, update its metadata, and delete technical trust relationships.

4.2.3 Conversion rule handling
Service providers may expect attributes, which may not be part of the IDP’s schema, i.e., the IDP cannot provide these attributes out of the box. In order to send them, IDP administrators utilize so-called user attribute data conversion rules, which will be used to extend the local attribute-resolver.xml definition of Shibboleth IDPs.
In the first step, a DataConnector from an IDP-internal data store, a LDAP-based directory service or a relational database pulls raw attributes. In the next step, these attributes are prepared for release in an attribute definition consisting two parts: the definition of the attribute itself and the so-called conversion rules. Typical attribute conversions encompass

- Renaming: the attribute is used with the same format, but another name. A simple example of renaming a source attribute gecos to a new displayName attribute would look like this:

  `<resolver:AttributeDefinition id="displayName" xsi:type="Simple"
  xmlns="urn:mace:[...]">
  sourceAttributeID="gecos">
  <resolver:AttributeEncoder[...]
  name="urn:mace:dir::displayName" />
  <resolver:AttributeEncoder[...]
  name="urn:oid:2.16.840.1.[...].3.1.241" friendlyName="displayName" />
  </resolver:AttributeDefinition>

- Transforming: this is typically used for attributes which indicates timestamps or dates, e.g. birth date if
the internally format is different from that of the SP. This allows the conversion from an international, ISO 8601-based format to a national format. Also different separator signs, like slashes or dashes are supported.

- Splitting: complex regular expression can be used to extract partial information from an attribute. This allows for example to extract a role name Administrator stored as an LDAP attribute with its usually seen distinguished name format (OU=Administrator, O=<company>, …)
- Merging: inter-connects two source attributes, e.g., givenName and surname, into a new one, e.g. commonName.

These conversion functions can be cascaded as well, i.e., one rule to prepare the attribute for internal use, then another one referencing the internal rule for the federation or communities schema, which can afterwards have a dependency to another schema. Administrators can re-use those conversion rules by utilizing the GNTB conversion rule repository. IDP administrator can upload XML files or search for reusable ones created by another IDPs. The data access layer provides the function

\[
gntb\text{ Conv\_FetchRule(Name)}\]

to download an appropriate conversion rule identified by its unique name. This can be, for example, used in a web based user interface providing a dropdown menu containing available rules for a specific provider entity. After downloading, the definitions within the rule are added to the local configuration, i.e., metadata-resolver.xml, by scripts, which are part of the software plugins GNTB provides. Due to the fact, that XML include tags would not work according to Scott Cantor in Shibboleth Users, as XInclude requires schema support in the original schema to mark where things can be included, and we want to reduce the manual work for administrators, local assembling of configuration files is reasonable. As described in section 4.1, we define source groups, which is one SP or a group of SPs needing specific attributes. The target group is one IDP or a group of IDPs, which can use this conversion rule. As mentioned above, conversion rules can be applied to different sources and targets, if a rule is applicable for different groups, i.e., federations, communities or project partners as well. For example, one conversion rule named 01203_2.xml was written from an IDP I belonging to the federation DFN-AAI, which is an example for a target group. For instance, if the source of a rule was SP B in the federation SWITCH-AAI, an IDP C in the Austrian federation ACONet, noticing that the rule 01203_2.xml can be used for their federation as well, can fetch this rule, updated its configuration and therefore this Austrian IDP C applies rule 01203_2.xml for the Austrian federation.

5. Protocol specification

Besides the workflows for metadata and attribute conversion rule management a protocol specification needed for the trust establishment between provider entities using GNTB (core workflow) is designed. As described in section 3.1 the core workflow is close to SAML, therefore GNTB needs to extend it. Based on a regular SAML workflow (Web-browser SSO Profile), GNTB works similar as an IDP Discovery Service / Where are you from (WAYF) service, which is usually based on the Identity Provider Discovery Service Protocol and Profile, we show the typical SAML workflow in Figure 5, its steps (numbered arrows) and the four types of protocol messages:

- HTTP GET request (colour: grey)
- HTTP POST message (colour: purple)
- HTTP 302 Redirect message (colour: green)
- HTTP 200 OK message (colour: orange)

Following the Identity Provider Discovery Service Protocol and Profile allows within the HTTP GET request, which redirects the user agent to the Discovery Service URL, to interactively select an appropriate IDP to specify a few optional parameters, which we investigate. One of these is the Boolean value isPassive, which indicates whether the discovery service is allowed to visibly interact with the user agent in the next communication steps or not. Additional redirection is permitted, but there must not be any interaction with the user, which is contradictory to the user-centric approach. GNTB is planned to redirect information about the user chosen IDP back to the SP. If this IDP is acceptable from the SP’s perspective then the SP sends a SAML AuthnRequest message to GNTB instead of the IDP. This first difference between SP and Discovery Service interaction of the normal SAML workflow and the GNTB workflow enables GNTB to store the user authentication request temporarily, i.e., GNTB acts like an IDP. This could be done by on board means of SAML.
Redirecting the request to the appropriate IDP requests user’s authentication. If successful, the IDP’s software plugin is triggered to fetch metadata information. If the configuration has been finished successfully, the stored authentication request will be responded by GNTB and the authentication and attribute assertions will be sent.

In addition, it is possible to rely on saved state, such as HTTP cookies to determine the appropriate identity provider (!_idp_sami_).

![Diagram](image)

Figure 5: User Login at SP with use of Discovery Service

### 6. Current status and outlook

Géant-TrustBroker enables the on-demand, user-triggered exchange of metadata and user attribute data conversion rules across identity federations’ and inter-federations’ borders. The scalability of the metadata exchange in federations and inter-federations is further improved. The so far aggregated metadata file is not needed any longer. GNTB supports the fully automated technical setup of FIM-based user authentication and authorization (AuthNZ) data exchange and therefore increases the automation of the former manual implementation by SPs and IDPs. Consequently, users can immediately start using a new service outside of their federation and have no waiting time until the administrators have finished the manual setup process. The further needed protocols between IDP and the GNTB for the core workflow as well as for the conversion rule exchange will be designed. Furthermore, the Géant-TrustBroker core workflow will be formally specified as an IETF Internet-Draft and submitted for standardisation as IETF Request for Comments (RFC). The GNTB prototype and implementation of the workflows for the FIM software package Shibboleth will be made available as open source and used for pilot operations in 2015.

Further research questions relate to the combination of technical and behavioural trust for the establishment of dynamic virtual federations as well as to quality assurance, e.g., measure for Level of Assurance (LoA) guarantees, of entities in the dynamic virtual federations.

### Acknowledgement:

The research leading to these results has received funding from the European Community’s Seventh Framework Programme under grant agreement nº 605243 (GN3plus).

### References


GÉANT, 2014. eduGAIN Homepage. Available through:
G UIDELINES – FORMATTING INSTRUCTIONS FOR FINAL PAPERS FOR PUBLICATIONS

20 January 2014

<http://www.geant.net/service/eduGAIN/Pages/home.aspx> [Accessed 7 April 2014]


Biographies

Daniela PÖHN received a university diploma degree in Computer Science from the University of Hagen, Germany, in 2012. She was engaged in the IT industry as a full-time software developer during her studies, before she joined LRZ as a Ph.D. candidate in September 2012. She is involved in the identity management research activity (JRA3) in GÉANT3+ since April 2013.

Stefan METZGER received a university diploma degree in Computer Science from Technical University Munich. Before he joined LRZ in July 2009, he oversaw diverse risk and compliance management projects as a security consultant. Holding a CISSP certification, his focus as a Ph.D. candidate is on information security as well as identity management in inter-organizational environments.

Wolfgang HOMMEL is a research group leader at Leibniz Supercomputing Centre. He has a Ph.D. as well as postdoctoral lecture qualification from Ludwig-Maximilians-University, Munich, where he teaches information security lectures and labs. His research focuses on information security, IT service, and network management in complex large-scale and inter-organizational scenarios.