AARC Blueprint Architecture 2019

Publication Date 2019-11-06
Authors: AARC Community members;AppInt members;Nicolas Liampotis (ed.)

Document Code: AARC-G045
DOI: https://doi.org/10.5281/zenodo.3672785

The research leading to these results has received funding from the European Community's Horizon2020 Programme under Grant Agreement No. 730941 (AARC2).

Abstract

The AARC Blueprint Architecture (BPA) provides a set of building blocks for software architects and technical decision makers who are designing and implementing access management solutions for international research collaborations. This document describes the evolution of the AARC Blueprint Architecture, starting with a summary of the changes since AARC-BPA-2017. The current iteration of the BPA focuses on the interoperability aspects, to address an increasing number of use cases from research communities requiring access to federated resources offered by different research and e-Infrastructures. Hence the introduction of the Community AAI, which streamlines researchers' access to services. These typically include services offered to members of a specific community, as well as infrastructure services that may be shared with other communities. Users can authenticate to the Community AAI primarily via institutional credentials from national identity federations in eduGAIN, but, if permitted by the community, can also use other Identity Providers.

Table of Contents

1. Introduction

The AARC Blueprint Architecture (BPA) provides a set of interoperable architectural building blocks for software architects and technical decision makers, who design and implement access management solutions for international research collaborations.

The new developments and initiatives (such as the European Open Science Cloud) are striving for long-term sustainability, interoperability, and seamless access across resource providers, including Research and e-Infrastructures (or Infrastructures). The BPA became a de-facto model of facilitating access to resources between researchers, however, further guidance and developments is needed, in order to fulfil the stated goals of "reducing the fragmentation of the research and innovation ecosystem" [ESFRI].

Currently, research collaborations use an AARC BPA-compliant AAI for managing access to their resources. Access to resources protected by a different AAI should be seamless for researchers, i.e. it should not require re-registration of users, or any additional management of user attributes by the other AAI. Therefore, this document focuses on the interoperability of AAIs that are operated by different research and e-Infrastructures. This functionality is needed by research communities requiring access to federated resources offered by different infrastructure providers. Hence the introduction of the Community AAI that aims at streamlining how researchers can access services across different infrastructures. Researchers continue using primarily their institutional credentials from national identity federations in eduGAIN, but also from other sources, as allowed by the community. This information is typically enriched with authorisation information managed by the Community AAI.

The remainder of this document is structured as follows: Chapter 2 provides relevant terms and definitions. Chapter 3 describes the layers of the latest iteration of the BPA and provides an overview of the revisions since AARC-BPA-2017 [AARC-G012]. Chapter 4 describes the interoperability architecture for enabling access across infrastructures via the Community AAI. Finally, in Chapter 5, conclusions are drawn.

2. Definitions

Community: A group of users, organised with a common purpose, and jointly granted access to resources. It may act as the interface between individual users and the resources. (see also [WISE-SCI])

Digital identity: Information that represents an entity (subject) within a domain. It contains information about the subject's attributes and relationships.

Community identity: A user's digital identity that may be enriched by the community with additional attributes such as a shared user identifier, profile information, and community attributes such as group membership and role information (see [REFEDS-R&S] and [REFEDS-Sirtfi]).

AAI service: A service that enables authenticated and authorised access to resources.

Community AAI: An AAI service that also enables the use and management of community identities for access to resources. It comprises three (3) AARC BPA component layers: the Access Protocol Translation, the Community Attributes Services, and the Authorisation.

Infrastructure Proxy: An AAI service of a research infrastructure or e-Infrastructure (hereafter termed infrastructure) that enables access to resources offered by Service Providers connected to that infrastructure. This AAI service does not provide community membership management[1]. Specifically, the Infrastructure Proxy comprises two (2) AARC BPA component layers: the Access Protocol Translation and the Authorisation.

Generic service: A service provided to users, possibly as members of different communities.

Community service: A service provided to members of a specific community.

Infrastructure service: A service provided by a research infrastructure or e-Infrastructure to members of one or more Community AAI which receives the required attributes through an Infrastructure Proxy.

[1] The Infrastructure proxy, as any other end service, can apply its own authorisation process concerning its protected resources. As such it may add Infrastructure specific attributes to the incoming identities. In other words, authorisation to access services may be based on a combination of information coming from more than one source, including the Infrastructure proxy.

3. Architecture layers

The latest iteration of the AARC BPA (AARC-BPA-2019) includes five (5) component layers: User Identity, Community Attribute Services, Access Protocol Translation, Authorisation, and End Services.

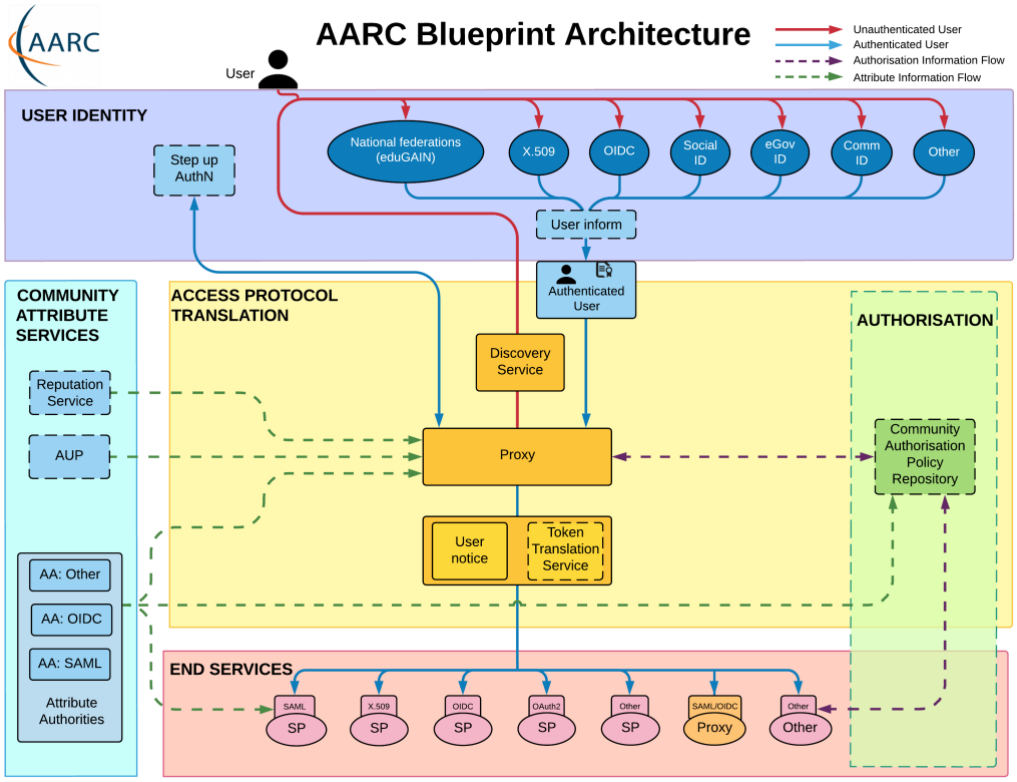

Figure 1: Component layers of the AARC Blueprint Architecture (AARC-BPA-2019)

As illustrated in Figure 1, each of these layers includes one or more functional components, grouped by their complementary functional roles:

- *User Identity* - Contains services for the identification and authentication of users. In existing implementations in the research and education space, these services typically include Security Assertion Markup Language (SAML) identity providers, certification authorities and, more recently, OpenID Connect (OIDC) or OAuth2 Providers (OPs). Although the focus of the services in this layer is to provide user authentication, often some end-user profile information is released as part of the authentication process.
- *Community Attribute Services* - Groups components related to managing and providing information (attributes) about users. Typically, this information includes community group memberships and roles, which is added on top of the information that might be provided directly by the identity providers from the User Identity Layer.
- *Access Protocol Translation* - Addresses the requirement for supporting multiple authentication technologies. It includes the following services:

    - SP-IdP-Proxy (proxy), which serves as a single integration point between the Identity Providers from the User Identity Layer and the Service Providers in the End Services Layer. Thus, the proxy acts as an SP towards the Identity Federations for which this proxy looks like any other SP, while towards the internal SPs it acts as an IdP.
    - Token Translation Services, which translate identity tokens between different technologies.
    - Discovery Service, which enables the selection of the user's authenticating IdP.
    - User notice, which allows users to be informed regarding the processing of their personal data
- *Authorisation* - Contains components for controlling access to the End Services Layer. The AARC BPA allows the implementers to delegate many of the complex authorisation decisions to central components, which can significantly reduce the complexity of managing authorisation policies, and their evaluation for each service individually.
- *End-services* - Contains the services users want to use. Access to these services is protected (using different technologies). These services can range from simple web-browser-based services, such as wikis or portals for accessing computing and storage resources, to non-web-browser-based resources such as APIs, login shells, or workload management systems.

3.1. Revisions since AARC-BPA-2017

The current version of the AARC blueprint architecture builds upon the previous one [AARC-G012] (depicted in Figure 2), while retaining full backwards compatibility. As shown in Figure 2, it retains the same five layers, each of which includes one or more functional components, grouped by their complementary functional roles. The User Identity Layer, the End Services Layer and the Authorisation Layer are still there, while the User Attribute Services Layer has been renamed Community Attribute Services Layer (see definition of Community Identity in the Glossary) and the Identity Access Management (IAM) Layer has been renamed Access Protocol Translation Layer and retains its prominent role in the architecture. Within the Access Protocol Translation Layer, the layout of the Token Translation Service (TTS) has been updated to better visualise the role of the TTS in the flow of attributes between the proxy and the connected services.

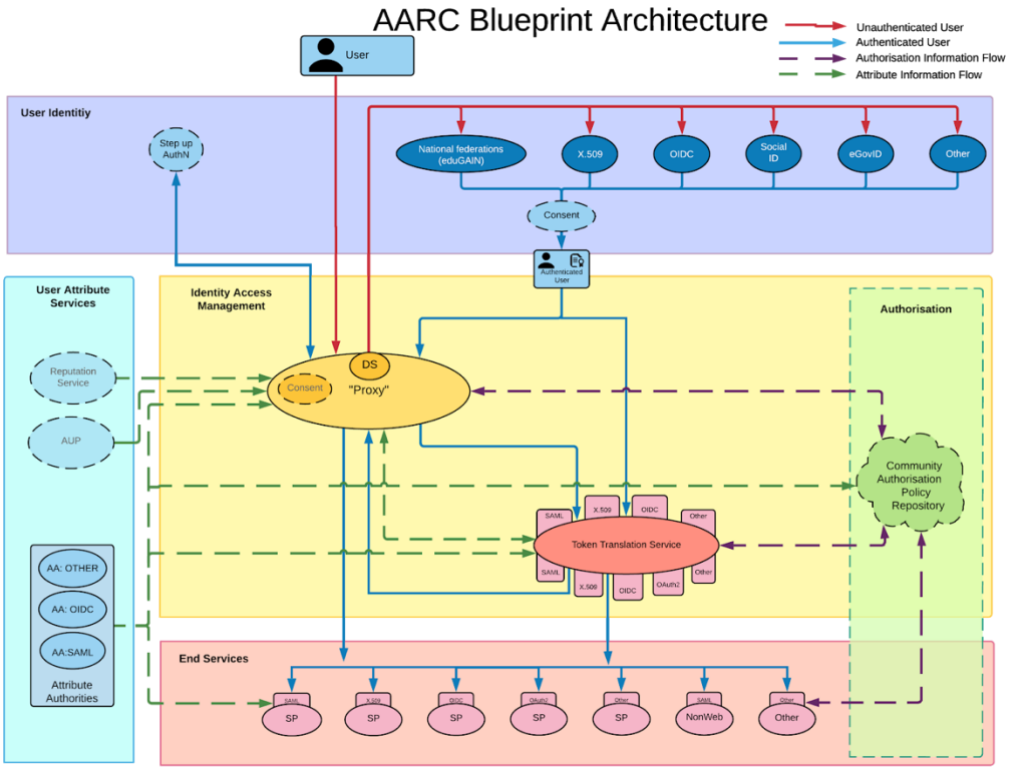

Figure 2: Component layers of the AARC Blueprint Architecture (AARC-BPA-2017)

It is worth noting that, in the new version of the architecture, "User consent" has been renamed "User notice" to indicate the points where users (data subjects) need to be informed regarding the processing of their personal data. This change is in line with the current consensus [CORMACK1, CORMACK2, AARC-G016, AARC-G042] which considers *legitimate interest*, rather than *consent*, the correct legal basis (Article 6.1(f) of [GDPR]) for the processing of personal data in the context of granting access to resources for collaborative and research communities, which is typically done for professional reasons.

The reader will note another proxy among the end services. A proxy is by definition a *service* for the IdPs facing it, and it is sometimes possible to daisy-chain proxies. This approach enables access to resources offered by infrastructures through infrastructure proxies, as described in the following chapter.

4. Architecture for interoperable AARC BPA implementations

We can distinguish between two types of AAI services. One focuses on the infrastructure management, while the other focuses on the community management. Both types of AAI services may comprise the same interfaces (e.g. a proxy), but their functionality and their organisational purposes differ.

The purpose of the Community AAI is to streamline researchers' access to services, both those provided by their own infrastructure (if they have one) as well as the services provided by infrastructures that are shared with other communities. User authentication to the Community AAI uses primarily institutional credentials from national identity federations in eduGAIN, but, if permitted by the community, can also use other IdPs.

The Community AAI follows the proxy-based architecture shown in Figure 1. It can therefore add attributes to the federated identity that in turn can enable services to control access to their resources. Furthermore, the Community AAI is responsible for dealing with the complexity of using different identity providers with the services offered to the community. We can distinguish among three types of services (see also definitions in Chapter 2):

1. *generic services* - provided to members of different communities, or individuals (e.g. the RCauth.eu Online CA service)
2. *community services* - provided only to members of a given community
3. *infrastructure services* - provided by a given research infrastructure or e-Infrastructure, typically through an infrastructure proxy

The architecture, from the perspective of a researcher is shown in Figure 3. This illustrates how community-specific services only need to connect to a single identity provider, i.e. their Community AAI. In contrast, generic services connect to multiple Community AAIs in order to serve different communities. Being connected to multiple Community AAIs requires those generic services to provide some form of IdP discovery, to be able to redirect the user to the relevant Community AAI. Additionally, to allow "community branding" of the service and automatically redirecting the user to the corresponding Community AAI, the generic services may support some means of doing "IdP hinting" (see [AARC-G049]).

Communities may also require access to various services which themselves are behind (another) proxy. This could for example be resources offered by e-Infrastructures or Research Infrastructures (Infrastructures hereafter). These "Infrastructure Proxies" can be connected to multiple Community AAIs (see Figure 3). So, just as for the generic services, Infrastructure services should be able to hint to the Infrastructure Proxy which Community AAI to use (see [AARC-G049]).

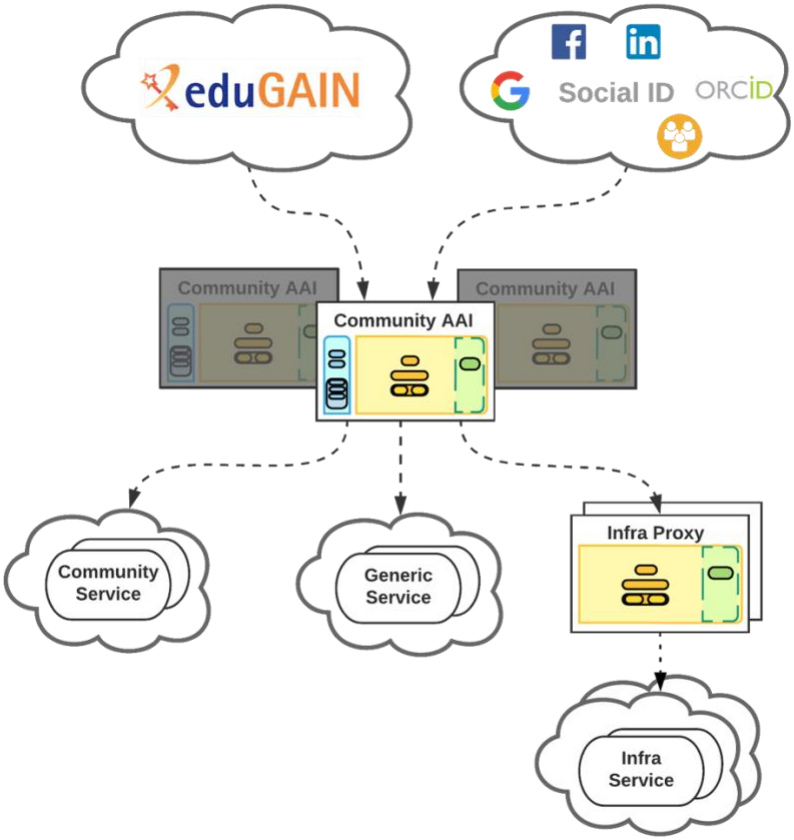

Figure 3: Researchers access services using their institutional (eduGAIN), social or community-managed IdP via their Community AAI. Community services are connected to a single Community AAI, whereas generic services can be connected to more than one Community AAI. Infrastructure services are connected to different Community AAIs through a single Infrastructure Proxy.

It should be noted that this approach does not impose a requirement on communities to deploy and operate a Community AAI on their own. Communities could make use of either dedicated or multi-tenant deployments of AAI services operated by a third-party, typically a generic e-Infrastructures. A multi-tenant AAI service deployment supports different communities, as depicted in Figure 4. It typically appears as a single entity to its connected IdPs and SPs. Such multi-tenant deployments are aimed at medium-to-small research communities/groups or individual researchers. Yet it should be emphasised that also in the multi-tenant AAI scenario, the community managers are responsible for managing their community members, groups and authorisation attributes.

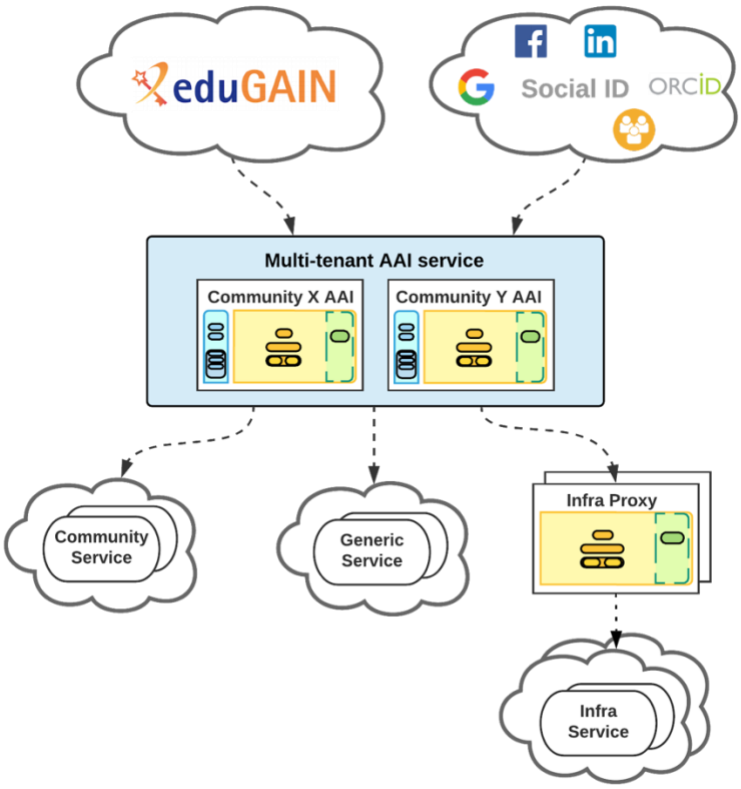

Figure 4: Multi-tenant deployment of AAI services following the AARC BPA architecture.

5. Conclusions

The AARC Blueprint Architecture provides a reference architecture for implementing an AAI that supports common use cases within research collaborations. The evolution of the architecture focuses on interoperability across BPA-compliant AAIs and provides a broader view for addressing an increasing number of use cases from research communities requiring access to federated resources offered by different infrastructure providers.

It should be stressed that AARC-BPA-2019 is backwards compatible with all previous versions of the BPA which have already been adopted by many e-infrastructure providers, research infrastructures and collaborations [AEGIS].

References

AARC-G012 AARC Blueprint Architecture 2017 (AARC-G012); https://aarc-community.org/guidelines/aarc-g012/

AARC-G016 AARC guidelines: Recommendations on the exchange of personal data in accounting data sharing (AARC-G016); https://aarc-community.org/guidelines/aarc-g016/

AARC-G042 AARC guidelines: Data Protection Impact Assessment – an initial guide for communities; https://aarc-community.org/guidelines/aarc-g042/

AARC-G049 AARC guidelines: A specification for IdP hinting (AARC-G049); https://aarc-community.org/guidelines/aarc-g049/

AEGIS AARC Engagement Group for Infrastructures; https://aarc-community.org/about/aegis/

CORMACK1 A. Cormack, "Federated Access Management and GDPR"; https://community.jisc.ac.uk/blogs/regulatory-developments/article/federatedaccess-management-and-gdpr

CORMACK2 A. Cormack, "Legitimate Interests and Federated Access Management"; https://community.jisc.ac.uk/blogs/regulatory-developments/article/legitimate-interests-and-federated-access-management

ESFRI European Strategy Forum on Research Infrastructures; https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en

GDPR General Data Protection Regulation on eur-lex; https://data.europa.eu/eli/reg/2016/679/2016-05-04

REFEDS-R&S REFEDS Research and Scholarship Entity Category; https://refeds.org/category/research-and-scholarship

REFEDS-Sirtfi Security Incident Response Trust Framework for Federated Identity (Sirtfi); https://refeds.org/sirtfi

WISE-SCI Security for Collaborating Infrastructures (SCI) Trust Framework; https://wise-community.org/sci/