29-04-2019

Deliverable D3.3: DNA3.4 Recommendations for e-Researcher-Centric Policies and Assurance

Deliverable D3.4

Contractual Date:	30-04-2019
Actual Date:	29-04-2019
Grant Agreement No.:	653965
Work Package:	WP3 (NA3)
Task Item:	TNA3.3
Lead Partner:	STFC
Document Code:	D3.3

Authors: Ian Neilson (STFC), David Groep (Nikhef), Petr Holub (BBMRI-ERIC), David Kelsey (STFC), Mikael Linden (CSC), Hannah Short (CERN), Uros Stevanovic (KIT)

Abstract

These Recommendations provide a set of frameworks and guidelines that support, involve, and affect researchers and research communities in order to more effectively use federated identity for accessing services in a blueprint-based proxy architecture. At the same time, through support of the (existing) Federated Identity Management for Research (FIM4R) community in expressing and prioritizing the requirements that structured research communities themselves bring forward for their use of federated identity, we ensure that the policies and guidelines proposed meet the research community need – as well as proving the basis for future work in both policy and complementary areas (e.g. in architecture, usability, or governance).

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

Table of Contents

Executive Summary

These Recommendations provide a set of frameworks and guidelines that support, involve, and affect researchers and research communities in order to more effectively use federated identity for accessing services in a blueprint-based proxy architecture. At the same time, through support of the (existing) Federated Identity Management for Research (FIM4R) community in expressing and prioritizing the requirements that structured research communities themselves bring forward for their use of federated identity, we ensure that the policies and guidelines proposed meet the research community need – as well as proving the basis for future work in both policy and complementary areas (e.g. in architecture, usability, or governance).

The recommendations discussed here – and released as AARC Guidelines and Informational Documents – support the AARC2 use cases for high-assurance access to services (and work towards addressing the FIM4R recommendation on "sensitive research user experience") and the urgent issue of addressing the AUP cascading for "multi-BPA" architectures and composite services.

The recommendations here should be read in conjunction with AARC guidelines and information documents that are applicable to or relate to researchers and research communities (https://aarc-project.eu/guidelines/#policy).

1. Introduction

The security policy framework supporting collaboration amongst the infrastructures for research (SCI) and its application to the federated authentication and authorization domain by construction involved both the service providers in the infrastructure as well as the researchers and research communities themselves. And whereas trust both between and within the service provider collaborations can be controlled through the proxy model in the Blueprint Architecture – and can be modelled for example in terms of *Snctfi* – the needs of researcher and research community are broader. Requirements on federated identity management (FIM) for research encompass a far broader category of topics: the identity life cycle and mobility of researchers, the method by which federated identity is discovered and how usable it is, requirements for authorisation, provisioning of access to services (and the de-provisioning at the end), attribute release, security incident response, assurance of identity and the 'strength' of authentication, interoperability between web and non-web services, and the relative ease of 'on-boarding' new collaborations and researchers.

At the same time researchers, even more than service provider collaborations and e-Infrastructures, are global, and benefit from coherence between 'adjacent' communities working in similar fields of research. In recognition of the fact that the best practice in FIM for research is achieved by having the community itself have its say in identification and prioritisation of the requirements on federated authentication and authorisation, the AARC project supported and reinvigorated the "FIM4R" community and worked with that community to extend both its geographical as well as domain scope – taking care not to unduly influence the research communities in their goal of expressing their proper issues and priorities. The result of that process, the second FIM4R white paper [doi:10.5281/zenodo.1307551] and the FIM4R online presence (fim4r.org) are described herein.

Complementing the requirements of the research communities, two key elements were identified that would most readily aid the researchers and research communities in achieving interoperability and seamless access to (increasingly complex, valuable, and globally distributed) infrastructure services. The first is a consistent way to define and express identity assurance – so that a wide range of services can grant access to eligible researchers without the need for each service to independently perform identity vetting or issue high-assurance authentication tokens to the same set of researchers. The second is a common baseline for the 'acceptable use policy' (AUP) that each service provider needs to bind their research users to the intended use of the service, but which in case of composite services and distributed infrastructures would quickly lead to a veritable jungle of partial statements, leading to interrupted workflows where the researcher would continuously have to click through new sets of AUP interstitial notices.

Here, we bring together the set of policy recommendations that involve the researchers and research communities that are most appropriately addressed through policy. There are other requirements identified through the FIM4R process that are more appropriately dealt with through the architecture recommendations (e.g. those of the use of proxies or the consistency of identifiers and account linking) or are more effectively addressed within the realm of the federated identity providers (such as attribute release for research and scholarship applications and discovery guidelines). The recommendations discussed here – and released as AARC Guidelines and Informational Documents – support the AARC2 use cases for high-assurance access to services (and work towards addressing the FIM4R recommendation on "sensitive research user experience") and the urgent issue of addressing the AUP cascading for "multi-BPA" architectures and composite services.

The recommendations here should be read in conjunction with AARC guidelines and information documents that are applicable to or relate to researchers and research communities:

- *AARC-I050* Comparison Guide to Identity Assurance Mappings for Infrastructures
- *AARC-G041* Expression of REFEDS RAF assurance components for identities derived from social media accounts
- The REFEDS Assurance Framework (RAF)
- *AARC-I044* Implementers Guide to the WISE Baseline Acceptable Use Policy

The work is part of the larger set of policies that includes also service-centric policies – discussed in the “Accounting and Traceability in Multi-Domain Service Provider Environments” and its guidelines, as well as guidelines on operational security capabilities – in particular AARC-G048, the “Guidelines for Secure Operation of Attribute Authorities and other issuers of access statements”, that addresses how research community membership data and the resulting attributes is to be protected, and that thereby has direct impact on the member researchers and communities.

2. The REFEDS Assurance Framework

Development of REFEDS Assurance Framework started during AARC project which, in November 2015, published recommendations on minimal assurance level relevant for low-risk research use cases [AARC-MNA3.1]. The research and education federations community [REFEDS] reacted by establishing a REFEDS Assurance working group which took the requirements and based on them developed a REFEDS Assurance Framework [REFEDS-RAF] which was approved and published in October 2018. The AARC2 project funded the RAF specification editor’s work.

2.1 Specification

RAF covers three components of identity assurance

- **uniqueness of user identifier**, including if it represents a single natural person (i.e. not a shared account), if its holder can be contacted, if the identifier values are re-assigned (i.e. recycled to other persons over time) and which attributes or claims are used to express the identifier on SAML2 and OpenID Connect protocols.
- **identity proofing**, including if the identity is self-asserted by the user themselves or verified by a trusted third party. The RAF builds on existing specification, such as Kantara [KANTARA], IGTF [IGTF] and eIDAS [EIDAS].
- **attribute freshness**, in particular how promptly an identity provider and its back-end identity management system is able to close or update a person’s eduPerson(Scoped)Affiliation attribute value when they depart or otherwise change their role in the home organisation.

An identity provider can signal and a relying service observe these components independently. For the relying services that prefer simplicity, two aggregate profiles combining component values are defined:

- **Cappuccino**, to match the low-risk research use cases, originally defined in AARC MNA3.1.
- **Espresso**, to match more demanding research use cases.

The component values and profiles building on them are expressed using the eduPersonAssurance attribute which can be delivered to relying services on SAML2 and OpenID Connect protocols.

RAF is supplemented with two authentication profiles, describing how the user was authenticated by the identity provider in the beginning of the session

- **REFEDS MFA** [REFEDS-MFA], which is an interoperability profile defining conditions for multi-factor authentication. This specification was approved by REFEDS in 2017.
- **REFEDS SFA** [REFEDS-SFA], which defines the conditions satisfying minimum requirements for acceptable single-factor authentication, such as passwords. This specification matches the requirements in AARC MNA3.1 and was approved by REFEDS together with the RAF.

2.2 Implementation pilot and dissemination activity

Already in the first requirements paper by the research communities in 2012 (FIM4Rv1), the “ability to express Levels of Assurance was deemed of high importance and, although much effort was spent to define appropriate protocols, they have yet to be adopted or propagated” [FIM4Rv2]. Demonstrable interoperability and active promotion of the adopted frameworks thus deserves significant effort. The structure of the REFEDS Assurance Suite (RAF, REFEDS MFA and REFEDS SFA) already emphasises feasibility of implementation of the policy aspects. To complement this, a technical pilot was run in collaboration with several Infrastructures and academic identity providers to demonstrate readiness of implementations and address the challenge of identity provider configuration. The RAF Pilot, run in early 2018 with ELIXIR (EU), EGI (EU), CILogon (USA), and SWITCHaai (CH) as service providers, and UChicago (US), XSEDE (US), Aalto University (FI), and CSC (FI), showed the ability of all prevalent software implementations to support the Assurance Suite, and of both R&E identity federations and Infrastructure proxies to consume and act upon the assurance information conveyed. A report was published on the REFEDS web site [RAFpilot] to ensure this information was available to those entities that will need to implement the RAF functionality.

To ensure visibility of the framework and foster adoption, the REFEDS Assurance Suite (RAF, REFEDS MFA and REFEDS SFA specifications) has been disseminated in

- 38th REFEDS meeting in Trondheim (https://refeds.org/meetings/38th-meeting)
- 39th REFEDS meeting in Orlando (https://refeds.org/meetings/39th-meeting)
- TechEx 2018 session in Orlando (https://meetings.internet2.edu/2018-technology-exchange/detail/10005176/)
- Webinar December 2018 (https://www.youtube.com/channel/UCussxbcR_OxG1e_kRp0pjpA)
- FIM4R meeting in February 2019 (https://indico.cern.ch/event/775478/)
- REFEDS blog (https://refeds.org/a/1918, https://refeds.org/a/2108)

2.3 High Assurance Use Cases

Research in many domains builds on data that is sensitive for various reasons other than just personal preferences of researchers. The biomedical and social sciences often work with personal data (pseudonymized data also counting as personal), some of which even fall into the highest sensitivity under data protection legislation (e.g., genetic data, health data, or political opinions and religious beliefs under European General Data Protection Regulation [GDPR]). When dealing with personal data, the access of researchers is also given for a particular purpose only, depending on the legal basis of the collected material. Some research also deals with data of national security concerns (e.g., highly pathogenic biological material). Authorization decisions are then dependent on high-enough assurance of authentication and subject to risk assessment and adopting relevant risk mitigation measures.

Examples of such high assurance use cases are given below, modelled based on experiences from BBMRI-ERIC, European research infrastructure facilitating access to biological samples and associated data.

2.3.1 Example 1: Retrieval of data from medical data repository

Specific examples of this use case: BBMRI-ERIC Colorectal Cancer Cohort.

In this use case, a researcher wants to retrieve certain pseudonymized data set containing clinical and genetic information about a cohort of patients for a particular research project (i.e., for a particular purpose only). The data repository is operated by a research infrastructure – which is a data custodian (e.g., a biobank or an international organization). The research project has been favourably assessed by an ethics review board and is hosted at the home institution of the researcher. From a European data protection perspective, where GDPR is the key regulation, the custodian typically acts as a data controller and the researcher receives the data as a data processor for a particular purpose, and this data transfer is covered by a contract between the data controller and data processor.

Identity vetting assurance	Identity of the requesting researcher needs to be verified at least to NIST SP 800-63-2 or equivalent. For projects that are born by the researcher's home institution, affiliation of the researchers to the institution needs to be asserted, as the data transfer agreement is eventually signed by the two institutions (data custodian institution and researcher's home institution) and these institutions bear legal liability.
Authentication instance assurance	Multi-factor authentication instances are typically required to mitigate risks of a single credential being stolen.
Authorization process	Committee-controlled access, where the committee makes a decision based on the project (= proposed purpose) and legal basis of their data collection (e.g., purpose and limitations of processing specified in the consent, if an informed consent is the legal basis).

Notes: The authentication currently requires manual validation of user identity and manual assignment of multi-factor authentication credentials, as sufficient LoA is not available in the federated academic AAIs.

2.3.2 Example 2: Processing personal data on secure computing infrastructures

Specific examples of this use case: BiobankCloud, MOSLER, TSD.

In this use case the data has been deposited in a secure data processing infrastructure, be it a private cloud or a high-security public cloud complying with requirements on processing personal data. The data can be transferred to such an infrastructure as a part of implementation of the previous use case.

Identity vetting assurance	Assurance is ideally similar to the previous use case; in certain cases it can be lower if the user is solely responsible for handling the data (e.g., uploading the data herself/himself and thus it means s/he was previously already authorized to receive the data). In case the identity vetting and institutional affiliation is used for authorization decision in this use case at each access, timely updates of assurance are important; when a user leaves the institution, s/he should lose the access rights based on losing the institutional affiliation attribute.
Authentication instance assurance	Multi-factor authentication instances are typically required to mitigate risks of a single credential being stolen.
Authorization process	Authorization is based on identity vetting of the user, plus

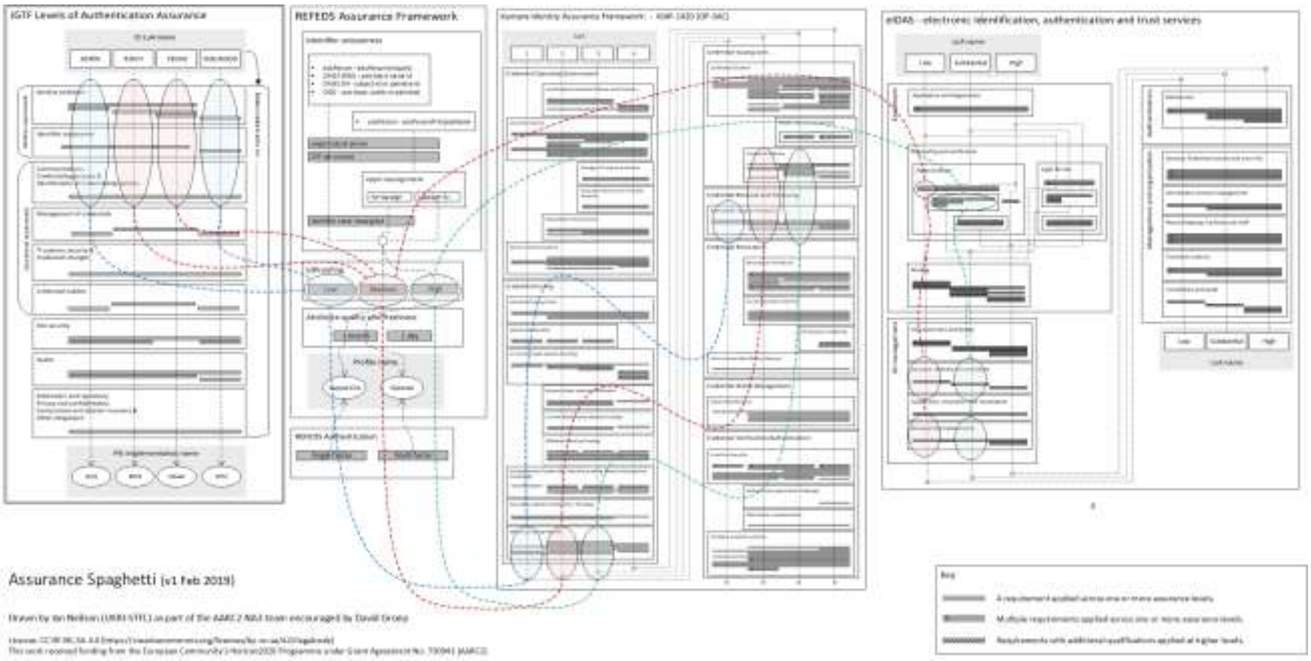
Notes: ISO27001/27018 certification might be an example of standards addressing such public clouds. Demonstrating compliance is, however, a more complex process and the ISO certification is only one of the components that may or may not be sufficient.

2.4 Comparison Guide to Identity Assurance Mappings for Infrastructures

The REFEDS RAF, described above, added to an already wide range of identity assurance frameworks available. The most appropriate choice of profile for a use case (one that meets both the risk assessment and the social and community context in which the assurance is needed) may be viewed as confusing. Such a choice may be between, for example, Cappuccino or Espresso from the REFEDS Assurance Framework, Assam from the AARC social media assurance, Birch and Dogwood from the Interoperable Global Trust Federation [IGTF], Silver and Bronze from InCommon [INCOMMON], or one of Levels 1 through 4 from both Kantara [KANTARA] and NIST SP 800-63 [NIST].

In response to the request for a matrix showing the different assurance levels, in the context of the AARC Guidelines and Deliverables, to clarify the parameters through which a selection of the most appropriate assurance level might be made, the AARC2 NA3 Team created a whitepaper - Comparison Guide to Identity Assurance Mappings for Infrastructures [AARC-I050].

In this whitepaper, the implicit trust assumptions (in research and collaboration frameworks, the R&E identity federations, general private sector frameworks and e-government schemes) were identified and presented as a way of comparing these frameworks together with a suite of diagrammatic representations illustrating their overall structure and relative complexity to assist in understanding and analysis. Finally, a representation of the relationship between the profiles for the identity-proofing component of assurance, relative to the REFEDS RAF, was created and is illustrated below. The reader is referred to the whitepaper for full details.


IGTF Levels of Authentication Assurance
REFEDS Assurance Framework
eIDAS - electronic identification, authentication and trust services
Key
Assurance Spaghetti (v1 Feb 2019)

Mapping between the assurance frameworks from REFEDS, the IGTF, Kantara, and eIDAS. The relations between the various frameworks and the REFEDS identity vetting profiles are represented by the coloured lines and ovals in the diagram and documented in AARC-I050. For an explanation of the diagram and methods, the reader is referred to that informational white paper.

3. Baseline Acceptable Use Policy

Security for Collaborating Infrastructures Trust Framework [SCI] states that *"Each infrastructure has the following: ... An Acceptable Use Policy (AUP) addressing at least the following areas: defined acceptable and non-acceptable use, user registration, protection and use of authentication and authorisation credentials, data protection and privacy, disclaimers, liability and sanctions"* [SCI-V2,section 6]. Acceptable use policy (AUP) and terms and conditions are necessary instruments in the regulation of infrastructure access. They bind the user to the 'purpose' for which the services and resources they use have been provided. Yet, as with privacy notices, the reader is rather inclined to click through and proceed with the actual task at hand. Thus, to reduce the burden on the user and increase the likelihood that they will read the AUP, the number of times a user is presented with such notices must be kept to a minimum, preferably just a single time. Yet the notice should cover as much of the user's potential use of the infrastructure as possible: the more services and resources deem an AUP as sufficient for their policy purposes, the better it will be. This will allow users to use resources from multiple service and resource providers without the need to confirm acceptance of additional AUPs.

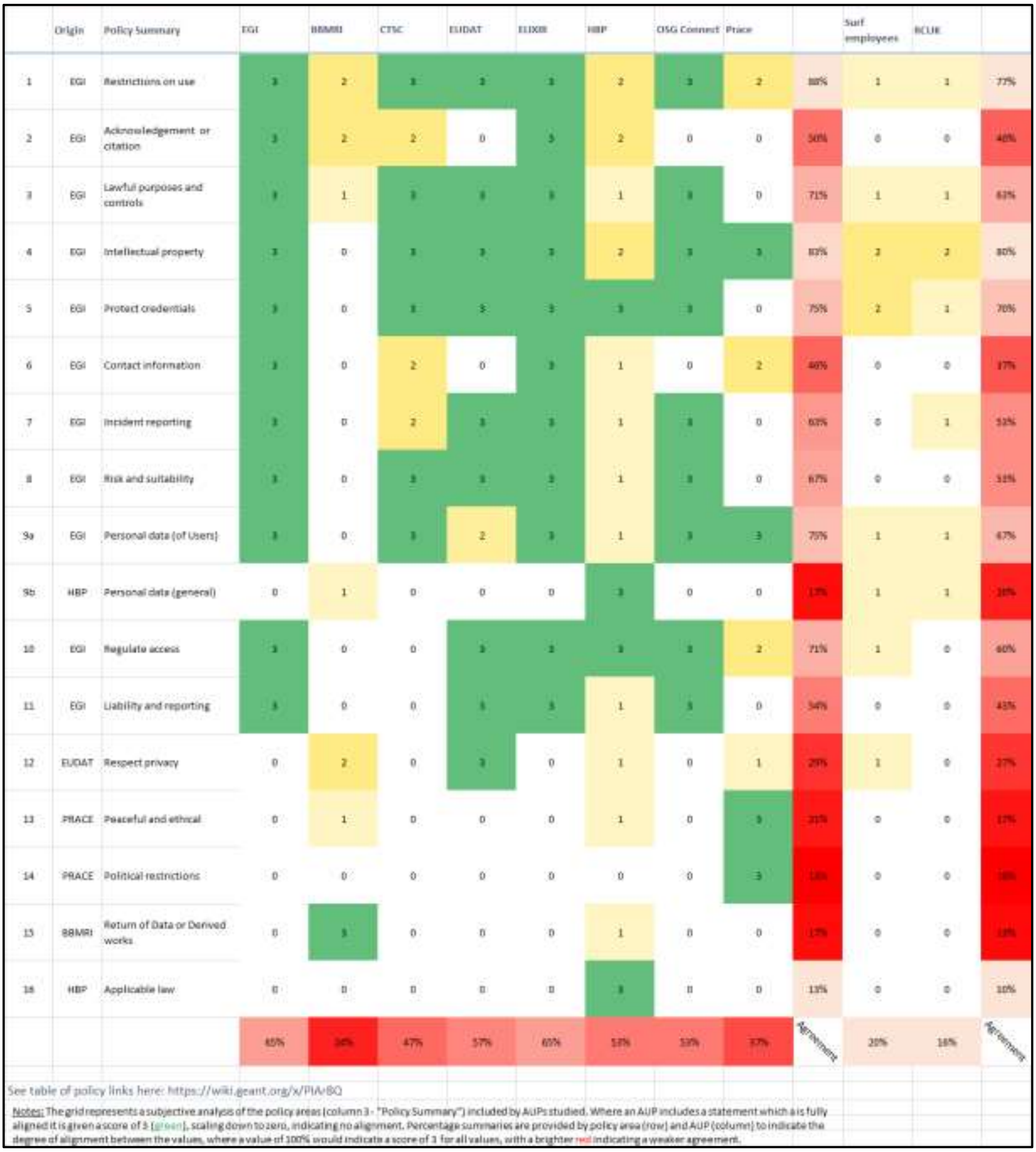
The aim of the Baseline AUP is to

- provide a common baseline set of criteria for acceptable use and terms and conditions for the professional use of IT infrastructures for research globally – and thereby ease the trust of users across infrastructures: services within an infrastructure have a common framework describing the behaviour of users coming from multiple communities;
- facilitate a presentation format that allows necessary privacy notices (in Europe for GDPR compliance) to be presented at the same time and remain easily available thereafter;
- support services with varying levels of support and quality guarantees;
- provide for augmentation of the baseline AUP with community and infrastructure-specific terms and conditions
- be applicable to both community-first and user-first AAI membership management services.

The AARC2 AUP alignment study, followed by a series of consultations and drafts, resulted in the common minimum, or 'baseline', AUP text to satisfy these requirements. Use of this baseline should facilitate rapid community infrastructure 'bootstrap', easing the trust of users across an infrastructure and providing a consistent and more understandable enrolment for users as they move between communities and projects.

3.1 The AUP Study

Given the many possible AUPs in use which might form a basis for a common baseline, a study was conducted comparing 11 existing community and infrastructure AUP texts, looking for commonalities and discrepancies to inform the project. From an initial reading, it was apparent that a majority of the texts already shared a common heritage. This derived from policy created by the Joint Security Policy Group [JSPG] which defined and maintained policies for a number of Grid projects in the early 2000's. The most direct descendant of this work is the EGI AUP [EGI-AUP] and, for this reason, the clauses of the EGI AUP were chosen as comparators, against which the other 10 AUPs could be assessed.



	Origin	Policy Summary	EGI	BBMRI	CTSC	EUDAT	ELIXIR	HBP	OSG Connect	Prace		Surf employees	RCUK	
1	EGI	Restrictions on use	3	2	3	3	3	2	3	2	88%	1	1	77%
2	EGI	Acknowledgement or citation	3	2	2	0	3	2	0	0	50%	0	0	40%
3	EGI	Lawful purposes and controls	3	1	3	3	3	1	3	0	71%	1	1	63%
4	EGI	Intellectual property	3	0	3	3	3	2	3	3	83%	2	2	80%
5	EGI	Protect credentials	3	0	3	3	3	3	3	0	75%	2	1	70%
6	EGI	Contact information	3	0	2	0	3	1	0	2	46%	0	0	37%
7	EGI	Incident reporting	3	0	2	3	3	1	3	0	63%	0	1	53%
8	EGI	Risk and suitability	3	0	3	3	3	1	3	0	67%	0	0	53%
9a	EGI	Personal data (of Users)	3	0	3	2	3	1	3	3	75%	1	1	67%
9b	HBP	Personal data (general)	0	1	0	0	0	3	0	0	17%	1	1	20%
10	EGI	Regulate access	3	0	0	3	3	3	3	2	71%	1	0	60%
11	EGI	Liability and reporting	3	0	0	3	3	1	3	0	54%	0	0	43%
12	EUDAT	Respect privacy	0	2	0	3	0	1	0	1	29%	1	0	27%
13	PRACE	Peaceful and ethical	0	1	0	0	0	1	0	3	21%	0	0	17%
14	PRACE	Political restrictions	0	0	0	0	0	0	0	3	13%	0	0	10%
15	BBMRI	Return of Data or Derived works	0	3	0	0	0	1	0	0	17%	0	0	13%
16	HBP	Applicable law	0	0	0	0	0	3	0	0	13%	0	0	10%
			65%	24%	47%	57%	65%	53%	53%	37%	Agreement	20%	16%	Agreement

See table of policy links here: https://wiki.geant.org/x/PIArBQ

Notes: The grid represents a subjective analysis of the policy areas (column 3 - "Policy Summary") included by AUPs studied. Where an AUP includes a statement which a is fully aligned it is given a score of 3 (green), scaling down to zero, indicating no alignment. Percentage summaries are provided by policy area (row) and AUP (column) to indicate the degree of alignment between the values, where a value of 100% would indicate a score of 3 for all values, with a brighter red indicating a weaker agreement.

Comparison matrix scoring the origin or similarity of AUP policy clauses against the EGI-AUP from selected infrastructure and organisational AUPs

The resulting comparison chart, shown here, presents each AUP policy clause as a horizontal row and each AUP as a column. Where an AUP (column) includes a statement which is fully aligned with the given clause at the row/column intersection it is given a score of 3 (green), scaling down to zero, indicating no alignment. A colour scale of green (3) through yellow (1,2) to white (zero) have been applied. Percentage summaries are provided by policy area (row) and AUP (column) to indicate the degree of alignment between the values, where a value of 100% would indicate a score of 3 for all values, with a brighter red indicating a weaker agreement.

Two policies, from SURF and RCUK, representing organisational AUPs rather than project or infrastructure policies, were included in the study. Due to the different nature of the target audiences these results were tabulated separately. Also, additional rows, being clauses not present in the EGI policy, but which were seen a representing further useful input for the study were added to the comparison.

Although subjective in its analysis, the relatively high correlation (3-green) across the rows of the comparison study did provide data to indicate that selection of clauses from the current EGI Acceptable Use Policy would form a reasonable basis of a ‘baseline’ AUP of more general applicability.

3.2 Development of the AUP template

Adopting the result of the AUP Study described above - that the so-called ‘JSPG-evolved’ EGI AUP clauses could be used as a baseline template - an initial AARC Baseline AUP was created for further work. This is included below for reference.

“JSPG-evolved” AUP

By registering as a user you declare that you have read, understood and will abide by the following conditions of use:

1. You shall only use the resources/services to perform work, or transmit or store data consistent with the stated goals, policies and conditions of use as defined by the body or bodies granting you access.
2. You shall provide appropriate acknowledgement of support or citation for your use of the resources/services provided as required by the body or bodies granting you access.
3. You shall not use the resources/services for any purpose that is unlawful and not (attempt to) breach or circumvent any administrative or security controls.
4. You shall respect intellectual property and confidentiality agreements.
5. You shall protect your access credentials (e.g. private keys or passwords).
6. You shall keep all your registered information correct and up to date.
7. You shall immediately report any known or suspected security breach or misuse of the resources/services or access credentials to the specified incident reporting locations and to the relevant credential issuing authorities.
8. You use the resources/services at your own risk. There is no guarantee that the resources/services will be available at any time or that their integrity or confidentiality will be preserved or that they will suit any purpose.
9. You agree that logged information, including personal data provided by you for registration purposes, may be used for administrative, operational, accounting, monitoring and security purposes. You agree that this logged information may be disclosed to other authorised participants via secured mechanisms, only for the same purposes and only as far as necessary to provide the services.
10. You agree that the body or bodies granting you access and resource/service providers are entitled to regulate, suspend or terminate your access without prior notice and without compensation, within their domain of authority, and you shall immediately comply with their instructions.
11. You are liable for the consequences of your violation of any of these conditions of use, which may include but are not limited to the reporting of your violation to your home institute and, if the activities are thought to be illegal, to appropriate law enforcement agencies

During the course of 2018 and early 2019, centred around meetings of AARC2, EUGridPMA [EUGRIDPMA], WISE [WISE] and at joint infrastructure workshops involving representatives from EGI, EOSC-hub, EUDAT, WLCG and others, the AARC2 NA3 Policy Team refined, and disseminated for discussion, updated versions derived from this text. Consideration was also given to how the AUP would be maintained following the end of the AARC project. To this end, the WISE Community [WISE] was approached and agreed to consider the policy for formal inclusion in its published set of documents.

The following events were used for both dissemination and work on integration of received comments -

- Presentation of AUP alignment work in progress at the WISE security group meeting in Abingdon, UK 27/2/2018
- Presentation of AUP alignment work at AARC2 third project meeting in Athens, Greece 10-13/4/2018 and 43rd EUGridPMA meeting in Karlsruhe, Germany 23-26/5/2018 with additional material
- Presentation of AUP alignment work at EOSC-hub/AARC2/EGI/EUDAT/WLCG Joint Security Policy Workshop at CERN, Switzerland 18-20/7/2018
- Presentation of AUP alignment work given by Dave Kelsey at the WISE session at the 2018 NSF Cybersecurity Summit in Alexandria, VA, USA 21/8/2018
- Presentation of AUP alignment work at the 44th EUGridPMA meeting in Toulouse, France 26/09/2018
- Resulting AUP document (v1.2) after discussions at the EOSC-hub/AARC2/EGI/EUDAT/WLCG Joint Security Policy Workshop at Forschungszentrum Jülich, Germany, 14-16/11/2018
- Discussions at the AARC2 Fourth Meeting at Reti, Busto Arsizio, Italy, 19-22/11/2018 resulted in the drafting of AARC-I044 Implementers Guide to the WISE Baseline Acceptable Use Policy (preliminary) formatted document here.
- Discussion of adoption and sustainability of AUP and other AARC outputs at 45th EUGridPMA meeting at CERN, Switzerland 21-23/01/2019
- Resulting draft WISE AUP document (v1.3) after discussions at the EOSC-hub/AARC2/EGI/EUDAT/WLCG Joint Security Policy Workshop in Abingdon, UK 19-21/02/2019

In March 2019 the WISE community formally adopted the AUP and published the text on its website. This final text is included below for reference.

WISE Baseline AUP template v1.3

When using the baseline AUP text below, curly brackets "{ }" (coloured blue) indicate text which should be replaced as appropriate to the community, agency or infrastructure presenting the AUP to the user. Angle brackets "< >" (coloured green) indicate text which is optional and should be deleted or replaced as indicated. Other text should not be changed.

Acceptable Use Policy and Conditions of Use

This Acceptable Use Policy and Conditions of Use ("AUP") defines the rules and conditions that govern your access to and use (including transmission, processing, and storage of data) of the resources and services ("Services") as granted by {community, agency, or infrastructure name} for the purpose of {describe the stated goals and policies governing the intended use}.

<To further define and limit what constitutes acceptable use, the community, agency, or infrastructure may optionally add additional information, rules or conditions, or references thereto, here or at the placeholder below. These additions must not conflict with the clauses 1-10 below, whose wording and numbering must not be changed.>

1. You shall only use the Services in a manner consistent with the purposes and limitations described above; you shall show consideration towards other users including by not causing harm to the Services; you have an obligation to collaborate in the resolution of issues arising from your use of the Services.
2. You shall only use the Services for lawful purposes and not breach, attempt to breach, nor circumvent administrative or security controls.
3. You shall respect intellectual property and confidentiality agreements.
4. You shall protect your access credentials (e.g. passwords, private keys or multi-factor tokens); no intentional sharing is permitted.
5. You shall keep your registered information correct and up to date.
6. You shall promptly report known or suspected security breaches, credential compromise, or misuse to the security contact stated below; and report any compromised credentials to the relevant issuing authorities.
7. Reliance on the Services shall only be to the extent specified by any applicable service level agreements listed below. Use without such agreements is at your own risk.
8. Your personal data will be processed in accordance with the privacy statements referenced below.
9. Your use of the Services may be restricted or suspended, for administrative, operational, or security reasons, without prior notice and without compensation.
10. If you violate these rules, you may be liable for the consequences, which may include your account being suspended and a report being made to your home organisation or to law enforcement.

<Insert additional numbered clauses here.>

The administrative contact for this AUP is: {email address for the community, agency, or infrastructure name}
The security contact for this AUP is: {email address for the community, agency, or infrastructure security contact}
The privacy statements (e.g. Privacy Notices) are located at: {URL}
Applicable service level agreements are located at: <URLs>

3.3 Guidance for use of the AUP template

Supplementing the basic instructions on usage of the AUP template included within the text itself, the NA3 policy team produced the Implementer’s Guide to the WISE Baseline Acceptable Use Policy [AARC-I044]. This guide provides detailed advice for implementers wishing to adopt the Baseline AUP, distinguishing two possible implementation models. The first, more straightforward, route to adoption, is a so-called “community-first” use-case. Here, a community (or other body) has its own ‘AAI entry point’ and can easily adapt the AUP template for its own purposes for presentation to its members. A second model, so-called “user-first” use-case, is used where the community, in which the user has yet to become registered, makes use of a (multi-tenant) membership management service. The reader is referred to the Implementers Guide for full details.

4. Federated Identity Management for Research Collaboration

Effective engagement with the community of researchers with respect to AAI policies and practices has traditionally been fraught with challenges: identity management is not in itself of interest to researchers but merely needed to gain access to services, data, or collaborations – and as such easily considered a hindrance rather than an enabler for research. And conversely, the needs of researchers and research communities as to what they need from an identity management solution are often expressed piecemeal or come with implicit assumptions as to how a solution applies to the community that are not explicitly captures as requirements. Addressing this gap is the “FIM4R” (Federated Identity Management for Research) community, the group that in 2012 brought together “research communities and infrastructures with a shared interest in enabling Federated Identity Management for their research cyber infrastructures”. After publishing the initial white paper in 2012 [FIM4Rv1], the group retained a coordination function and met (mostly on a biannual basis) to exchange ideas, challenges and best practices in FIM. It is intentionally composed from a broad range of research stakeholders, including photon/neutron facilities, social science & humanities, high-energy physics, atmospheric science, bioinformatics and fusion energy. To a certain extent, the AARC project series itself may be seen as one of the results of the white paper published in 2012. The paper similarly influenced the direction of the REFEDS group and the AAI developments in infrastructures such as GEANT.

The FIM4R community thus provides the most logical place to validate the results of AARC activities, to disseminate to and support the researchers and research communities in effectively using the (policy and best practice) results from AARC, and to ensure their future requirements are collected and effectively anchored in both AARC as well as in the continuation of its work in future forums. The AARC project thus identified FIM4R as the primary mechanism for Community Engagement, and decided to support the reinvigoration of the group by supporting the showcasing of cross-domain demonstrators (the AARC pilots), by promoting its meetings in both Europe and globally, and by encouraging all communities to review and assess the original requirements and chart the way for the future FIM developments for research and collaboration.

At the same time, identified immediate needs of the FIM4R community – including for example the need for ‘less hassle’ (e.g. by aligning acceptable use policies) and the call for help in drafting policy suites for new communities which resulted in the Policy Development Kit – could be expressed and taken up by the AARC policy and best practice activity.

Successful developments in FIM since the FIM4R version 1.0 publication included:

- Development and implementation of the Research & Scholarship Category [REFEDS-R&S], a globally adopted program that defines a set of user attributes and helps to manage user privacy by disclosing them only to federated services independently vetted to be purposed for research or scholarly use.
- Development and implementation of Sirtfi [REFEDS-Sirtfi], a security incident response framework adopted by R&E Federations, Snctfi [Snctfi], a suite of policies that facilitate successful integration of cyber infrastructures with R&E Federations, and the GÉANT Data Protection Code of Conduct [GEANT-DPCoCo] to address data privacy compliance needs of federated organisations in the EU.
- Various solutions to non-browser federated access needs.
- Experiments with defining and fielding solutions to Level of Assurance needs [INCOMMON] [REFEDS-RAF].
- Emergence of the AARC BPA proxy architecture as the approach taken by multiple research cyber infrastructures to simplify their integration with R&E Federations.

Yet likely the most significant outcome of AARC support to the FIM4R community was the new “version 2” of the FIM4R White Paper, issued in July 2018, which both shows the evolution of FIM4R which was co-enabled by AARC, and identifies specific recommendations to all stakeholders for its future evolution.

4.1 The FIM4R V2 Paper

Although the developments since FIM4R version 1, described above, are quite substantial, they have not fully addressed all of the problems at which they are aimed. Most R&E organisations do not yet participate in the Research & Scholarship Category or Sirtfi programs, and the Data Protection Code of Conduct has had to be revised in light of the General Data Protection Regulation [GDPR], a process that is not yet completed. The InCommon Federation in the United States developed its Bronze and Silver Levels of Assurance but discovered that most of its member organisations found them too onerous to implement.

Whereas in 2012 many practitioners envisioned that every service in a research cyber-infrastructure would directly join an R&E Federation, experience has shown that it is more practical and scalable to implement a proxy for them that is joined to an R&E Federation. This model is articulated in the AARC Blueprint Architecture and centralises credential translation, authorisation management, and other functions in one place, avoiding the need to do so in each service within a cyber infrastructure and join it to the R&E Federation. A proxy helps to mitigate the shortcomings of the Research & Scholarship Category program by providing an alternate locus for managing needed user attributes. There are now good open source proxy platforms that address these needs, and services such as ORCID [ORCID], for example, that provide new approaches to meeting some of the needs of research communities.

In early 2017, FIM4R members, encouraged and supported by AARC2 NA3 staff, determined that it was time to look anew at how integration of FIM and research cyber infrastructures should continue to evolve and began a new cycle of gathering input from research communities and cyber infrastructures.

Representatives of 14 research fields across physics, astronomy, climate and planetary science, life sciences, infectious diseases, and humanities, to name but a few, and their supporting research cyber infrastructures, provided input. Five face-to-face meetings focused on this effort took place in Europe and North America. At three of them presentations by research communities and cyber infrastructures were heard, followed by discussion to appropriately integrate their specific requirements within a single catalogue. At another meeting sets of specific requirements were assigned to break-out groups to reconsider whether they were the right requirements, which led to some requirements being removed, others merged, and sharpening of the language used to express those remaining. At the final face-to-face meeting, research communities were asked to endorse requirements, ensuring that the published list reflected genuine needs. More details of the approach and methods used are available in our paper at the CHEP2018 conference [FIM4R-CHEP2018].

An editorial team was established to complete the final paper, which was published in June 2018 on Zenodo[FIM4Rv2], in line with the group's affiliation with Open Research. Members of the editorial team prepared a set of presentations at meetings in Europe (TNC18, RDA, CHEP), North America (Internet2 Technology Exchange), and Asia Pacific (ISGC) to inform the wider community and seek further input.

Summary of FIM4R V2 Recommendations

The FIM4R white paper version 2.0 has identified sets of recommendations that the authors strongly believe is beneficial to the academic community. As in version 1, the paper starts with a comprehensive set of requirements that were identified together with all the stakeholders. These requirements address, among others, attribute release, web and non-web access and technologies, security, authentication, and authorisation. An expansive analysis of requirements, both by their distribution (i.e. who has expressed particular requirements) and their importance (i.e. how important are the requirements), has produced a set of recommendations aimed to address the identified problems and to increase the uptake of FIM. The paper further seeks to facilitate the adoption of best practices by mapping the identified concerns to relevant groups and stakeholders best suited to address them.

- **Governance and coordination** Representation of researchers and research e-Infrastructure operators within large Infrastructures should be increased. This would help to ensure the continued alignment of interests of the researchers with the intended mission of large Infrastructures, i.e. supporting research and scholarly aspects. It is also essential that the FIM services are operated sustainably, reliably, and with the appropriate user support. Due to the ever changing technological and operational environment, providing a forum, owned and attended by research communities, where common issues can be discussed is beneficial.
- **Baseline of Research User Experience** This section identifies several well-established practices, of which increased adoption would significantly boost the usability of federated access mechanisms. Releasing a sufficient set of attributes (as identified in R&S entity category) would increase the value of federations, and reduce the impediments the researchers face in accessing remote services. Better harmonisation of import/export practices by R&E operators of their metadata is recommended, and also providing a process through which certain research organisations (that are not legal entities) can be admitted, either at the national or international level. Usability can further be increased by better presenting errors to the users and by providing means to support user mobility (e.g. ORCID).
- **Security Incident Response Readiness** All participants in FIM and federations should support best practices for operational security (such as Sirtfi). Having a security incident response plan is strongly recommended, as well as periodic testing of such abilities. The issue of cooperation for security purposes is also recognised and encouraged by legislation, e.g. Recital 49 of the GDPR.
- **Harmonisation of Research Community Proxy Operations and Practices** As previously mentioned, proxies have emerged as an answer to needs unmet by R&E federations. They have by now matured as a solution, hence their stability, support, and sustainability is becoming significantly important. This includes following the identified best practices, such as the AARC Blueprint Architecture and related guidelines, and reuse of certain AAI services, when applicable.
- **Sensitive Research User Experience** Employing strong controls of authentication and access management is paramount. This is necessary, for example, to ensure confidentiality of restricted research data, integrity of basic research data, or fine grained access to expensive instruments and computing resources. REFEDS Assurance Framework in its first version has issued guidance responding to these needs, and it is encouraged that these instructions are applied by relevant parties.

5. Summary and outlook

Policy and the associated set of practices is there to support researchers and the research communities, but in order to be effective needs the engagement of the researcher end-users. The AARC activity on researcher-centric policy specifically addressed those elements of policy that are most visible to the users and where a specific push towards adoption of coordinated standards would significantly ease the implementation of federated identity processes: a common standard for identity authentication assurance (obviating the need to perform time and human-intensive identity vetting repeatedly for the same users), and a harmonised baseline Acceptable Use Policy (AUP) that reduces the amount of 'interstitial' screens that have to be presented to the user and that would otherwise interrupt the research workflow in multi-infrastructure scenarios.

Guidelines, specifications, and implementation guides were published by or with significant support from AARC to support these standards: the REFEDS Assurance Framework (which formed part of the larger suite including standards for single- and multi-factor authenticator assurance); a comparison guide to assurance mappings to put the REFEDS framework (focussing on feasible assurance for IdPs) in the larger context of e-Infrastructure assurance requirements, the Kantara Assurance Framework, and the eIDAS levels of assurance; guidance on expressing social media assurance for researchers; and implementation guidance on the WISE Baseline AUP for community-first and user-first membership management services.

In order to ensure alignment with the research community requirements and promote global harmonisation, the policy development and engagement work liaises with the research communities through the FIM4R ("Federated Identity Management for Research") collaboration, a collection of research communities and infrastructures with a shared interest in enabling Federated Identity Management for their research cyber infrastructures spanning almost all of the key Research Infrastructure clusters in Europe and with strong global participation. The second FIM4R White Paper, published in July 2018, both validated the success of the AARC BPA proxy architecture and accompanying policies and practices, and sets out the important requirements for the future.

By construction, most of the activities in the research (and also service) centric policy development were executed in the context of existing or broadly established groups that will 'outlive' the AARC project proper. The work on assurance where it targets identity federations and institutional providers was done in the context of REFEDS, where it will continue to be developed and promoted on a global scale. The interaction with the research infrastructures is supported through the Interoperable Global Trust Federation IGTF which has adopted the service-centric assurance and its evolution. The WISE Information Security for e-Infrastructures community SCI working group provides the home for the Baseline Acceptable Use Policy, and with support from ongoing projects including EOSC-HUB and GEANT4-3, as well as international partners in the Americas and the Asia-Pacific region, will promote its global adoption.

The FIM4R community has already demonstrated its long-term viability, and version 2 of the FIM4R white paper makes specific recommendations as to how research can be best supported by federated identity management. The AARC activities are happy to have been able to contribute to the FIM4R goals and views the continuity of its activities in adjacent and succeeding projects as a demonstration of its usefulness in promoting harmonisation across all sectors and geographic regions.

References

[AARC-I044]	https://aarc-project.eu/guidelines/aarc-i044/
[AARC-I050]	https://aarc-project.eu/guidelines/aarc-i050/
[AARC-MNA3.1]	https://aarc-project.eu/wp-content/uploads/2015/11/MNA31-Minimum-LoA-level.pdf
[EGI]	https://egi.eu
[EGI-AUP]	https://documents.egi.eu/document/2623
[EIDAS]	https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2014.257.01.0073.01.ENG
[EUGRIDPMA]	https://www.eugridpma.org/
[FIM4R]	https://fim4r.org/
[FIM4R-CHEP2018]	“Federated Identity Management for Research”, Thomas Barton et al. accepted for publication in the Proceedings of the 23rd International Conference on Computing in High Energy and Nuclear Physics, 9-13 July 2018, Sofia, Bulgaria
[FIM4Rv1]	https://cds.cern.ch/record/1442597
[FIM4Rv2]	https://doi.org/10.5281/zenodo.1307551
[GDPR]	https://eur-lex.europa.eu/eli/reg/2016/679/oj
[GEANT-DPCoCo]	https://wiki.refeds.org/display/CODE/Data+Protection+Code+of+Conduct+Home
[IGTF]	https://www.igtf.net/ap/authn-assurance/
[INCOMMON]	https://www.incommon.org/assurance/
[JSPG]	http://proj-lcg-security.web.cern.ch/proj-lcg-security/
[KANTARA]	https://kantarainitiative.org/confluence/display/LC/Identity+Assurance+Framework
[ORCID]	https://orcid.org/
[REFEDS]	https://refeds.org
[REFEDS-MFA]	https://refeds.org/profile/mfa
[REFEDS-SFA]	https://refeds.org/profile/sfa
[REFEDS-RAF]	https://refeds.org/assurance
[REFEDS-R&S]	https://refeds.org/research-and-scholarship
[REFEDS-Sirtfi]	https://doi.org/10.5281/zenodo.1256531
[RAFpilot]	https://wiki.refeds.org/display/GROUPS/RAF+pilot+final+report
[SCI]	https://wise-community.org/sci/
[SCI-V2]	https://wise-community.org/wp-content/uploads/2017/05/WISE-SCI-V2.0.pdf
[Snctfi]	https://www.igtf.net/snctfi/
[NIST]	https://pages.nist.gov/800-63-3/
[WISE]	https://wise-community.org/