

Preliminary Proxy Policy Recommendations (application to the LS AAI)

Publication Date: 2018-02-12 (DRAFT v2)
Authors: David Groep
Internal Reference: AARC-initial-LSAAI-policy-recommendations.docx
Lead Partner: Nikhef
Document Code: AARC2-NA3

© GÉANT on behalf of the AARC project.
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 Pilot covering the Life Sciences AAI Infrastructure Proxy, developed in joint collaboration with EGI, EUDAT and GÉANT, is a multi-staged pilot that will result in a production-equivalent service to be operated for the Life Sciences community by the joint e-Infrastructures. As the pilot enters its second phase, a practical policy related issue is that the LS AAI has to declare R&S and CoCo. In this document, NA3 aims to provide preliminary guidance for the operators of the pilot.

It must be understood that **this guidance may and likely will change**, in particular if and when the GEANT Data Protection Code of Conduct has been formally approved by the European Data Protection Board, and when relevant components of the Policy Development Kit and the Aligned Acceptable Use Policy for Infrastructures will be adopted.

1. Introduction

The Life Sciences AAI Infrastructure Proxy, developed in joint collaboration with EGI, EUDAT and GÉANT, will result in a production-equivalent service to be operated for the Life Sciences (LS) community by the joint e-Infrastructures. In order to gain acceptance of the Proxy by the R&E federations, and to ensure that both LS-specific and generic e-Infrastructures can provide services to the LS community via the proxy, the LS AAI has to ensure and declare adherence to REFEDS Research and Scholarship (R&S) [RS] and the GEANT Data Protection Code of Conduct (DPCoCo) [DPCOCO].

Moreover, we need to keep in mind that the policy adopted by the LS AAI Proxy should also satisfy the (security) requirements of its underlying service providers, in particular the generic e-Infrastructures that already have both security and acceptable use policies in place.

In this document, NA3 aims to provide preliminary guidance for the operators of the LS AAI Proxy. To provide this guidance, we leverage the following mechanisms

- Scalable Negotiator for a Community Trust Framework in Federated Infrastructures (Snctfi) [SNCTFI]
- Acceptable Use Policy alignment study (draft) [AUPSTUDY]
- the JSPG (evolved) version of the Acceptable Use Policy [JSPGAUP2]
- the draft GÉANT Data Protection Code of Conduct v2 (29Jan2018) [DPCOCO2]

It must be understood that *this guidance may and likely will change*, in particular if and when the GEANT Data Protection Code of Conduct has been formally approved by the European Data Protection Board, and when relevant components of the Policy Development Kit and the Aligned Acceptable Use Policy for Infrastructures will be adopted.

2. Research and Scholarship

In order for the LS Proxy to conform to the R&S entity category requirements, it must itself be and connect subordinate service provided that “*are operated for the purpose of supporting research and scholarship interaction, collaboration or management, at least in part*”. The federation that registers the Proxy additionally has to execute a series of checks to ensure compliance to the entity category, but since such checks are done solely against the Proxy itself (and not its subordinate services), these compliance checks need not be propagated downward ‘as-is’, but can be qualified.

The LSAAI Proxy complies with the REFEDS R&S criteria and best practice, and can request assertion of this entity category under the following conditions:

- the Proxy operators themselves ensure refreshing of the meta-data of the proxy at least daily

- the Proxy operators submit for consideration only a SAML2 end-point that supports SAML V2.0 HTTP-POST binding
- the Proxy operators provide an `mdui:DisplayName` and `mdui:InformationURL` in metadata (and should provide a logo if at all possible)
- the Proxy operators provide at least technical contact in the meta-data
- the Proxy operators SHOULD provide a security contact in compliance with the Sirtfi criteria [SIRTFI]

Since the LSAAI and any services connected to it are by definition enhancing scholarship, eligibility is satisfied for the SP end of the proxy.

The LSAAI should commit to not connecting community services that are not offering services to the community. The generic e-Infrastructures already satisfy R&S eligibility by default.

For the *DisplayName*, *informationURL*, and the *logo*, we recommend names and URLs that reflect the LS community (not the operators), as these will be seen and used as guidance by end-users.

2.1. Subordinate service providers

The R&S eligibility criteria are SAML-specific, but this does not necessarily translate to requirements on subordinate SPs, who can be either OIDC or SAML, but even for SAML not all requirements are ‘pushed down’ by the R&S spec. To ensure that updating the meta-data daily makes sense, we recommend that all downstream SPs and generic Infrastructures comply – at least in spirit even if not technically feasible – with Sirtfi [SIRTFI] as well. For the e-Infrastructures and the LSAAI, we recommend that security incident response is harmonised in the spirit of joint trust groups, a coordinated CSIRT information exchange, and the Framework for a coordinated response to security incidents [FEDINCR3.2].

More specific technical AARC guidelines apply to attribute release to downstream services, but there at least other key identifiers are passed or offered for translation (in particular the life science ID of the user). Here the R&S compliance does not impose any restrictions on the relation between the Proxy and the subordinate SPs, so the LSAAI operators are free to choose the appropriate mechanisms (unique identifiers, assurance profiles, &c) as per the pertinent AARC recommendations.

2.2. Identity Provider support

For the IdP entity category support, R&S states that specific identifiers of the users are released to the service (in this case, the proxy). These should be used to uniquely identify the user within the LSAAI system – in particular to identify the entity and associate the proper Life Science ID.

3. Data Protection Code of Conduct

The GEANT Data Protection Code of Conduct is currently being reviewed, so there is some level of ambiguity as to which one is the most appropriate. We assume that adherence to DPCoCo version 1 is needed in the short term (for assertion in the meta-data of the Proxy), but take interpretation predominantly from the draft version 2 since its scope is broader and

not incompatible with version 1 (it is mainly more specific). For both the Proxy and for virtually all of the SPs, compliance with DPCoCo version 1 is implicit anyway, since it reflects the Data Protection Directive and means compliance with applicable European rules. For those (few) SPs that are non-European, they are typically bound in a policy framework like the one for EGI, which has been evaluated for DPCoCo v1 compliance and found to be materially compliant with it.

There is however an item that needs to be addressed to satisfy even DPCoCo version 1 compliance, and that is that each privacy notice should include a reference to compliance with the Code of Conduct. And although it is not explicitly mandatory to push this requirement 'down' to the subordinate SPs and e-Infrastructures, it would be good practice and it would simplify any assessment that the Proxy operators need to do. Following the methodology of DPCoCo version 2:

“The Service Provider shall not to transfer Attributes to any third party (such as a collaboration partner) except:

- a) if mandated by the Service Provider for enabling access to its Service on its behalf, or*
- b) if the third party is committed to the Code of Conduct or has undertaken similar duties considered sufficient under the data protection law applicable to the Service Provider or*
- c) if prior Consent has been given by the End User.”*

Since the LS Proxy itself does not constitute the Service, (a) is not a basis that would be pertinent. And although the end user should be informed – at sign-up time and periodically – that the LS Proxy acts on behalf of a group of service providers and infrastructures (c) is not applicable as the user is not in a position to give consent as defined by WP29. This leaves as a basis for transfer that *the third party is committed to the Code of Conduct or has undertaken similar duties*. Stating compliance with the Code of Conduct is the easier of the two.

3.1. Snctfi

However, we propose that the commitment is strengthened, and to some extent can be demonstrated without further ado, provided that the underlying service provider has a set of policies in place that meet the requirements of Snctfi [SNCTFI].

Also Snctfi emphasizes data protection (through criteria DP1 and DP2), and recommends adherence to the Code of Conduct or mechanisms that come close to it in spirit: adoption of a binding and self-consistent security policy framework with mechanisms to evict non-compliant participants [BCRLIKE].

Snctfi also addresses elements identified as important in DPCoCo v2, in particular information security (compliance with *Sirtfi* [SIRTFI]) and the information duty towards the user. Although this is *not* 'consent' (the legal basis for the CoCo is *legitimate interest* under 6.1f of the GDPR, as is also suggested to the Home Organisations), the privacy notice must be shown to the end-user on first access to the service. We recommend to show this notice

(and get from the user a positive indication that the message has been received) at the same time as an acceptable use policy is shown (RU1, RU2, and RU3 in *Snctfi*).

3.2. Requirements on the Proxy itself

The Proxy itself must meet the requirements of *Sirtfi*, and the operators must be subject to infrastructures that adhere to documented security practices as specified in OS1 and OS2.

The Proxy must show a privacy notice and designate the relevant contact points for the Proxy itself for the data it processes.

In many cases, the Proxy conveys personal data to third parties (SPs and the e-Infrastructures), and in doing so make sure that these SPs and e-Infrastructures are committed to the Code of Conduct or have undertaken similar duties. The simplest way to demonstrate this is to collect, e.g., on a public web page that is easily findable, a list of references to the privacy notices of all connected services. These may be automatically collected (e.g. from “policy_uri” metadata_statements in OIDC federation) or maintained manually.

The Proxy operators shall post on a public URL, and present to the end-user during signup, a Privacy Notice compliant with the *privacy notice template* as shown in DPCoCo v2 draft 2 annex 1 [DPCOCO2] under “*privacy notice template*”.

The Proxy operators should present the privacy notice in conjunction with the Acceptable Use Policy (described below).

The Proxy operators should establish a web page, preferably linked to the above-mentioned page, that lists the privacy notices of all connected SPs and Infrastructures.

The Proxy operators should also review the Code of Conduct and the GDPR for their own compliance. In almost all cases adherence to sound security practices and *Snctfi* will be adequate, since the proxy itself does not deal with sensitive information (for the AAI, the data itself may of course be very sensitive). It should however be pro-active in addressing security concerns and in identifying potential data breaches, since the mere fact that researchers participate in life sciences research could in specific circumstances be confidential (not only for IPR reasons but also because societal pressure groups at times take views vis-à-vis researchers that could expose the researchers to danger).

At the minimum, the Proxy operators should know where personal data is stored and processed within their own systems. This is an explicit requirement of the GDPR. It is also recommended that (at least for internal reasons) a record of the *balancing test* performed to justify legitimate interest is documented. This need not be long, but it helps assess the risk even if no full Data Protection Impact Assessment (DPIA) needs to be performed.

The Proxy operators shall maintain a registry of where personal data is stored and processed within the Proxy service components themselves. This should include a list of systems, processing organisations, and a security contact.

It is recommended that the Proxy operators establish and periodically test a communications flow for use in security incidents and for reporting potential data breaches to the (life sciences community) Data Controller.

The end-user will have to be contacted periodically, e.g. for reconfirmation of the AUP or when the a change in the set of subordinate service providers results in the need to ask for *consent* as per DPCoCo version 2. This is addressed by Snctfi RC1, which must be implemented by the Proxy registration process (in Perun). This information is also needed to effectuate collective incident response (RC4, RC5).

The Proxy operators shall record for all end-users enough information to contact the user in case of security incidents, and to inform these users in case the acceptable use or data transfer to third parties change in a way that must be communicated to the users.

Both the GDPR as well as the Code of Conduct require the designation of a Data Controller. The LS AAI Proxy is operated by the e-Infrastructures, but the purpose as well as means of processing (technical and – within the context of a funding proposal coordinated by the LS community – materially) are entirely determined by the LS community – who has set both the requirements and has performed the selection of the consortium. This means that in the sense of the GDPR and the Code of Conduct the Life Sciences community is the *Data Controller*. The LS AAI Proxy operator consortium (EGI, the EUDAT partners, GEANT) processes the personal data in the proxy on behalf of the controller only. This includes both the hosting of the personal data (in Perun) and in transit in the other Proxy elements. However, until now it has been unclear which natural or legal person has actually requested the proxy services.

The Life Sciences community must designate a natural or legal person to take on the role of Data Controller for the LS AAI.

The operators of the LS AAI Proxy must agree to the responsibilities placed on a Data Processor and communicate promptly regarding data protection and data breach issues with the LS-community-designated Data Controller.

3.3. Acceptable Use Policy

The end-user should be informed at initial contact, and at specific moments thereafter, both about the processing of personal data (privacy notice), and for good measure also about acceptable (and non-acceptable) use. We suggest that these notice are always shown together, and that both are permanently available and easily findable.

The end-user shall give unambiguous indication (“INFORM”) that the AUP is accepted during initial enrolment in the Proxy, when the AUP materially changes, or whenever additional consent is required as per DPCoCo version 2.

Users should be asked to re-confirm acceptance at least once a year (in order to satisfy subordinate SP and Infrastructure requirements on AUP reconfirmation).

Acceptable Use Policies traditionally vary a lot between organisations, service providers, and infrastructures. An AUP alignment study [AUPSTUDY] is currently ongoing, and its preliminary results indicate there is one ‘family’ of AUPs that are roughly similar, but beyond that a wider range of quite disparate AUP models. Of these disparate AUPs, many are either project specific and name specific services, or include managerial content (such as sanctions) that are specific to the Infrastructure or organisation. Many organisational AUPs in addition include references to personal use that are not appropriate in this case.

The one ‘family’ of AUPs are all derived from a single source, the Joint Security Policy Group Acceptable Use Policy (2006), whose signature has been preserved over time. It is evident from the ELIXIR AUP, the EGI AUP, but also from e.g. Open Science Grid and XSEDE. Its best evolved form is available from the AARC policy pages [JSPGAUP2], but this version will be evolved as part of the *Policy Development Kit* that is future-scheduled work by the AARC project. This proposed text is compatible with the ELIXIR AUP and the EGI AUP.

In addition, for specific Life Science services there are community-specific acceptable and non-acceptable uses. For example, the BBMRI-ERIC IT AUP explicitly adds the condition that “the user will avoid any attempts to reverse privacy enhancing technologies (i.e. pseudonymization, anonymization) applied to the data”.

We expect (but this is an assumption) that the general requirements on ethical behaviour of the end-user as expressed in the BBMRI-ERIC AUP are similarly applicable and not harmful to the other research infrastructures served by the LS AAI Proxy. This allows a *layered approach* to the construction of the AUP, where the AUP presented to the end-user (on enrolment or later) comprises both the generic JSPG-evolved version plus the community-specific additions.

The model of the JSPG AUP additionally assumes that the community (“the body that grants you access”) defines its common aims and purposes, i.e. the research or scholarship goals. This also addresses the *Snctfi* requirements RC6.

The Proxy shall present an Acceptable Use Policy also on behalf of its connected services and infrastructures.

The Proxy operator shall present as the AUP:

- the common aims and purposes, i.e. the research or scholarship goals of the Life Sciences Research Infrastructures (in a few high-level sentences)

This text must be supplied by the Life Sciences community.

- the list of 11 (eleven) items from the Evolved JSPG AUP [JSPGAUP2]
- the ethical requirements for all the community, for which the text should be endorsed by the Life Sciences community, e.g.:
*“You agree to be a bona fide researcher and use the Infrastructure only for activities that are not inconsistent with legal and ethical requirements or widely recognised good research practice.
 You will avoid any attempts to reverse privacy enhancing technologies (i.e., pseudonymization, anonymization) applied to the data and/or to (re-)identify individual natural persons (such as patients or donors who have consented to and contributed her/his data or biological material to be used in research) contributing the data and/or donating the biological material.”*
- The note that specific services *may* ask the user to sign additional conditions of use.

In addition to a ‘global’ AUP, one may also consider having supplementary AUPs when enrolling in a particular subgroup (gaining access to additional sensitive services). How to handle that is still under study, and lacking a concrete risk assessment no policy guidance can yet be given in this area. Technically, it would imply implementation of mechanisms in the LS AAI user registry (Perun) to present additional more-specific AUPs (on top of the general one) during a subgroup-enrolment flow, and keeping records thereof. Without a proper risk assessment, we cannot give guidance here, and have to defer to the specific services.

More specific conditions of use, above and beyond the one presented by the LS AAI Proxy, are the responsibility of the individual SPs and Infrastructures. It is recommended to keep such cases the minimum necessary.

3.4. Requirements on subordinate SPs and Infrastructures

All service providers and infrastructures (“*Constituents*” in Sncfci) must provide, in a visible and accessible way, a Privacy Policy covering their processing of personal data for purposes that are necessary for the safe and reliable operation of their service, compliant with the Infrastructure policy (or policy framework). This is a requirement both in Sncfci [DP2] but also in the DP Code of Conduct.

The Proxy operators ensure that all connected SPs and Infrastructures have a Privacy Notice that is compatible with DPCoCo

Assessment of compatibility with DPCoCo by connected SPs and Infrastructures is primarily their own task – it would be practically impossible for the Proxy operators to do that assessment externally. Since DPCoCo is a self-assertion (albeit with a mechanism to complain and request action by a management body), it seems fair to ask the SPs and Infrastructures to state compliance. Moreover, especially for version 1, it merely encodes legal requirements for those SPs and Infrastructures that are established in an EU memberstate or EEA country. The assessment has been done already even for global



infrastructures such as EGI, if the BCR-like [BCRLIKE] mechanism is considered sufficient; the EGI SPG policy suite effectively implements all of CoCo version 1.

The privacy notices should be written by and published by the subordinate SPs and Infrastructures themselves, so that they can be linked.

The Proxy operators should also ensure that the service providers have compatible information security practices. This is particularly important in case of data breaches, for which formal time limits are specified in the GDPR. DPCoCo version 2 incorporates *Sirtfi* to address this.

The Proxy operators shall ask a self-assertion to *Sirtfi* from all connected SPs and Infrastructures, and shall ensure and maintain correctness of the security contact address provided.

The Proxy operators shall ask all connected generic e-Infrastructures to state sufficiency of the common AUP for any use made of the Infrastructure with identities mediated by means of the LS AAI Proxy.

The Proxy operators shall recommend that all LS community specific AUPs consider sufficient the common AUP, and shall clarify to any SPs that require additional conditions of use that it is their own duty to inform the end-user and ensure compliance.



Table of Contents

1. Introduction	2
2. Research and Scholarship	2
2.1. Subordinate service providers.....	3
2.2. Identity Provider support	3
3. Data Protection Code of Conduct.....	3
3.1. Snctfi.....	4
3.2. Requirements on the Proxy itself	5
3.3. Acceptable Use Policy	6
3.4. Requirements on subordinate SPs and Infrastructures	8
Table of Contents.....	10
References	11

References

- AUPSTUDY** <https://wiki.geant.org/pages/viewpage.action?pageId=86736956>
- BCRLIKE** https://aarc-project.eu/wp-content/uploads/2016/12/AARC-DNA3.5_Recommendations-for-Processing-Personal-Data_2016_11_07_v4_DG.pdf
- DPCOCO** <https://www.geant.org/uri/Pages/dataprotection-code-of-conduct.aspx>
- DPCOCO2** GÉANT Data Protection Code of Conduct v2_29Jan2018.pdf, see <https://wiki.refeds.org/display/CODE/GEANT+Data+Protection+Code+of+Conduct+workshop+6+February+2018>
- FEDINCR3.2** Framework for a coordinated response to security incidents (DNA3.2) <https://aarc-project.eu/wp-content/uploads/2017/02/DNA3.2-Security-Incident-Response-Procedure-v1.0.pdf>
- JSPGAUP2** <https://wiki.geant.org/pages/viewpage.action?pageId=97945151>
- RS** <https://refeds.org/category/research-and-scholarship>
- SIRTFI** <https://refeds.org/sirtfi>
- SNCTFI** <https://www.igtf.net/snctfi/>