# Research and Scholarship FAQ

## Research & Scholarship Entity Category FAQ



## General Information

### Where is the official Research and Scholarship definition?

The formal, approved definition of the *REFEDS Research and Scholarship (R&S) Entity Category* is published on the REFEDS website:

http://refeds.org/category/research-and-scholarship

(Note that the URI value of the REFEDS entity attribute resolves to the R&S specification.)

### Where can I find support materials?

Advice from federation operators that have already implemented R&S is available. REFEDS has also prepared guidance on the legal justification for R&S and a presentation that can be reused for local training (cf. training material from the Attribute Release Workshop for Federation Operators, TNC2015). If you would like help to provide training for your members please do not hesitate to contact us.

### What do I do if I think a Service Provider is misusing R&S?

Please either contact your local Federation Operator or the REFED Steering Committee (contact@refeds.org). The REFEDS SC has a process in place for reviewing such complaints and working with Federation Operators to address.

## For SP Owners

### What types of services are considered R&S services?

The category definition says that:

> *Candidates for the Research and Scholarship (R&S) Category are Service Providers that are operated for the purpose of supporting research and scholarship interaction, collaboration or management, at least in part.*
>
> *Example Service Providers may include (but are not limited to) collaborative tools and services such as wikis, blogs, project and grant management tools that require some personal information about users to work effectively. This Entity Category should not be used for access to licensed content such as e-journals.*

Broadly this means that R&S is intended for platforms and services used by researchers or scholars where some sort of collaboration, discussion or other interaction between users is required, making the release of personally identifiable information necessary for the service to work properly. This services may be both paid for or freely available services - the focus of the category is on the nature of the service offering and legitimate requirements for attributes.

Think about issues like:

- Is it necessary for a name to be displayed in order for work to be attributed to the user or to show them as the contributor? (a wiki is a prime example)
- Is it necessary for a service to have a user's email address for correspondence such as updates about a grant application? (optional services such as alerting systems that are not part of the core offering would not be considered a good reason for R&S membership).

As a federation, you should be satisfied that the release of personal data is an essential part of the operation of the service (and not purely to activate added features) and that the service in question, whether commercial or not, exists to support research and scholarship as a primary function.

Services that should **not** be included in this category include:

- e-Journal, ebook or other data access, where content may be accessed based on a users affiliation without a need for personal information.
- Services selling products or offering discounts to staff or students based on their affiliation.

### What are the distinguishing characteristics of an R&S service?

*Collaboration* is a sufficient condition for inclusion in the R&S category. Thus a service that functions as a collaborative tool (at least in part) meets the intent of this category.

A wiki is probably the most obvious example of a collaborative service. Other examples include (but are not limited to): calendaring and scheduling tools, content and document management systems, and mailing list software.

Scientific research (broadly defined) is inherently a collaborative endeavor, and so web apps, portals, and computational tools for researchers clearly satisfy the intent of R&S. Collaborative learning platforms for research or education are also candidates for the R&S category.

An important characteristic of collaborative tools and services is that they require the user's name to function effectively. Hence, the R&S attribute bundle includes a name-based identifier (`eduPersonPrincipalName`) and person name as essential attributes. The user's email address is also included in the bundle, to facilitate communication among the users of the service and between the service and its users.

### What exactly is meant by a "production SAML deployment?"

The following REFEDS R&S requirement:

> *4.3.1 The Service Provider is a production SAML deployment that supports SAML V2.0 HTTP-POST binding.*

may be interpreted as the following pair of requirements:

- The Service Provider supports standard SAML V2.0 Web Browser SSO. In particular, the Service Provider has an endpoint in metadata that supports the SAML V2.0 HTTP-POST binding.
- The Service Provider is a production deployment or one of a group of services that together comprise a production deployment.

The latter includes dev and/or staging instances of the overall Service Provider deployment.

### Are SPs allowed to request attributes other than R&S attributes?

Service Providers should only request attributes that the service actually uses. The specification does not explicitly prevent Service Providers from requesting attributes outside the R&S attribute bundle. R&S works best for both Identity Providers and Service Providers when the bundle is treated as the maximal set of attributes requested. The specification gives the following advice:

> *Service Providers SHOULD limit their data requirements to the bundle of attributes defined in Section 5, but MAY negotiate for additional data as required via mechanisms that are outside the scope of this specification.*

The category specifies "SHOULD" so as to not unintentionally disallow scenarios where there is a very good reason to ask for an extra attribute, although providers are encouraged to stick to the R&S bundle where-ever possible. An example exception might be where a contractual arrangement exists and specific attributes (e.g. eduPersonEntitlement) are used to help flag this contractual arrangement.

That said, if an SP requests an attribute outside the R&S attribute bundle, an IdP that supports R&S is by no means required to release it.

### Will I definitely get the attributes requested?

Release of data from organisations is governed by data protection laws that provide a variety of mechanisms to ensure that people and organisations have choice over the data that is released. R&S is designed to safely and securely release appropriate and required data and all IdPs are encourage to release requested attributes. There may however be legitimate reasons for attributes not be release (e.g. user consent, data not available for all users in IDM systems etc.). SPs are encouraged to consider providing helpful error message screens where this may impact service provision.

### Are attributes single or multi valued?

Service Providers should reference the eduPerson specification for details on values that may be received per attribute, but in general terms:

- eduPersonPrincipalName, eduPersonTargetedID, displayName are single-valued.
- givenName + sn, email address, eduPersonScopedAffiliation can be mutli-valued.

## For IdP Operators

### What attributes have to be released by an R&S IdP?

The Research & Scholarship specification defines a bundles of attributes that Identity Providers supporting R&S must release to R&S services:

- email address, person name (either displayName or givenName+sn; ideally both forms to help applications that can only deal with one form), eduPersonPrincipalName
- If eduPersonPrincipalName values may be re-assigend at a given IdP (from one person to another, even after a grace period) a SAML 2.0 persistent NameID (or eduPersonTargetedId attribute, though deprecated) must also be released by that IdP.
  (Persistent NameIDs may not be re-assigned, so R&S SPs that

The bundle also includes one *optional* attribute (everything else above being mandated by the specification):

- eduPersonScopedAffiliation

Service Providers should therefore be prepared to not receive affiliation attributes under R&S, due to their optional nature.
Affiliations in the form of eduPersonScopedAffiliation attribute values have long known to be not widely interoperable (REFEDS Whitepaper, A.Cormack, M. Linden, 2009) particularly in cross-institutional, cross-cultural or international uses. (As pointed out in the conclusion of said whitepaper affiliations are also unsuitable for most kinds of authorization use-cases, them being too "high level"). For these reasons their use within R&S is not emphasized or recommended.

Category support is defined as follows:

> *An Identity Provider indicates support for the R&S Category by exhibiting the R&S entity attribute in its metadata. Such an Identity Provider MUST, for a significant subset of its user population, release all required attributes in the bundle defined in Section 5 to all R&S Service Providers, either automatically or subject to user consent or notification, without administrative involvement by any party.*

Section 7 of the Entity Category gives details of how IdPs should implement support. Effectively managing use of eduPersonPrincipalName and eduPersonTargetedID in relation to reassignment is one of the areas that causes the most confusion. For the avoidance of doubt, REFEDS recommends that if you support both, release both.

### How do I configure an IdP to release attributes to R&S SPs?

To release attributes to all current and future R&S SPs with a one-time configuration, an IdP leverages entity attributes (instead of entity IDs). Thus the configuration steps documented in the R&S IdP Config topic require Shibboleth IdP v2.3.4 or later, which fully supports using entity attributes in SP metadata as part of an attribute release filter policy. No other SAML IdP software is known to support entity attributes at this time.

IdPs are broadly taking one of two approaches to releasing attributes to R&S SPs:

- Configure an IdP to Release a Fixed Subset of R&S Attributes. This releases the same subset to every R&S SP.
- Configure an IdP to Release a Dynamic Subset of R&S Attributes. This releases a different subset to each R&S SP based on the `<md:RequestedAttribute>` elements in SP metadata.

### If an IdP restricts attribute release to some subset of R&S SPs, can that IdP declare support for R&S?

The short answer is no. An IdP must release attributes to **all** R&S SPs before it can assert the REFEDS R&S entity attribute in metadata.

Consider, for example, an IdP that releases the minimal subset of the R&S attribute bundle to any SP that is a member of **both** the Code of Conduct category **and** the Research & Scholarship category. That IdP is **not** eligible to receive the REFEDS R&S entity attribute in its metadata.

As another example, consider an IdP that releases the minimal subset of the R&S attribute bundle to any R&S SP in the InCommon Federation (but no other federation). That IdP may **not** receive the REFEDS R&S entity attribute in its metadata.

Finally, consider the following counterexample. Suppose an IdP releases the minimal subset of the R&S attribute bundle to any R&S SP provided the user is a non-student. That IdP may indeed receive the REFEDS R&S entity attribute in its metadata since it supports the R&S category "for a significant subset of its user population," as required by the REFEDS R&S specification.

## For Federation Operators

### How would a federation operator implement the Research & Scholarship Category?

Some tips and suggestions for implementing the Research & Scholarship Category are given in a separate document. There is also guidance and advice on attribute release and a useful seven step programme that could be used when adding service providers to an entity category.

### What is the difference between Research & Scholarship and the Code of Conduct?

The GÉANT Data Protection Code of Conduct is a process that allows Service Providers to commit to a series of declarations of support for data protection within the context of the EU Data Protection Directive. Like R&S, it results in the application of an entity category tag and is intended to give greater confidence to IdPs when releasing data.

- The Code of Conduct is designed to help IdPs feel more comfortable with the SPs intentions to abide by existing data protection law and therefore have relationship with them, but does not define attribute release and does not work outside of Europe in its current form, although an international version is being explored.
- R&S is designed to help IdPs that are struggling to define any sort of attribute release policies have an easier way of mitigating the risk and designing policies for a small subset of Service Providers that have been through some minimal vetting. It can be used by any federation globally.

## What Federations are Using R&S?

This can be determined using the entities search on https://met.refeds.org/.