eduGAIN Policy Framework Data Protection Profile (OPTIONAL)

Version date	Editor	Change
21.9.2010	ML	Changes based on the Call for Comments

1. Introduction

When releasing Attributes from a Home Organisation to a Service Provider, the Data Protection Directive (Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data) needs often to be taken into account. The directive imposes requirements, some of which are best covered by Home Organisations' and Identity and Service Providers' coordinated functioning. An introduction to relevant articles of the directive and their interpretation in the context of federated identity management is provided in Appendix B.

In the eduGAIN Policy, it has been recognized as an important goal to introduce policies and practices that adapt the data protection directive to the technical infrastructure. This optional data protection profile of eduGAIN defines two categories for Service Providers with different positions with regards to the directive. An Identity and Service Provider and their Home Federation use SAML 2.0 metadata tags defined in this document to indicate their support for this profile.

As the controller of its end user's personal data, the ultimate responsibility of releasing personal data to a Service Provider is in the Home Organisation, who makes its decision based on its local risk management procedures. However, eduGAIN may have thousands of Service Providers, which has made it practical to seek for ways to ease the Home Organisations' burden. This profile may help the Home Organisations in their decisions by mediating privacy-related information on the service from the Service Provider to the Home Organisations using well-defined syntax and semantics.

2. Requirements and categories for Service Providers

~~Service Providers have different characteristics with regards to the end users accessing the Service Provider.~~ Considering the data protection directive's implications, Service Providers are ~~divided~~ grouped into these ~~following~~ two categories:

- category PII: the Service Provider processes personal data
- category non-PII: the Service Provider processes no personal data

PII stands for 'Personally Identifiable Information'. The categories are further elaborated below and summarized as a table in Appendix A.

2.1. Registering to a category

A Service Provider who has adopted this data protection profile registers to either of the two categories. Registering to a category implies that the Service Provider is committed to the functionality this profile requires in that category. A Service Provider cannot manifest conformance to this profile without registering to either of the two categories.

The Service Provider's Home Federation records and mediates the category to other Participant Federations and their Identity Providers ~~using~~ via the exposed eduGAIN~~'s~~ SAML 2.0 metadata. If a Service Provider is registered to the category non-PII, ~~it~~ the Service Provider takes the responsibility of ensuring that:

- the Attribute Requirements registered for it do not contain any personal data in the Service Provider's jurisdiction, and
- the Service Provider resides in an EU/EEA country or a country which ensures adequate level of data protection or that the Service Provider is otherwise committed to an adequate level of protection (e.g. the Service Provider is committed to the US Safe Harbour privacy principles).

Depending on the jurisdiction, some attributes either do or do not count as personal data, or, due to lacking court decisions, the status is unknown. For this data protection profile, it is ~~adviced~~advised to assume that SAML 2.0 Persistent NameID (a.k.a. eduPersonTargetedID) is personal data.

In some ~~juristictions~~jurisdictions, IP ~~adresses~~addresses are considered personal data. IP ~~adresses~~addresses are not released via eduGAIN, but collected directly from the end user. Service Providers who reside in jurisdictions where IP ~~adresses~~addresses are personal data should treat them as such and have adequate legal grounds (i.e. consent or necessity) from end users before collecting them. This also applies to other data that given other added information (e.g. an identifier) can become personal data.

2.2. Service Providers manifesting no category

If a Service Provider does not manifest any category, it is assumed that the Home Organisations, ~~and~~ Identity Providers and Service Providers ~~have~~ will fulfil~~led~~ the obligations set by the data protection directive using an out-of-band mechanism. ~~This is the default for Home Organisations and Identity and Service Provides who have not adopted this profile.~~

2.3. Category PII: SP processes personal data

In category PII, the Service Provider is processing personal data because it receives Attributes ~~which are considered personal data~~ from the Identity Provider which are considered personal data.

The Service Provider can be either

- a data processor, processing personal data on behalf of a Home Organisation, in which case the Home Organisation and Service Provider are supposed to have a written agreement as the basis for processing personal data. For instance, the Service Provider is providing software as a service (SaaS) or licensed contents (e.g. library content) to the Home Organisation, and a related contract is in place between the Home Organisation and Service Provider.

- a data controller, processing personal data not on behalf of the Home Organisation. Instead, release of personal data from the Home Organisation to the Service Provider initiates a new and separate processing of personal data in the Service Provider.

Whether ~~The~~ the Service Provider ~~being~~ is a data processor or data controller may ~~depend on~~vary per the Home Organisation. With some Home Organisations in eduGAIN, ~~T~~the Service Provider may have a data processing agreement and acts as~~with some Home Organisations in eduGAIN, making the Service Provider~~ a data processor~~ for those Home Organisations~~. For the ~~rest of the~~other Home Organisations, the Service Provider may be a data controller.

2.3.1. Purpose of processing

In eduGAIN confederation, personal data is processed in order to support the goal of eduGAIN as defined in the eduGAIN constitution.

A bilateral ~~The~~ data processing agreement~~s~~ signed by ~~the~~ a data controller~~s~~ and a data processor~~s~~ is likely to ~~may~~ be more specific on ~~what is~~ the purpose of processing.

2.3.2. Relevance of personal data processed

See section "2.5. Relevance of Attributes".

2.3.3. Informing the data subject

The Service Provider must make the service's Privacy Policy publicly available. The Service Provider's Home Federation must register a URL to a place where the privacy policy can be found and expose ~~it~~ this URL to the eduGAIN metadata. The privacy policy must be available at least in English and address the issues presented in Article 11 of the data protection directive:

a. the identity of the controller and of his representative, if any;
b. the purposes of the processing;
c. any further information such as
   — the categories of data concerned,
   — the recipients or categories of recipients,
   — the existence of the right of access to and the right to rectify the data concerning him

Before releasing the end user's Attributes to the Service Provider

- for the first time, or
- for the first time after an extension in the Attribute set for this Service Provider

the Home Organisation ~~Identity Provider~~ must provide the Service Provider's ~~clickable~~ privacy policy URL to the end user. This can be done, for instance, when an end user consents, if necessary, to Attribute release (see next section~~)~~ 2.3.4).

The data controller is responsible for informing the end user on processing his/her personal data. If the Service Provider is a data processor~~ and the Home Organisation is the data controller~~, the Service Provider may refer to the Home Organisation in its privacy policy web page.

2.3.4. Criteria for making data processing legitimate

Releasing personal data from a Home Organisation to a Service Provider may be based on necessity or end user's consent.

In Category PII, the Service Provider, being an expert of the service and its use scenarios, makes a proposal on the legal grounds for processing. The Service Provider's Home Federation registers the proposal to the Service Provider's metadata and exposes it to eduGAIN. Based on the proposal, the Service Provider's privacy policy and other information available on the Service Provider, the Home Organisation decides if Attribute release is based on consent or necessity.

To assist Providers in the decision-making, following guidelines and good practice is provided:

- A service that is related to an employee doing his/her work is usually based on necessity
- A service that is related to a student taking his/her ~~cources~~courses and otherwise being educated is usually based on necessity

The process for informing the end user (see section 2.3.3) and asking his/her consent for attribute release may vary. If the end user is a child, giving the consent may also involve his/her parents.

When an end user logs in to a Service Provider for the first time,

- If Attribute release is based on consent, the Home Organisation~~Identity Provider~~ provides the end user the following or equivalent text "I am informed on release of my personal data to the service and consent to it <OK> <Cancel>"
- If Attribute release is based on necessity, the ~~Identity Provider~~Home Organisation provide~~r~~s the end user the following or equivalent text "I am informed on release of my personal data to the service <OK>"

In both cases, the ~~Identity Provider~~Home Organisation must provide to the end user a clickable link to the Service Provider's privacy policy (see the previous section 2.3.3). A way to integrate this to the login sequence in an Identity Provider is proposed in section 4.5.

If an end user wants to withdraw his/her consent later, he can use the contact information in the privacy policy to submit a request to the Service Provider to remove his/her personal data.

2.4. Category non-PII: No personal data processed

In Category non-PII, the Service Provider does not process personal data and the directive is not applied to the Attribute release and the Service Provider.

2.4.1. Relevance of personal data processed

See the next section "2.5. Relevance of Attributes".

It is a responsibility of the Service Provider to ensure, that:

- the Attribute Requirements registered for it does not contain any personal data in the Service Provider's jurisdiction, and
- the Service Provider resides in an EU/EEA country or in a country which ensures adequate level of data protection or that the Service Provider is otherwise committed to an adequate level of protection (e.g. the Service Provider is committed to the US Safe Harbour privacy principles).

2.5. Relevance of Attributes

Irrespective of which category PII or non-PII the Service Provider belongs to, the Home Federation must register the Attribute Requirements of a Service Provider. The Home Federation publishes the Attribute Requirements in the Service Provider's SAML 2.0 metadata entry exposed to eduGAIN.

It is assumed that the Service Provider, which is the expert of the service, carefully balances its Attribute Requirements with the data protection directive and its national implementation before registering it to the Home Federation. The Home Federation and eduGAIN confederation takes no legal responsibility on the Attribute Requirements a Service Provider has registered.

Additionally, the Service Provider may register one or several statements made by one or several trusted third parties (TTP) on Attributes the TTP deems relevant for the service. It is up to the Home Organisation

- to decide if it trusts the statement and
- make an out-of-band agreement with the TTP on any legal responsibilities the TTP takes by the statement.

Attributes revealing data that the data protection directive defines as sensitive personal data ~~racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life~~ should not be released in eduGAIN.

3. Registering a Home Organisation's conformance

The Home Federation registers a Home Organisation's manifest that is has adopted this data protection profile. Registering implies that the Home Organisation is committed to the functionality this profile requires from a Home Organisation. A Home Organisation can manifest support to the ~~categoty~~category non-PII Service Providers, category PII Service Providers, or both.

The Home Federation records and mediates the Home Organisation’s manifest of conformance to this profile to other Participant Federations and their Service Providers using eduGAIN’s SAML 2.0 metadata.

If a Home Organisation does not manifest conformance to this profile, it is assumed that the Home Organisation and the Service Providers ~~have~~ will fulfil~~led~~ the obligations set by the data protection directive using an out-of-band mechanism. This is the default for Home Organisations and Identity and Service Provides who have not adopted this profile.

Technical implementation

This section defines how the data protection mechanisms introduced in this document are technically expressed in the Identity and Service Providers' SAML 2.0 metadata entity elements. A new XML namespace mddp is introduced with one XML element `DataProtectionProperties`, having three child elements:

1. `Category` to indicate the category an SP belongs to and the categories an IdP supports,
2. `LegalGrounds` to indicate the legal grounds for processing as suggested by the Service Provider, and
3. `saml:Assertion` to embed any signed Trusted Third Party statements to the metadata.

Additionally, SAML 2.0 metadata specification is used to indicate the attributes the Service Provider requests, and La Joie: IdP Discovery and Login UI Metadata Extension Profile (version 1.0, DRAFT 03, 29 March 2010) to indicate the Service Provider's Privacy Policy's URL.

4.1. Provider's category indication

A Service Provider uses Category element to indicate which category the Service Provider belongs to. An Identity Provider uses the same element to indicate which categories the Identity Provider supports.

The element is placed to the Provider's metadata extensions element as a child element of the DataProtectionProperties element. The category is expressed using the values "non-PII" and "PII" and implementations should ignore the case.

Example (Service Provider):

```
<SPSSODescriptor>
   <md:Extensions>
      <mddp:DataProtectionProperties>
         <mddp:Category>PII</mddp:Category>
      </mddp:DataProtectionProperties>
   </md:Extensions>
```

Example (Identity Provider):

```
<IDPSSODescriptor>
   <md:Extensions>
      <mddp:DataProtectionProperties>
         <mddp:Category>PII</mddp:Category>
         <mddp:Category>non-PII</mddp:Category>
      </mddp:DataProtectionProperties>
   </md:Extensions>
```

4.2. Relevance of personal data

In its eduGAIN SAML 2.0 metadata element, the Service Provider uses the RequestedAttribute element defined by SAML 2.0 Metadata standard to indicate the Service Provider's Attribute Requirements. The isRequired XML attribute should be set to "true" if the service does not open to the user (not even using some lower level of functionality) without releasing the Attribute.

Example:

```
<SPSSODescriptor>
   <AttributeConsumingService ...>
      <RequestedAttribute
         NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri"
         Name="urn:oid:2.5.4.4" isRequired="true"/>
      <RequestedAttribute
         NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:uri"
         Name="urn:oid:2.5.4.42" isRequired="false"/>
   </AttributeConsumingService>
```

Additionally, if the Service Provider wants to register a trusted third party's statement on necessary Attributes to its metadata entry, it can place a SAML attribute assertion in its EntityDescriptor. The contents and the semantics of the assertion are out of scope for this profile.

Example:

```
<SPSSODescriptor>
   <md:Extensions>
      <mddp:DataProtectionProperties>
```

```
            <saml:Assertion>
               ... a TTP statement here...
            </saml:Assertion>
         </mddp:DataProtectionProperties>
     </md:Extensions>
```

4.3. Informing the data subject

In its eduGAIN SAML 2.0 metadata element, the Service Provider uses the PrivacyStatementURL element defined in La Joie: IdP Discovery and Login UI Metadata Extension Profile (version 1.0, DRAFT 03, 29 March 2010) to indicate where the Service Provider's Privacy Policy can be found. The element must be in place for any category PII Service Provider.

Example:

```
   <SPSSODescriptor>
      <md:Extensions>
         <mdui:UIInfo>
            <mdui:PrivacyStatementURL xml:lang="en">
               http://www.example.org/privacypolicy.html
            </mdui:PrivacyStatementURL>
         </mdui:UIInfo>
      </md:Extensions>

```

4.4. Criteria for making data processing legitimate

In its eduGAIN SAML 2.0 metadata element, the Service Provider must use a LegalGrounds element to indicate what the Service Provider proposes as the legal grounds for processing personal data in the Service Provider (see section 2.3.4). The values should be in lowercase. The implementations must ignore the case.

consent The data subject gives unambiguously his consent (see article 11 a of the directive)
necessity Release of personal information is necessary (see article 11 b-f of the directive)
The Home Organisations may use this information to decide if Attribute release is based on necessity or consent. The element must be in place for any category PII Service Provider.

Example:

```
   <SPSSODescriptor>
      <md:Extensions>
         <mddp:DataProtectionProperties>
            <mddp:LegalGrounds>consent</mddp:LegalGrounds>
         </mddp:DataProtectionProperties>
      </md:Extensions>
```

. . .

4.5. Identity Provider behaviour

An Identity Provider relying on the data protection mechanisms provided in this profile must, before releasing any Attributes, ensure that:

- the Service Provider manifests conformance to category PII or non-PII, and
- only Attributes a category non-PII Service Provider requests are released to it, and
- only necessary Attributes are released to a category PII or non-PII Service Provider. The Identity Provider may use the RequestedAttribute information, privacy policy URL and trusted third party statements available in the Service Provider's metadata entry to construct its Attribute Release Policy.

Sections 2.3.3. and 2.3.4 introduced two requirements for Home Organisations:~~For a Service Provider in category PII, the Identity Provider must also:~~

- inform the end user by providing him/her a clickable link to the Service Provider's privacy policy and
- ask him/her to consent, if necessary, to the Attribute release.

The Home Organisations may, of course, use any processes (e.g. printed and signed documents) to fulfil these requirements. However, in the front-channel binding of SAML 2.0 web single sign-on, a practical way could be that after authenticating the end user but before releasing his/her attributes to the Service Provider, s/he is presented a web dialogue which covers the two steps.

4.6. Service Provider behaviour

~~A Service Provider relying~~Relays a Service Provider on the data protection mechanisms defined ~~provided~~ in this document and ~~belonging~~ belongs to category PII, the Service Provider must~~, before accepting any Attributes,~~ ensure that the Identity Provider manifests conformance to category PII before the Service Provider accepts any attributes.

4.7. “Multi faced” Service Providers~~ which have "multiple faces"~~

It is possible that an SP's category and other properties vary depending on from which Home Organisation the end user logs in. For instance, some library content may be licensed to some Home Organisations as an expensive site license and as a cheaper per-user license to another.

For such Service Providers, it is suggested that

- the Service Provider registers ~~several~~ multiple entries (with ~~separate~~ distinct entityIDs)~~ in the metadata~~, or
- the SP is not registered to eduGAIN at all

Not so many SPs are assumed to face this issue.

APPENDIX A: A summary of Service Provider categories

	No category (default) Data protection covered out-of-band	Category PII: the SP processes personal data	Category non-PII: the SP processes no personal data
1. Description			
	eduGAIN is not involved in fulfilling the obligations imposed by the data protection directive. The providers must use an out-of-band mechanism.	The SP processes personal data, which may be released to the SP from an IdP. For a Home Organisation, the SP may be - a data processor, processing personal data on behalf of the Home Organisation, or - a data controller, not processing personal data on behalf of the Home Organisation.	No personal data is passed to the SP from the IdP.
2. The directive and how the category covers it			
2.1.Purpose of processing (Directive's article 6.1(b))	N/A	"To support the goal of eduGAIN." If the Service Provider is a data processor, the data processing agreement may define a more specific purpose.	N/A. Personal data is not processed
2.2.Relevance of personal data processed (Article 6.1 c)	N/A	The SP's Home Federation registers the SP's Attribute Requirements and provides them as part of the SAML2 metadata. Additionally, the metadata may contain a trusted third party's statement on what Attributes it deems necessary for the service. Sensitive personal data should not be released.	<- the same The SP must ensure that the Attribute Requirements do not incorporate personal data.

	No category (default) Data protection covered out-of-band	Category PII: the SP processes personal data	Category non-PII: the SP processes no personal data
2.3.Informing the data subject (Article 11)	N/A	The SP's Home Federation registers the SP's Privacy policy's location in the <PrivacyStatementURL> element in the SAML2 metadata. When the Attribute release from the Home Organisation to the SP takes place for the first time, the IdP must provide this clickable link to the end user.	Not needed. Personal data is not processed
2.4.Criteria for making data processing legitimate (Article 7)	N/A	The SP proposes the criteria for making data processing legitimate. Based on the proposal, the Home Organisation decides if Attributes are released based on consent or necessity. It is assumed that in most use scenarios in eduGAIN, Attribute release is based on necessity.	N/A. Personal data is not processed
2.5.Withdrawal of consent	N/A	If the Attribute release is based on consent, the end user can contact the data controller's representative e.g. by mail. It is assumed that withdrawal of consent is not very frequent procedure.	N/A. Personal data is not processed.
2.6.Release of personal data to 3rd countries	N/A	Personal data can be released to countries with adequate level of data protection just as it is released to EU/EEA countries. The Service Provider ensures that it resides in EU/EEA or in a country with adequate level of protection.	N/A. Personal data is not processed. Attributes can be freely released to 3rd countries.
2.7. Receiving personal data from 3rd	N/A	Personal data can be received from 3rd countries in a similar way they are received from Home	N/A. Personal data is not processed. Attributes can be freely received from 3rd

	No category (default) Data protection covered out-of-band	**Category PII: the SP processes personal data**	**Category non-PII: the SP processes no personal data**
countries		Organisations in EU/EEA.	countries.
3. Technical implementation			
3.1.How providers manifest conformance to this category	This is the default category which is implied if a provider does not manifest any other categories.	IdPs and SPs manifest their conformance to this category by adding a tag "PII" to its metadata.	IdPs and SPs manifest their conformance to this category by adding a tag "non-PII" to its metadata.
3.2.IdP behavior during login	The IdP must use an out-of-band mechanism to ensure that the obligations imposed by the data protection directive are fulfilled.	Before releasing Attributes to a Category PII SP, the IdP must ensure that the SP manifests conformance to Category PII in the SAML2 metadata.	Before releasing Attributes to a Category non-PII SP, the IdP must make sure that the SP manifests conformance to Category non-PII in the SAML2 metadata.
3.3.SP behavior during login	The SP must use an out-of-band mechanism to ensure that the obligations imposed by the data protection directive are fulfilled.	Before accepting Attributes from a Category PII IdP, the SP must ensure that the IdP manifests its conformance to Category PII in the SAML2 metadata.	No requirements to the SP behaviour.
4. Examples			
		eduroam trouble ticketing system (TTS), eduroam wiki, CLARIN	Library contents

Table A.1~~2~~: Service Provider Categories.

APPENDIX B: Selected sections of the directive and a confederation

This appendix discusses the directive's articles which are particularly interesting for federated identity management and the eduGAIN confederation. The other provisions of the directive and its implementations are naturally binding, as well, but the provisions presented here have been identified as those who need coordinated functionality from the Home Organisations and Identity and Service Providers in eduGAIN.

B.1. Objective of the directive

The objective of the directive is to protect natural persons' fundamental rights while guaranteeing the free flow of personal data between member states. Thus, the directive can be seen as an enabler, not as a disabler, of eduGAIN, provided the Attribute release in eduGAIN is implemented in a way that follows the provisions of the directive.

> *Article 1*
>
> *1. In accordance with this Directive, Member States shall protect the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of personal data.*
>
> *2. Member States shall neither restrict nor prohibit the free flow of personal data between Member States for reasons connected with the protection afforded under paragraph 1.*

B.2. Definition: personal data (Article 2a)

> *'Personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;*

It is obvious that common Attributes such as end user's full name (cn), email address (mail) and unique identifier (eduPersonPrincipalName) are personal data. However, it is questionable if Attributes such as privacy-preserving bilateral identifiers (eduPersonTargetedID/SAML2 Persistent identifier) are personal data. This question is shortly discussed next.

The only property the eduPersonTargetedID Attribute has is that it has the same value when the same end user visits the same service again. The interpretation of the expression *relating to an identified or identifiable natural person* seem to vary country by country. The directive seems to make no difference between *identification* and *recognition*, the latter meaning that the service notices the end user is the same one who has visited the service earlier, although it does not know who s/he is in the real life.

This case is fundamentally similar to the use of an IP address; the end user is recognized by his/her IP address, but an end user's identity cannot be deduced from it. There is some case law available in the Member States; some German court (Berlin Regional Court 23 S 3/07) has decided IP address being personal data, another German court has decided that it isn't (Munich district court 133 C 5677/08). It is obvious that it is hard to get a pan-European interpretation if IP address or eduPersonTargetedID is personal data. Thus, to be in the safe side, in eduGAIN project, it should be assumed that eduPersonTargetedID is personal data.

It is also worth noticing that if several Attributes are coupled together (as they usually do) and one of them is personal data, then all the Attributes are personal data. For instance, an end user's role (the eduPersonAffiliation Attribute) in his/her Home Organisation is not personal data alone, but put together with his unique identifier (eduPersonPrincipalName) it becomes personal data, too.

B.3. Definition: processing of personal data (Article 2b)

> *'Processing of personal data' ('processing') shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;*

Based on the definition, it is obvious that a Home Organisation processes personal data and the directive is applied to it. However, as Home Organisations in eduGAIN confederation typically already maintain user accounts for their end users, being a Home Organisation who has registered an Identity Provider to a federation does not change their status there.

Service Providers are processing personal data if they collect any information (either Attributes from an Identity Provider or directly from the end user him/herself) that is considered to be personal data.

A common interpretation of the directive is that when an Identity Provider passes Attributes carrying personal data to the Service Provider, the Identity Provider disseminates an end user's personal data to the Service Provider - even althought, technically, in the front-channel binding of the SAML 2.0 authentication request protocol, it is the end user who uses his/her web browser to carry the SAML assertion to the Service Provider. There is no known case law where this assumption is verified. If it turns out that an Identity Provider is not passing personal data to the Service Provider but it is the end user him/herself, then most requirements presented in this document collapse. An end user can, naturally, do whatever with his/her personal data.

Some federations have a distributed architecture, each Home Organisation operating an Identity Provider of their own. The role of the federation operator is typically to maintain a trusted list of all registered Identity and Service Providers. In such a federation, the federation operator is not processing personal data (except possibly a list of Identity and Service Provider administrators' and their contacts). On the other hand, if the federation operator is also operating Identity Provider(s) on behalf of the Home Organisations, they are processing personal data, too (and, thereof, have probably a data processor status as will be introduced in the next section).

The confederation operator does not process personal data (except possibly a list of participant federations' administrators' and their contacts).

B.4. Definition: data controller and processor (Article 2d,e)

> *'Controller' shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law;*

> *'Processor' shall mean a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller;*

A research and higher education institution, which has registered an Identity Provider to eduGAIN, is typically processing affiliated end users' personal data in order to support research and education in the institution. In other words, the Home Organisation is a data controller and has determined that the purpose of processing is to support institutions primary functions which are, in general, research and education.

The Service Provider's position as a data controller or processor depends on the service. When the Service Provider is a subcontractor of the Home Organisation, the Service Provider is a data processor processing personal data on behalf of the Home Organisation. This is the case e.g. if the Service Provider provides licensed content (e.g. library content) or Software as a Service (SaaS) to the Home Organisation. Article 17 of the directive makes it explicit that the data processor must have a written contract with the Identity Provider.

> *3. The carrying out of processing by way of a processor must be governed by a contract or legal act binding the processor to the controller...*

> *4. For the purposes of keeping proof, the parts of the contract or the legal act relating to data protection and the requirements relating to the measures referred to in paragraph 1 shall be in writing or in another equivalent form.*

In a (con)federation, bilateral agreements between Home Organisations and Service Providers are not expected, in general. The scalability benefits of a (con)federation are questionable, if the directive enforces bilateral written agreements between each Home Organisation and Service Provider, anyway. Fortunately, the directive leaves the door open for Service Providers who are not data processors but data controllers. In this case, release of personal data from a Home Organisation starts a new and separate processing of personal data in the Service Provider.

In a (con)federation, it is also possible that a Service Provider is processing personal data on behalf of some Home Organisation(s) with whom it has a data processing contract, but is also willing to grant access to end users from other Home Organisations. In this case, the Service Provider is a data processor for some Home Organisations and an independent data controller with regards to the other Home Organisations. The (con)federation does not have built-in mechanisms to keep track of the bilateral agreements the Home Organiations and Service Providers may have. Thus, it is safe to assume that each Service Provider is a data processor for some Home Organisations and a data controller with regards to the other Home Organisations.

The data protection directive is applied both to data controllers and processors, but the obligations imposed differ slightly. For instance, it is the obligation of the data controller, not the data processor, to inform the end user on processing his/her personal data. In a confederation spanning multiple jurisdictions, it is also necessary to notice that the jurisdiction follows the data controller. More obligations are introduced in the next section.

B.5. Security of processing (Article 17)

> *1. Member States shall provide that the controller must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing.*
>
> *Having regard to the state of the art and the cost of their implementation, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the data to be protected.*

This section makes it an obligation of a controller to make necessary measures to protect personal data, in particular when it is transmitted over a network, which is the case in federated identity management. Having this eduGAIN data protection profile and Service Providers manifesting conformance to it in place is supposed to be part of the *appropriate technical and organisational measures* that Home Organisations can rely on.

On the other hand, the article lets the controllers balance the obligation with the implementation costs, risks and the nature of the data. It can be argued that personal data released via eduGAIN do not represent significant risks. Especially, there seems to be no need to release Attributes which Article 8 defines sensitive:

> *Article 8*
>
> *1. Member States shall prohibit the processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and the processing of data concerning health or sex life.*
>
> ...

B.6. Purpose of processing (Article 6.1b)

> *Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.*

As noticed above, the institution as the controller of affiliated end user's data has defined the purpose of processing personal data. In a research and education institution, the purpose typically follows from the institutions charter, and is, in general, to support research and education.

Following the directive, the institution needs to obey to this purpose also when it, acting as a Home Organisation, releases Attributes to a Service Provider. The purpose of processing personal data in the Service Provider may not conflict with the purpose of processing in the Home Organisation. For instance, a Home Organisation is not conflicting with the directive when releasing student's data to a Learning Management

System in another university, but releasing students' personal data to a gambling service is hardly "supporting research and education".

B.7. Relevance of the personal data processed (Article 6.1 c)

> *Personal data must be adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed.*

A Service Provider may process only those Attributes that are necessary for the service, whether gathered from the end user him/herself, from a Home Organisation or from some other source. In federated identity management, relevance of personal data translates to the principle of "minimal disclosure"; an Identity Provider may release only relevant Attributes to a Service Provider.

In an identity federation, the concept of an Attribute Release Policy (ARP, having its origins in the Shibboleth software) is commonly used for expressing which Attributes an Identity Provider releases to which Service Providers. For scalability reasons, in a large (con)federation, some centralized mechanism to mediate Service Providers' Attribute Requirements to all Home Organisations and their Identity Providers is desirable. It can be assumed that the Service Provider is in a key role here; the Service Provider is the expert of the service.

B.8. Informing the data subject (Article 11)

> *Information where the data have not been obtained from the data subject*
>
> *1. Where the data have not been obtained from the data subject, Member States shall provide that the controller or his representative must at the time of undertaking the recording of personal data or if a disclosure to a third party is envisaged, no later than the time when the data are first disclosed provide the data subject with at least the following information, except where he already has it:*
>
> a. the identity of the controller and of his representative, if any;
>
> b. the purposes of the processing;
>
> c. any further information such as
>
> — the categories of data concerned,
>
> — the recipients or categories of recipients,
>
> — the existence of the right of access to and the right to rectify the data concerning him
>
> *in so far as such further information is necessary, having regard to the specific circumstances in which the data are processed, to guarantee fair processing in respect of the data subject.*

The data controller needs to inform the end user on processing his/her personal data. For a Home Organisation, informing the end user is obvious and can be done when a new end user gets his/her account at the institution. The Service Provider's obligation depends on if it is a data processor or a controller. As a data controller, a Service Provider is responsible for providing this information to the end user. As a data processor a Service Provider can refer to the Home Organisation.

In the Internet, a standard practice to inform the end user on processing his/her personal data in services is to provide him/her a Privacy Policy web page in the service.

A convenient place to inform the end user is when the Attribute release takes place for the first time, and several federations in European higher education and research have already developed tools for that (e.g. the uApprove module implemented for Shibboleth,the consent module implemented for SimpleSAMLphp). Informing the end user can be conveniently bundled to the step where the end user, if necessary, consents to Attribute release, which is going to be discussed next.

B.9. Criteria for making data processing legitimate (Article 7). Withdrawal of consent

> *Personal data may be processed only if:*
>
> *(a) the data subject has unambiguously given his consent; or*
> *(b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; or*
> *(c) processing is necessary for compliance with a legal obligation to which the controller is subject; or*
> *(d) processing is necessary in order to protect the vital interests of the data subject; or*
> *(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed; or*
> *(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Article 1 (1).*

In short, this article can be summarised that data processing can be based either on consent or necessity. If based on consent, it must be freely given (an end user must have an option to say no) and informed (an end user must understand what s/he consents to. See the previous section). A property of consent is that it can be withdrawn any time:

> *(Article 2 h) 'The data subject's consent' shall mean any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed.*

Alternatively, data processing may be based on being necessary, for instance, for

- providing education to a student (c, e)

- a teacher, researcher or other employee to do the jobs his/her employer has assigned to him/her (b)

However, deciding if a service is necessary or not is cumbersome. If a student is taking a cource which is mandatory in his/her curriculum, release of personal data to the course's learning management system is probably necessary, but what if the course is optional? If a researcher is using licensed contents related to his/her subject of research, release of personal data is probably necessary, but what if the researcher is browsing contents outside his subject of research? As a result, decision if Attribute release is based on consent or necessity becomes a complex function of (Service Provider, end user, time).

There seems to be two interpretations of this article; in some countries, consent is the primary way of making data processing legitimate. In other countries, consent should be used only as a last resort, and the desirable way is to base processing of personal data on necessity, whenever possible.

After all, it is worth noticing that consent does not override the other obligations imposed by the directive, including the purpose of processing, relevance of personal data processed and informing the data subject. It is wrong to assume that anything can be done with an end user's personal data if s/he consents to it.

B.10. Release of personal data to 3rd countries

Personal data may be released to other EU and EEA (Norway, Iceland, Lichenstein) countries as it is released within an EU/EEA country. The directive recognises that also some non-EU/EEA countries (dubbed as 3rd countries in the directive) may have adequate level of data protection. Personal data can be released to those countries just as it is released to any EU/EEA country. In federated identity management, this principle is applied to non-EU/EEA Service Providers.

The European Commission publishes a list of countries with adequate level of protection. For instance, in Switzerland and Argentina, data protection laws ensure adequate level of protection. Canada has sector-specific data protection legislation, and the protection is adequate if the Canadian data controller is subject to the Personal Information Protection and Electronic Documents Act. In the United States, the level of data protection is adequate if the data controller is committed to the "Safe Harbor privacy principles" that the US Department of Commerce and the Commission have agreed on.

The Service Provider's jurisdiction follows the data controller. If the Service Provider is a data controller, the Service Provider's local laws on data protection are applied to the Service Provider. If the Service Provider is a data processor (i.e. processes personal data on behalf of the Home Organisation), the Home Organisation's laws are applied.

To release personal data to countries who do not guarantee adequate data protection, the level of protection must be ensured in an agreement with the data recipient. In federated identity management, the attribute release takes place between the Home Organisation and the Service Provider, who should sign a bilateral agreement which commits the Service Provider to an adequate level of protection. In a (con)federation, bilateral contracts are not expected in general, which suggests that this data protection profile cannot be used by Service Providers who are not bound to an adequate level of protection by the local law or the US Safe Harbour privacy principles. This does not exclude e.g. US Service Providers or even federations from eduGAIN, but their data protection issues must be solved using some other mechanism.

B.11. Receiving Personal data from 3rd countries

The directive is applied to processing personal data in EU/EEA, regardless of the Service Provider processing personal data on behalf of a data controller in a 3rd country or not. However, if the Service Provider is a data processor and the data controller is in a 3rd country, the directive expects the data processor to have a representative in EU/EEA to ensure the directive can be enforced:

> *(Article 4) 1. Each Member State shall apply the national provisions it adopts pursuant to this Directive to the processing of personal data where:*
>
> ...
> *(c) the controller is not established on Community territory and, for purposes of processing personal data makes use of equipment, automated or otherwise, situated on the territory of the said Member State, unless such equipment is used only for purposes of transit through the territory of the Community.*
>
> *2. In the circumstances referred to in paragraph 1(c), the controller must designate a representative established in the territory of that Member State, without prejudice to legal actions which could be initiated against the controller himself.*

If the Home Organisation outside EU/EEA has a data controller/processor relationship with any of the Service Providers in EU/EEA, it needs a representative in EU/EEA. On the other hand, if the Service Provider in EU/EEA is a data processor for a non-EU Home Organisation, it needs to have a written agreement with the non-EU Home Organisation anyway (see section B.4.), and the EU/EEA representative is covered there. Thus, for simplicity, in the (con)federation agreement, the requirement for a non-EU/EEA Home Organisation having a representative in EU/EEA can be omitted. The Home Organisation does not need to reside in a country which quarantees adequate level of data protection, either.

Appendix C.

C.1. Open issues

- Currently, Data protection profile covers only Identity and Service Providers, but the eduGAIN Policy Framework recognises also other Entities such as Attribute Providers. In principle, from Data Protection perspective,, Attribute Providers are like Identity Providers, but there isn't necessarily a front-channel binding for Attribute Requests, which makes the implementation of this Profile more difficult for them.~~What if an SP changes its requested attributes over time? How do IdPs get informed? Should requested attributes in the metadata have a timestamp, as well, and IdP-side consent modules store the timestamp value at the time user consents? At least uApprove does not need it, but everyone is not using uApprove...~~

- ~~If there is a distinction of "required attributes" (isRequired="true") and "desired attributes" (isRequired="false"), will it make a difference for a user consenting to attribute release?~~

- ~~References~~

- ~~**<Delete this section if not required>**~~

- ~~**[REF e.g. [GGF NM-WG]]** URL e.g. http://www.internet2.edu/presentations/jtcolumbus/20040720-piPEfitters-Simar.ppt~~

- ~~**[REFERENCE]** URL~~

- ~~**[REFERENCE]** URL~~

- ~~**[REFERENCE]** URL~~

•

Glossary

~~**<Delete this section if not required>**~~

~~**ACRONYM** Definition – which should automatically wrap to align correctly if the definition is longer than a single line, as this line should demonstrate~~

~~**ACRONYM** Definition~~

~~**ACRONYM** Definition~~

~~**ACRONYM** Definition~~

~~1.1. Terms~~

General terms:

AAI	Authentication and authorisation infrastructure.
Attribute Provider	An organisation which is responsible for managing additional identity data (attributes) for end users authenticated by a Home Organisation. Also a server that is acting in an Attribute Provider role as defined in SAML 2.0. In this document, an Attribute Provider refers to an attribute provider who is a Member of a Participant Federation and whom the Participant Federation has exposed to eduGAIN.
DANTE	Delivery of Advanced Network Technology to Europe. The GÉANT network is managed by DANTE.
Entity	Entity means an AAI endpoint described with a SAML 2 EntityDescription. An Entity can be, for instance, an Identity Provider, a Service Provider or an Attribute Provider. In this document, an Entity refers to an entity that a Participant Federation has exposed to eduGAIN.
Federation	(identity federation) An association of organisations that come together to exchange information as appropriate about their users and resources in order to enable collaborations and transactions.
GÉANT	Gigabit European Academic Network project, the pan-European data network dedicated to the research and education community. The GÉANT network is managed by DANTE.
Home Organisation	The organisation which the end users are affiliated to and which is responsible for managing end users' identity data (attributes) and authenticating them. Home Organisation is responsible for setting up and operating either one or more Identity Providers, either by itself or via an outsourced service. In this document, a Home Organisation refers to a home organisation who is a Member of a Participant Federation and whose Identity Provider the Participant Federation has exposed to eduGAIN.
Identity Provider	A server acting in an Identity Provider role as defined in SAML 2.0 specifications. In this document, an Identity Provider refers to the Identity Provider that a Participant Federation has exposed to eduGAIN.
Member	Any organisation that has signed an agreement with a federation operator to cover the

	verification and publication of metadata. In this document, Member refers to a member whose Entity is exposed to eduGAIN.
NREN PC	The Policy Committee of the GÉANT network and project, which consists of appointed representatives from each partner in the project. It is responsible for setting and overseeing overall policy of the GÉANT network and project.
OT	eduGAIN Operational Team, as defined in section 2.3.
Participant Federation	A Federation which has passed the joining process defined in section 3.3.
Policy Framework	(eduGAIN Policy Framework) This document, the profiles supplementing it and the eduGAIN Policy Declarations signed by Participant Federations.
Service Provider	An organisation that is responsible for offering the end user the service s/he is going to log in to. Also a server that is acting in a Service Provider role as defined in SAML 2.0. In this document, a Service Provider refers to a service provider who is a Member of a Participant Federation and whom the Participant Federation has exposed to eduGAIN.
TSG	eduGAIN Technical Steering Group, as defined in section 2.2.

Additional terms introduced in this Profile

Attribute	An end user's identifier (e.g. name, mail address, eduPersonPrincipalName or persistent NameID), role or other property that an Identity Provider releases or may release to a Service Provider.
Attribute Release Policy	An Identity Provider's decision and related configuration regarding which Attributes will be released to a given Service Provider.
Attribute Requirements	A list of Attributes a Service Provider requests from an Identity Provider
Home Federation	The eduGAIN Participant Federation to which an Identity or Service Provider has been registered and which exposes the Provider to the eduGAIN confederation
~~Home Organisation~~	~~The organisation which the end users are affiliated to and which is responsible for authenticating end users and maintaining their Attributes. Home Organisation is responsible of setting up and operating an Identity Provider, either by itself or as an outsourced service. In this document, a Home Organisation refers to an organisation whose Identity Provider a Participant Federation has exposed to eduGAIN~~
~~Identity Provider~~	~~A server acting in an Identity Provider role as defined in SAML 2.0 specifications. In this document, an Identity Provider refers to the Identity Provider that a Participant Federation has exposed to eduGAIN~~
~~Service Provider~~	~~An organisation that is responsible for offering the end user the service s/he is going to log in to. Also a server that is acting in a Service Provider role as defined in SAML 2.0. In this document, a Service Provider refers to a Service Provider that a Participant Federation has exposed to eduGAIN~~

~~Glossary~~