# GÉANT Data Protection Code of Conduct 2.0

Explanatory memorandum

2 January 2019



The research leading to these results has received funding from the Horizon 2020 programme under Grant Agreement No. 731122 (GN4-2).

# Contents

# Summary

Modern research is highly cross-national and collaborative. Researchers in thousands of research organisations access services in other organisations using electronic means. GÉANT develops and operates a pan-European e-infrastructure for authentication, authorisation and single sign-on for researchers (or "End Users") in research and education organisations in Europe.

Such e-infrastructure provides researchers with an easy access to cross-national research infrastructures without the use of additional credentials, such as usernames and passwords. Accessing the research e-infrastructure requires sharing some of a researcher's personal data from his/her research or education organisation (or "Home Organisation") to the service provider he/she is accessing in another Member State or beyond (or "Service Provider Organisation").

Home Organisations must comply with relevant data protection legislation to share End Users' personal data with Service Providers.. However, the lack of scalable, practical compliance guidelines has caused e-infrastructures in different Member States to adopt fragmented practices for personal data release, which is hindering cross-national access to services.

Discomfort about data protection compliance has also led to an approach where Home Organisations do not share End Users' personal data with the Service Provider as a precautionary measure. Consequently, End Users need to register and create separate credentials for each service, undermining the reliability of the credentials, inhibiting access to services, causing administration overheads and resulting in less efficient pan-European academic research.

GÉANT, as a cross-border European academic research and education e-infrastructure, drafted a pan-European Code of Conduct to facilitate the application of the data protection principles stemming from the General Data Protection Regulation, taking account the specific characteristics of the processing carried out in the academic sector, and respecting the national provisions adopted by Member States.

The attached GÉANT Data Protection Code of Conduct proposes a scalable approach to protect the End User's personal data when he/she logs in to a service provided by another Home Organisation. This Code of Conduct presents a harmonized approach to which Service Provider Organisations in the European Economic Area and beyond can adhere when receiving End Users' personal data from the Home Organisations. Home Organisations will feel more comfortable to release affiliated End-User personal data to

the Service Provider if they can see that the Service Provider has taken the necessary measures to properly protect the data.

A formal endorsement of the pan-European Code of Conduct by the competent Supervisory Authority, in accordance with the procedure provided by Article 40 of the GDPR, would enhance the credibility of the Code of Conduct. The GDPR endorses the use of codes of conduct to provide guidance on the GDPR's requirements and it acts as a signal to data subjects, third parties and regulators that our organization is in compliance with the GDPR. Moreover, an approved Code of Conduct provides with appropriate safeguards for international transfers of data.

For the purpose of the hereby Explanatory memorandum, Capitalized terms must have the meaning of the definitions provided by Appendix 4 of GÉANT Data Protection Code of Conduct

# 1. Context

## *1.1. Who we are*

GÉANT is the pan-European research and education backbone network that interconnects Europe's National Research and Education Networks (NRENs). Together with its national partners (the NRENs), the GÉANT network offers network connectivity and associated services (such as, an e-infrastructure for electronic identity) to over 10 000 research and education institutions in 42 countries, including all EU Member States.

The GÉANT network is developed and operated by the GÉANT project (GN4-3), which receives funding from the Horizon2020 programme of the European Commission. The project is coordinated, and the network is operated by GÉANT Association, the leading collaboration on network and related infrastructure and services for the benefit of research and education.

As the coordinator of the GÉANT project, GÉANT Association has submitted this Code of Conduct to the competent Supervisory Authority, on behalf of the GÉANT project partners.

## *1.2. Access to research infrastructures*

The success of the European economy is increasingly dependent on scientific and technological innovation. The European Research Area (ERA) was launched at the Lisbon European Council in March 2000. The development of ERA is needed to overcome the fragmentation of research in Europe along national and institutional barriers.

According to the EC's vision of the ERA, "*major infrastructures should be built and exploited in the form of joint European ventures. They should be accessible to research teams from across Europe and the world, with researchers working in Europe having access to international infrastructures and equipment in other parts of the world. These research infrastructures should be integrated, networked and accessed through the concomitant development of new generations of electronic communication infrastructures* (also known as "e-infrastructures")*, both in Europe and globally."*[1]

[1] European Commission. The European Research Area: New perspectives. Green Paper 4.4.2007.

In the strategy report on the European Research Infrastructure roadmap,[2] ESFRI identified 50 research infrastructure landmarks and projects from the broad disciplines of energy, environment, health & food, physical sciences & engineering and social & cultural innovation. The construction costs of the individual projects and the capital value of the landmarks are between EUR4 M and EUR1843 M, and annual operation costs range from EUR0,6 M to EUR234 M. Many of the research infrastructures rely on the availability of secure and reliable e-infrastructures.

For security reasons, researchers usually need electronic identities for authentication and verification of access rights before they can access the research infrastructure. In its communication on the ERA[3] in 2012, the European Commission invites Member States to "*adopt and implement national strategies for electronic identity for researchers giving them transnational access to digital research services* and research stakeholder organisations, to *implement and promote the uptake of electronic identity and digital research services".* The ERA communication also reflects the 2010 Digital Agenda for Europe, whose Key Action 3 calls for a framework for *cross-border "recognition and interoperability of secure eAuthentication systems".*[4]

Without a proper e-infrastructure for electronic identities, the ERA researchers need to manage credentials for thousands of services, inhibiting effective co-operation and research and creating administrative burdens.A new approach, **Federated Identity Management** has been introduced -which is described in the next section - to provide an e-infrastructure for secure authentication, authorisation and single sign-on of researchers and other End Users,

## *1.3. Federated Identity Management*

Federated Identity Management is crucial to the ability of e-Infrastructures to operate in a scalable, secure manner. However, Federated Identity Management requires a trust framework to exchange some of the researcher's personal data to operate.

[2] European Strategy Forum on Research Infrastructures. Strategy Report on Research Infrastructures. Roadmap 2016.

[3] European Commission. A Reinforced European Research Area Partnership for Excellence and Growth. COM(2012) 392, 17.7.2012.

[4] European Commission. A Digital Agenda for Europe. COM(2010) 245, 26.8.2010.

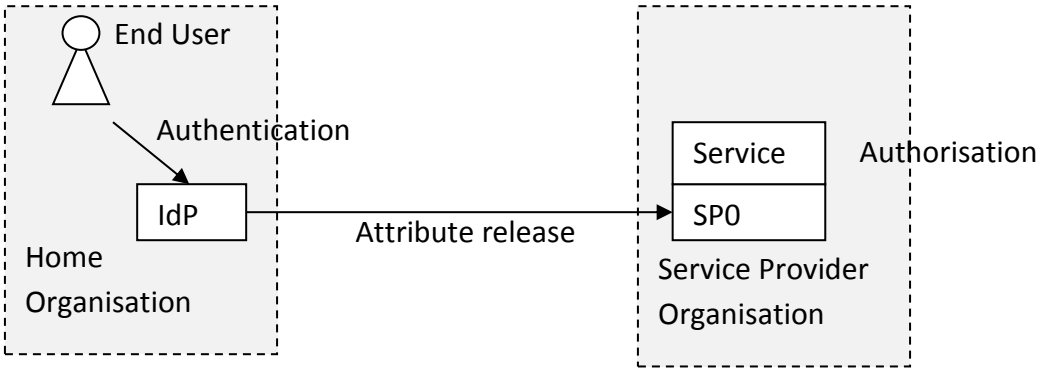



*Figure 1. The basic actors in Federated Identity Management and their division of responsibility.*

The basic division of responsibility is presented in Figure 1. **End Users** are researchers, teachers, students and other natural persons affiliated to a university, research institution or other Home Organisation. **Home Organisations** (such as universities and research institutions) are responsible for managing their End Users' personal data (called *Attributes*, such as name, e-mail address and role/position in the Home Organisation) and for delivering them credentials (such as usernames and passwords) to allow identified access to local and federated services. Now when an End User accesses a Service, his/her Home Organisation authenticates him/her using such credentials. The authentication is performed by an **Identity Provider** server (IdP) that the Home Organisation operates.[5] After authenticating the End User, the Identity Provider delivers the End User's Attributes to the **Service Provider Organisation** (SPO) the End User is accessing. The Attribute delivery takes place at the time when the End User accesses the service. Based on the Attributes, the Service Provider Organisation decides if the End User is entitled (authorised) to use the **Service**, and assigns him/her to his/her existing profile in the Service, if any.

By virtue of the Federated Identity Management, the End User only needs a single set of credentials to access all Services, and can enjoy a single sign-on experience. Service Provider Organisations do not need to maintain End Users' credentials, and are able to receive End Users' up-to-date and reliable attributes from their Home Organisations. When an End User departs, the Home Organisation closes his/her account, and access to the research infrastructures ceases.

[5] The Identity Provider server can be operated by the Home Organisation or the Home Organisation can outsource it. In some countries, the federation operator operates a centralized Identity Provider for the Home Organisations.

## *1.4. The relationship triangle*

From a data protection perspective, there are three parties and three relationships involved (see Figure 2).

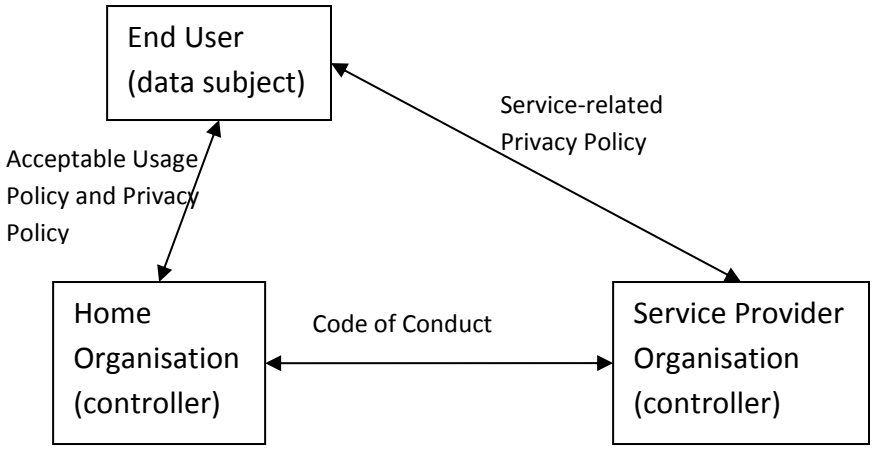



*Figure 2. The triangle of relationships involved in Federated Identity Management.*

The Home Organisation processes the End User's personal data to carry out its duties as an employer, to provide teaching to a student, or to fulfil other tasks of the institution. The relationship between the End User and the Home Organisation is governed by an **Acceptable Usage Policy** (typically including or linked to a privacy notice), to which the End User commits at the start of his/her employment, studies, or other relationship with the Home Organisation, and receives his/her credentials (such as, username and password) from the Home Organisation.

The Home Organisation delivers the End User's Attributes to the Service Provider Organisation to enable the researcher or student to conduct research, take courses, etc. The relationship between the End User and the Service Provider Organisation is set by the Service's purpose of data processing and is documented in the Service's **Privacy Notice**.

In some circumstances, the Home Organisation and Service Provider Organisation may have a bilateral agreement that also addresses protection in the End User's personal data. However, to guarantee a baseline for data protection in the instances where there is no bilateral agreement, the **Code of Conduct** helps the Service Provider Organisation to engage with the Home Organisations on data protection practices.

## *1.5. Academic Identity Federations*

The federated approach to identity management presented above suffers from scalability problems when End Users from thousands of Home Organisations need to access thousands of Services. To address this issue with scalability, e-infrastructures called

*Identity Federations* have been established in the research and education sector since 2005.

Today, academic Identity Federations are nationally focused and typically organised by NRENs, the non-profit organisations providing Internet connectivity and associated services to researchers, teachers and students in the institutions. Currently, there are 27 academic Identity Federations in EU Member States (Malta being the only country with no federation). In August 2018, those federations had together 6786 IdP and 8108 SPO servers registered.[6]

The role of the **Identity Federation** (Figure 3) is to set the rules for its **Members** (i.e. Home Organisations and Service Provider Organisations), such as the eligibility to join the federation, minimum requirements for strength of End User authentication, liability, technical specifications, etc. The Identity Federation also nominates an organisation to be responsible for operations of the federation. The **Federation Operator** (typically, the NREN) takes care of day-to-day issues, such as registering Members' Identity and Service Provider services to the federation and managing a current list of them, providing technical support to the Members, etc.

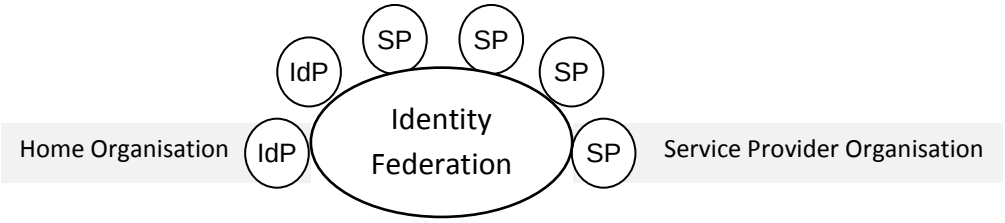



*Figure 3. An Identity Federation is an e-infrastructure service that allows Identity Provider (IdP) servers (managed by End Users' Home Organisations) and Service Provide Organisation (SPO) servers (managed by SPO organisations) to exchange Attributes on authenticated End Users.*

Typical services to which academic Identity Federations provide authentication and authorisation are scientific journals and other licensed content (such as university libraries' electronic magazine subscriptions for researchers), e-learning (teachers' virtual learning tools to support their teaching), collaboration (wikis and fileshares for cross-organisation workgroups) and SaaS services (universities' outsourcing of administrational services to the private sector). Recently, however, Identity Federations have started to face increasing demand to support multinational research infrastructures. This has created the need to bridge the formerly national federations into a cross-national interfederation service called eduGAIN.

[6] GÉANT: Metadata Explorer Tool. https://met.refeds.org/met

## *1.6. eduGAIN interfederation service*

The eduGAIN service (delivered by the EC-funded GÉANT GN4-3 project) connects the national academic federations into an interfederation service (see Figure). Within their Home Organisation, End Users can use their local usernames and passwords to access Service Provider Organisation in another Participant Federation (countries).

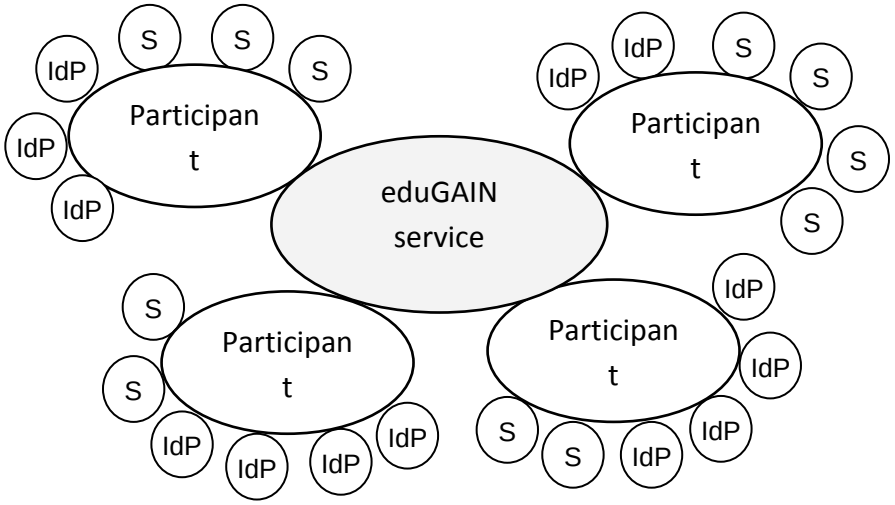



*Figure 4. eduGAIN interfederation service connects its participants and allows a direct exchange of Attributes between Identity Providers (IdP) and Service Provider Organisation (SPOO).*

The technical architecture is distributed as shown in the Figure above. eduGAIN mediates IdPs' and SPOs' technical descriptions (such as the address of the server endpoints and their trusted certificates) between the Participant Federations. There is no single proxy server that channels the entire message exchange in eduGAIN. Any release of personal data takes place directly between the IdP and SPO.

The eduGAIN interfederation service establishes a pan-European e-infrastructure for electronic identity in the ERA. It eases co-operation of researchers and other End Users throughout Europe and reduces the administrative burden needed to access research infrastructures. eduGAIN also supports the European single market introduced in Article 3 of the Treaty on European Union and is interconnected with similar e-infrastructures in other parts of the world.

In August 2018, eduGAIN had 2820 IdPs and 2144SPO registered. 1896 (67%) of the IdPs and 1443(67%) of the SPOs were registered by the Participant Federations that reside in the EU countries and the rest in other parts of the world, Including the United States, Brazil, Canada and Switzerland.

# 2. GÉANT Data Protection Code of Conduct

GÉANT Data Protection Code of Conduct aims to help GÉANT Community to apply the General Data Protection Regulation (GDPR) by providing specific and accessible guidance on how the GDPR should be applied in the research and education sector. Adherence to GÉANT Data Protection Code of Conduct will enable GÉANT's members to demonstrate compliance with the GDPR and facilitate enhance the free flow of personal data amongst them. The proposed Code of Conduct may be found in Appendix A.

This Code of Conduct observes the data protection principles stemming from the General Data Protection Regulation (GDPR), and respecting the national provisions adopted by Member States.

This Code of Conduct constitutes a binding community code for the Service Provider Organisations that have committed to it.

## *2.1. Goal and approach*

The 26 EU Member States' academic Identity Federations have been developed independently, with fragmented approaches to data protection issues often primarily embedded in specific national laws. A goal of the Code of Conduct is to

(i) seek for a pan-European interpretation and good practice approach for the issues related to data protection in Federated Identity Management in research and education, and
(ii) promote pan-European interpretation and practice to eduGAIN Participant Federations, in order to reduce obstacles hindering cross-national access to research infrastructures.

These requirements do not intend to replace or lower any requirements of specific data protection regulation for e-Infrastructures and parties involved in federated identity. Rather, the aim of the Code of Conduct is to identify and highlight key aspects for the research and education community in a scalable way.

The Code of Conduct covers issues related to the protection of the End User's personal data received from his/her Home Organisation when he/she accesses the Service. The Code of Conduct does not cover data protection issues related to the content of the service, such as, how biological research data can be shared for research purposes if it contains potentially sensitive patient data.

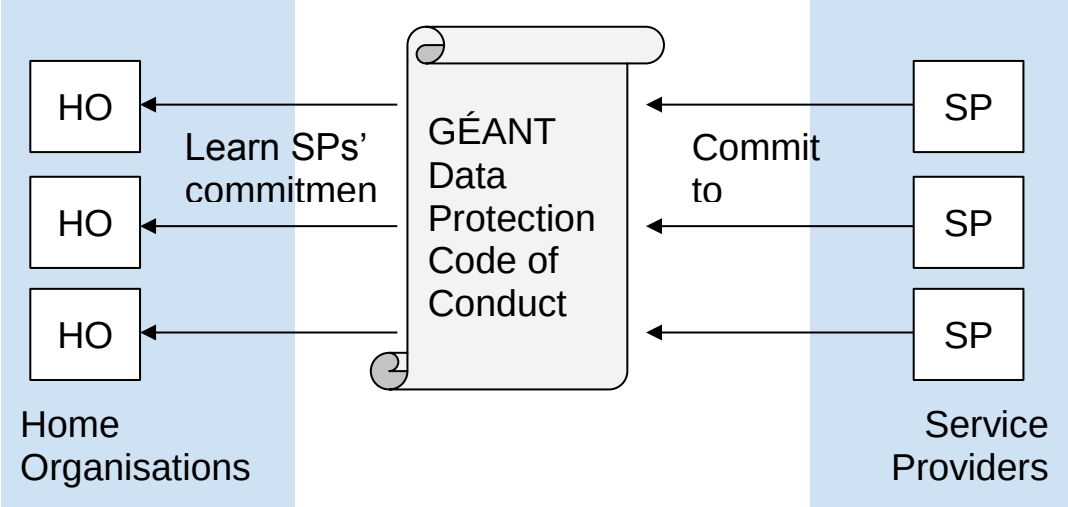



*Figure 6. Service Provider Organisation (SPO) commit to the Data protection Code of Conduct. Home Organisations (HO) learn the Service Provider Organisations' commitment by the indications relayed by the Identity Federation(s).*

The intention (depicted in Figure 6) is that the Service Provider Organisation commits to the Code of Conduct for a particular Service, the Identity Federations (and the eduGAIN interfederation service) relay an indication of the commitment to the Home Organisations, who then make the decision about releasing Attributes to the Service.

A pilot on the Code of Conduct has been started within the eduGAIN community. In September 2018, 157 Service Provider Organisations have expressed their commitment to the Code of Conduct version 1.0 which is based on the Data protection directive.

Although designed for the eduGAIN interfederation service, it is intended that the Code of Conduct can also be used locally within the national Identity Federations.

## *2.2 Who can adhere to the Code of Conduct?*

The Code of Conduct is addressed to any Service Provider Organisation established in any of the Member States of the European Union and in any of the countries belonging to the European Economic Area (all Member States of the European Union, Iceland, Liechtenstein and Norway).

Furthermore, Service Provider Organisations established in any third country offering an adequate level of data protection in the terms of the article 45 of the GDPR can also subscribe to this Code of Conduct.

The GDPR gives the opportunity to Service Provider Organisations that do not fall under the territorial scope of the Regulation (Article 3, territorial scope) and that are established outside of the EEA to join an approved Code of Conduct in order to provide

appropriate safeguards within the framework of transfers of personal data to third countries or international organisations under the terms referred to in point (e) of Article 46(2).

## *2.3. Principles of processing*

Following article 40.2 of the GDPR, the Code of Conduct specifies the application of the GDPR for online access management in the research and education sectorregarding the following principles:

- fair and transparent processing;
- the legitimate interests pursued by controllers in specific contexts;
- the collection of personal data;
- the pseudonymisation of personal data;
- the information provided to the public and to data subjects;
- the exercise of the rights of data subjects;
- the measures and procedures referred to in Articles 24 and 25 of the GDPR and the measures to ensure security of processing referred to in Article 32 of the GDPR;
- the notification of personal data breaches to supervisory authorities and the communication of such personal data breaches to data subjects;
- the transfer of personal data to third countries or international organisations; or
- out-of-court proceedings and other dispute resolution procedures for resolving disputes between controllers and data subjects with regard to processing, without prejudice to the rights of data subjects pursuant to Articles 77.

### *2.3.1. Personal data*

Pursuant to the GDPR, “personal data means any information that relates to an identified and identifiable individual.

Typical Attributes released to the Service by a Home Organisation include the End User’s online identifier (if possible, pseudonomised), name and email address. It is often necessary to also release Attributes describing the End User’s relationship with the Home Organisation, such as the name of the Home Organisation and the organisational unit the End User is affiliated to, and his/her role in the Home Organisation (e.g. student, researcher, member, library walk-in, etc). In some cases, there is a requirement to exchange more sophisticated Attributes, e.g. those based on a person’s entitlement to use a specific resource. Such Attributes, alone or together, qualify as personal data.

The Code of Conduct aims to provide guidance on how to process such personal data for online management purposes in compliance with the requirements provided by the GDPR.

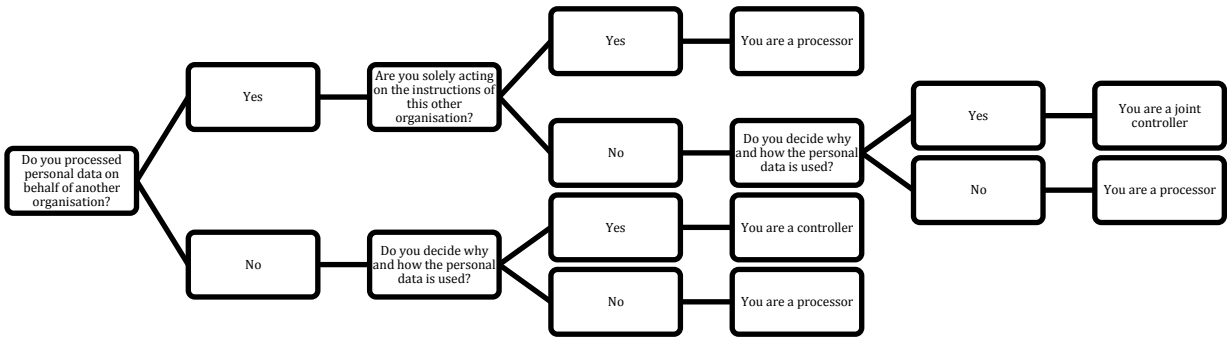
However, it is limited to the processing of Attributes which are released for enabling the End User to access the Service. That means processing activities of personal data for purposes other than enabling the End User to access the Service are not covered by the Code of Conduct. Nevertheless, Service Provider Organisation can decide to commit to the GDPR also for other Attributes.

### *2.3.2. Controller/Processor*

It is essential for organisations involved in the processing of personal data to be able to determine whether they are acting as a Controller or as a Processor in respect of the processing since it allows such an organisation to determine its responsibilities and obligations.

Pursuant to the GDPR, a “Controller” means a person who (either alone or jointly or in common with other persons) determines the purposes for which and the manner in which any personal data are, or are to be processed. That means that if a Service Provider Organisation decides the “why” and “how” the personal data should be processed it is the Controller. A “Processor” means any person who processes the data on behalf of the Controller.

In practice, it can be difficult to assess whether an organization is acting as a Controller or as a Processor. In order to assess whether they are acting as Controller or Processor, Service Provider Organisation may use the below decision tree:




The Code of Conduct is addressed to Service Provider Organisations acting as Controllers.

In the context of the Code of Conduct, a Home Organisation acts as a data controller as to the wider relationship with the End User (for example operating the Identity Provider (IdP) server in respect of the Attributes). An Agent who operates the IdP server on behalf of the Home Organisation acts as a data processor. This includes also the Federation Operators who operate a (potentially centralised) IdP server on behalf of the Home Organisation.

A Service Provider Organisation acts as a data controller in respect of the Attributes, processing them for the purposes of enabling access to the Service. In certain circumstances a Service Provider Organisation may be acting as a data processor, acting on behalf and as instructed by the Home Organisation. A Service Provider Organisation can also manage extra attributes of an End User and further become an Attribute Provider. In such a case, the Service Provider Organisation acts as a Data Controller.

### *2.3.3. Legal grounds*

Pursuant to the GDPR, for a processing of personal data to be lawful, a specific legal ground needs to be identified. There are six options depending on the purposes for the processing of personal data and the relationship among the individuals concerned. "Legitimate interests" is one of the possible grounds on which the processing of

personal data can be based(i.e. grounds on which the processing of personal data can relied upon).

Reliance on the controller's legitimate interests as the legal ground for the processing of personal data in the terms of Article 6.(f) is generally justified for both Home Organisations and Services.

Taking into account the WP 29 Opinion 06/2014 on the notion of legitimate interests of the Controller, we note that both Home Organisations and Service Provider Organisations have an interest in ensuring the Services offered by Service Providers Organisations to be available and accessible to all End Users in a way that limits the administrative burden for End Users. Furthermore, Home Organisations and Service Provider Organisations have an interest in avoiding a situation where the procedure to gain access to the Services becomes an impediment to the actual use of the Service.

We note that this interest (i) is lawful and not contrary to applicable EU and national laws; (ii) is clearly articulated; and (iii) represents a real and present interest for the Home Organisations and Service Provider Organisations; and as such, **constitutes a "legitimate interest"** in the sense of article 6.f of the GDPR.

Furthermore, we note that the Attribute-release mechanism, as described above, is indeed **necessary** to achieve the interest pursued. There are no reasonably workable, less-invasive means to reach the identified purpose of the processing. For instance, the Service asking these attributes directly from the End User may lead to an identity theft or unauthorised access as elaborated in clause d of the Code of Conduct.

In general, it is believed that there are **no overriding fundamental interests or rights** of End Users that would prevent reliance on the legitimate interests' legal ground, because a.o.

- The interest of the controllers, providing the End User with the service he/she has requested in an easily accessible manner, coincides with a wider, public interest, i.e. fostering scientific and technological innovation.
- Without the processing taking place, Service Provider Organisations would not be able to provide access to the Service for eligible End Users.
- The impact on End Users is very limited, compared to the benefit the system brings to End Users as well as Home Organisations and Service Providers Organisations.
- Furthermore, Federated Identity Management implements **additional safeguards.** Its design goal is to protect the privacy of End Users and to minimise the amount of personal data being exchanged. Instead of having End Users register and provide

credentials directly to Services, the Home Organisations will channel the data exchange through Identity Provider servers.

- The Code of Conduct fully supports the aim of privacy protection and data minimisation. For example, Service Provider Organisations must explicitly indicate which End User Attributes are required through the technical infrastructure and will not receive any additional information. Furthermore, pseudonymisation will be used, when possible.

### *2.3.4. Purpose of processing*

Pursuant to the GDPR, processing of personal data is only permitted if the data are collected for specified, explicit and legitimate purposes and no further processed in an incompatible manner with those purposes. This requirement aims to ensure that theController is transparent about the reasons for obtaining personal data, and that what it will do with the data is in line with the reasonable expectations of the individuals concerned.

In the context of Federated Identity Management, the purpose of processing personal data is to share only the necessary End User personal data from his/her Home Organisation with the Service Provider Organisation to allow the Service Provider Organisation to **enable access to the service for eligible End Users**. The term enabling access is further elaborated in clause b of the Code of Conduct. However, a Service Provider Organisation shall not further processpersonal data in a manner that is not compatible with the purpose of enabling access to the services for eligible End Users.

### *2.3.5. Data minimization*

In accordance with the GDPR, Service Provider Organisations must ensure that the personal data they process is:

- adequate – sufficient to properly fulfil their stated purpose;
- relevant – has a rational link to that purpose; and
- limited to what is necessary – they do not hold more than they need for that purpose.

#### *2.3.5.1. Relevance of Attributes*

The Service Provider Organisation has Service expertise as well as understanding of the Attribute requirements of the Service. For reasons of scalability, the Code of Conduct (clause d) takes an approach where the Service publishes a list of its Attribute requirements and takes full responsibility to ensure Attributes are adequate, relevant and not excessive, in relation to the purpose of processing.

#### *2.3.5.2. Keeping Attributes up-to-date*

It is assumed that the Home Organisation, which has a close relationship with the End User, is well placed to ensure that End Users' Attributes are kept up-to-date. Therefore, the Attributes are periodically refreshed from the Home Organisation to the Service, commonly every time the End User logs in to the service. The End User should contact his/her Home Organisation to rectify these Attributes, if necessary.

However, many Services also want to maintain some additional personal data in the service (for example, the End User's preferences in that particular service). The Service's Privacy notice instructs the End User how to rectify this data.

#### *2.3.5.3. Limitation of the information requested*

Only strictly necessary information should be obtained. However, in the context of the Code of Conduct, under no circumstances is a Service Provider Organisation is authorised to request End User's Attribute revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, genetic data, biometric data for the purposes of uniquely identifying a natural person or data concerning health or sex life or sexual orientation.

### *2.3.6. Informing the End User*

Individuals have the right to be informed about the collection and use of their personal data. This is a key transparency requirement under the GDPR. More specifically, the Service Provider Organisation shall provide at least the following information to the End User:

- the name, address and jurisdiction of the **Service Provider Organisation**; where applicable;
- the contact details of the data protection officer, where applicable;
- the purpose or purposes of the processing of the Attributes;
- a description of the Attributes being processed as well as the legal basis for the processing;
- the third party recipients or categories of third party recipient to whom the Attributes might be disclosed, and proposed transfers of Attributes to countries outside of the European Economic Area;
- the existence of the rights to access, rectify and delete the Attributes held about the End User;
- the retention period of the Attributes;
- a reference to the Code of Conduct;
- the right to lodge a complaint with a supervisory authority.

According to the clause f of the Code of Conduct, Service Provider Organisations are obliged to publish a prominent link to the service's privacy notice in the Service's landing page. End Users have an opportunity to study the privacy notice before they log in and their Attributes are released from the Home Organisation to the Service.

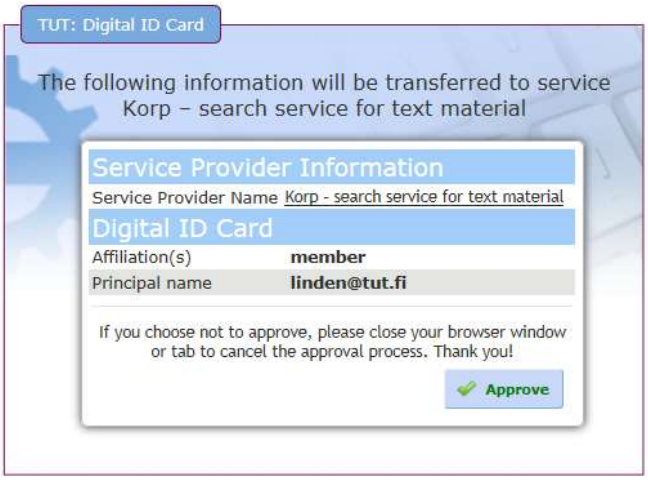



*Figure 5. Example of how a Home Organisation should inform End Users and provide an opt-out opportunity before Attributes are released to a new Service Provider. Clicking the Service Provider Name leads to its Privacy policy page.*

Additionally (see Figure 5), it is recommended that the Home Organisations ensure that their duty to inform the End User is fulfilled by providing End Users with the Service's name, description, clickable privacy notice link and requested Attribute names and values before the Attribute release takes place for the first time. This is described in detail Appendix B.

### *2.3.7. Security of processing*

A key principle of the GDPR is that personal data are processed securely by means of appropriate technical and organisational measures protecting the personal data against unauthorised or unlawful processing and against accidental loss, destruction or damage.

The technical specifications used in federations[7] recommend that message or transport layer encryption is used to ensure the confidentiality of the Attributes when transmitted over the network to the Service. The integrity of the Attributes is guaranteed by mandatory digital signatures created by the Identity Provider server. Confidentiality and integrity are supplemented by registering the Identity and Service Provider servers' trusted certificates. The appendix 2 of the Code of Conduct introduces further

[7] Currently, most federations make use of the SAML 2.0 standard (Security Assertion Mark-up Language) and the SAML2int profile derived from it.

information security, technical and organisational guidelines for the Service Provider Organisations, based on the well-established practices (Sirtfi framework) in use in the research and education sector.

The distributed architecture of eduGAIN has no central points that would be able to examine the Attribute exchange between the Identity Provider and Service.

### *2.4. Monitoring body*

The Monitoring body is an independent entity that oversees monitoring compliance with the Code of Conduct and of ensuring that the Code of Conduct is appropriately robust and trustworthy. The Monitoring body shall have an appropriate level of expertise in relation to the subject-matter of the Code of Conduct.

.

## 3. Requirements from the research and education sector

This section presents the requirements for the GÉANT Data Protection Code of Conduct from the Research and Education sector in Europe and beyond.

### *3.1. Scalability*

The Code of Conduct should scale to the size where thousands of Service Providers Organisation and Home Organisations exchange End Users' Attributes. Bilateral negotiations and/or contracts between each Home Organisation and Service Provider Organisation will not adequately fulfil this requirement.

### *3.2. Balance risk with the ease of collaboration*

The Code of Conduct should find an appropriate balance between minimising the data protection risks and enabling easy collaborations for End Users in different Home Organisations.

However, the Code of Conduct does not preclude a Home Organisation entering bilateral negotiations with a Service Provider Organisation in the case where the safeguards offered by the Code of Conduct are not enough e.g. for particularly sensitive work.

### *3.3. Sustain and recover from incidents and misbehaving entities*

In an ideal world, all parties would meet their respective data-protection-related requirements, and all parties would be able to assume that all other parties will also meet their data-protection-related responsibilities. There would be no security

incidents; no End Users would think that their privacy has been breached. In practice, however, that world does not exist. It is necessary to develop a design that provides all parties with sufficient assurance that other parties with which they interoperate are aware of and meet their data-protection-related responsibilities, and that all parties are able to tolerate and recover from incidents and misbehaving entities.

## *3.4. Minimise the risk of a Home Organisation's liability for a Service Provider Organisation's misbehaviour and vice versa*

it is the Home Organisation's role to release Attributes to a Service Provider Organisation when the Attribute released supports the End User's research, teaching and study-related activities. However, if it turns out that the Home Organisation may become liable for a data protection problem caused by a Service Provider Organisation, the Home Organisations may become hesitant and refrain from their role as an Identity Provider (and *vice versa*). Instead, the End Users would need to self-register and learn local usernames and passwords, which would defeat the benefits of Federated Identity Management.

Thus, it is important that the Code of Conduct clearly allocates responsibility for compliance with data protection rules and for possible breach of these rules to Home Organisations and Service Provider Organisations. It must be avoided that there is (a) a loophole where not all obligations imposed by the GDPR are covered by a party; or (b) a situation where a party can be held liable for another party's misbehaviour.

The Code of Conduct must therefore provide a clear and effective allocation of obligations and responsibilities between all parties.

## *3.5. Suggest good practices to the Home Organisations*

The intention of the Code of Conduct is to reduce Home Organisations' hesitation to release Attributes to the Service Provider Organisations by providing a tool to support risk analysis. This is achieved by clearly documenting behavioural rules for Service Provider Organisations. The rules enable the Home Organisation to decide whether it is safe to release personal data to the Service Provider Organisation. It should be noted that this only applies to the Home Organisation's release of data to the Service Provider Organisation. The Code of Conduct is a Code of Conduct for *Service Provider Organisations*, not for Home Organisations. The Code of Conduct may introduce good practices that can help a Home Organisation to reduce its data-protection-related risks, but if the Home Organisation decides to ignore them, it is first and foremost the Home Organisation, not the Service Provider Organisation, that is exposed to risk.

## *3.6. Minimise the Federation Operator's role*

The Attribute exchange takes place between the Home Organisation and the Service Provider Organisations. In most academic Identity Federations, the Federation Operator is not involved in the transaction.[8] The Code of Conduct therefore does not cover the Federation Operator's liability for data protection issues caused by the Home Organisation or the Service Provider Organisation.

## *3.7. The importance of a global approach*

The research and education community are global, and research collaborations cross national and federation borders. The approach adopted should be applicable globally, or at least as widely as possible.

[8] In some countries, the Federation Operator operates a centralised Identity Provider on behalf of the Home Organisations. In those countries, the Federation Operator is involved in the transaction as a data processor.

## APPENDIX A: GÉANT DATA PROTECTION CODE OF CONDUCT

## APPENDIX B: HOW THE HOME ORGANISATION SHOULD INFORM THE END USER ON THE ATTRIBUTE RELEASE

The Data protection laws create a set of requirements for the INFORM interactions with the user. This Data protection Code of Conduct recommends a division of responsibility where the INFORM interaction is carried out by the **Home Organisation** of the user, for instance, in an INFORM Graphical User Interface (GUI) installed at the Identity Provider server, whereas the data release is requested and subsequently processed by the Service Provider Organisation.

However, the Data protection regulators and the groups developing and enforcing these regulations recognise that there is a balance between full disclosure to meet the requirements and usability. A poor design of the user interaction screens can reduce the likelihood that users will understand what is happening.

## LAW REQUIREMENTS

### INFORMING THE END USER ("INFORM INTERACTION")

For a Home Organisation, informing the end user can be done when a new end user gets his/her account at the institution. At that time, the Home Organisation has the first opportunity to inform that the user's Attributes may also need to be released to a Service Provider Organisation when he/she wants to access it. However, the law requires the End User to be informed about the specific Attribute release every time the Attributes are to be released to a new Service Provider Organisation.

The Service Provider Organisation's obligation to inform the End User depends on whether it is acting asProcessor or a controller. As a Controller, the Service Provider Organisation is responsible for communicating with the End User the issues above; which Attributes it will be using, and what it will be doing with them. As aProcessor, a Service Provider Organisation can refer to the Home Organisation.

The European Data Protection Board, the EU advisory body contributing to the uniform application of the General Data Protection Regulation , took the view that the information must be given directly to individuals - it is not enough for information to be "available[9]". In the Internet, a standard practice to inform the end user on processing their personal data in Services is to provide them with a Privacy Notice web page in the Service.

In the Web Single Sign-On scenario of SAML 2.0, a convenient place to inform the end user is at the Home Organisation before the Attribute release takes place for the first time. Several federations supporting the European higher education and research communities have already developed tools implementing this approach (e.g. the uApprove module implemented for Shibboleth, Consent-informed Attribute Release system (CAR) module implemented for Shibboleth, the consent module implemented for SimpleSAMLphp). This allows the user's decision to directly affect the transfer

[9] Opinion 15/2011 on the definition of consent, p.20.

of Attributes to the Service Provider Organisations; if the Service Provider Organisations were communicating with the user it might have already received all the Attributes and values.

## GENERAL PRINCIPLES FOR INFORMING THE USER

Information dialogues should be short and concise.

The UK information commissioner proposes a "layered approach"[10], the basic information should appear on the main page, and a hyperlink shall be provided for detail. Merely having a clickable link labelled "Privacy Notice here" probably wouldn't be enough.

The goal is to provide a human readable form as the primary interface with the ability to click further to see what the 'technical' data is. The Acceptable Usage Policies presented by most Internet services do not suffice as they are rarely read nor understood by the users. The basic information should be provided as short accurate "user-friendly" descriptions; detailed information about "exactly what's going on" can be provided as a link.

Consequently, this profile recommends displaying the **Service Provider Organisation's** name, description, logo and requested attributes on the main page. If a user wants to learn more, they can click a link resolving to the **Service Provider Organisation's 's** Privacy policy. It is possible that users will actually not do the latter, but at least they have the choice to inform themselves of what is going on.

Layered notices can be particularly useful when describing the attribute values which will be released. In general, LDAP-style attributes are transferred to the **Service Provider Organisation** However, very few users have any familiarity with the conventions and usage of LDAP attributes. Instead, the Identity Provider could ask the user to release "name"; the link would take the user to a page listing all of the LDAP name attributes and values.

There are other attributes where the values are intentionally opaque (e.g. ePE="urn:mace:rediris.es:entitlement:wiki:tfemc2"). It is NOT reasonable to expect the end user to understand what this value means and to pick up a particular value to be released. Instead, natural language descriptions of the values should be provided.

A good way to explain to a user why there is a transfer of information is "your email, name and affiliation will be transferred".

## RECOMMENDATIONS

[10] *"A layered notice usually consists of a short notice plus a longer notice. The short notice contains basic information, such as the identity of the organisation and the way in which the personal information will be used... The short notice contains a link to a second, longer notice which provides much more detailed information."* (the UK information commissioner's Privacy Notices Code of Practice, page 18).

For all Attributes (INFORM interaction):

1. The user MUST be informed ofn the attribute release separately for each Service.

2. The user MUST be presented with the mdui:DisplayName value for the Service, if it is available.

3. The user MUST be presented with the mdui:Description value for the Service, if it is available.

4. The user SHOULD be presented with the mdui:Logo image for the Service, if it is available.

5. The user MUST be provided with access (e.g. a clickable link) to the document referenced by the mdui:PrivacyStatementURL.

6. The IdP MUST present a list of the RequestedAttributes defined as NECESSARY. No user consent is expected before release. (However, given how web browsers work, the user may have to click a CONTINUE button in order to continue in the sequence.)

   The IdP MAY list the NECESSARY attributes on the same screen as the username/password entry boxes, making clear that *if* you login then this is what will happen. It MUST be clear to the user that the consequence of their next action will be to release the attributes. NOTE -- the attribute values for the specific user are not available when the login screen is presented, since the user's identity is not yet known.

7. The display software SHOULD provide the ability to configure and display localised descriptions of the attributes (e.g. what PersistentID means) and their values (e.g. what eduPersonEntitlement="urn:mace:rediris.es:entitlement:wiki:tfemc2" means)

8. The display software MAY inform the user of the release of an "attribute group" (eg attributes expressing the user's "name"), and then release all requested attributes in the group (e.g. various forms of the user's name such as cn, sn, givenName and displayName).

9. The display software MAY give the user the option to remember that they have been INFORMed of the release of the necessary attributes.

10. If any of the following has changed since the user accessed this SPO for the last time, the user MUST be prompted again for the INFORM interaction

    a. the list of attributes the SPO requests
    b. the DisplayName of the SPO
    c. the Description of the SPO

## INTERNATIONALISZATION

The *lang* attribute of the *mdui* elements can be used to match the user's preferred language settings.

## SAMPLE NOTIFICATION

Example of how a Home Organisation should inform End Users and provide an opt-out opportunity before Attributes are released to a new Service Provider Organisation. Clicking the Service Provider Organisation's name leads to its Privacy policy page.


## APPENDIX C: SUPPORTING DOCUMENTS

The eduGAIN project has developed extensive additional material[11] to support Service Provider Organisations, Home Organisations and Identity Federations in adopting the Code of Conduct.

The normative documents consist of the Code of Conduct (Appendix A) and technical specifications on how to use the SAML 2.0 standard to convey information on Service Provider Organisations' commitment to the Code of Conduct to the Home Organisations in the most reliable way. The technical specifications also describe how the Service presents to the Identity Provider server the link to its privacy notice page and the list of Attributes it requests from the Home Organisation. This information is sufficient for the Home Organisation to configure its Identity Provider server to inform the End User of the Attribute release, and to release a limited Attribute set to Services without administrational involvement, for each Service.

In addition to the Code of Conduct, the normative documents are complemented with guidelines and good practice:

Guidelines for Identity Federations. The guideline describes how the Identity Federations are expected to record a Service Provider Organisation's commitment to the Code of Conduct and to convey it to the Home Organisations.

A Code of Conduct cookbook with recipes for Service Provider Organisations, Home Organisations and federation operators. The recipes provide step-by-step guidance on how to deploy the Code of Conduct.

## APPENDIX D: SUPPLEMENTARY TOOLS

[11] https://wiki.refeds.org/display/CODE

As the operator of the eduGAIN interfederation service, the GÉANT project runs a monitoring service[12] of all Services that have expressed commitment to the Code of Conduct. The monitoring tool performs some technical basic checks to the Services to make sure they have published a privacy notice document and the list of necessary Attributes that Home Organisations need to release. The tool provides a page where Service administrators can check their technical conformance (see Figure 7).

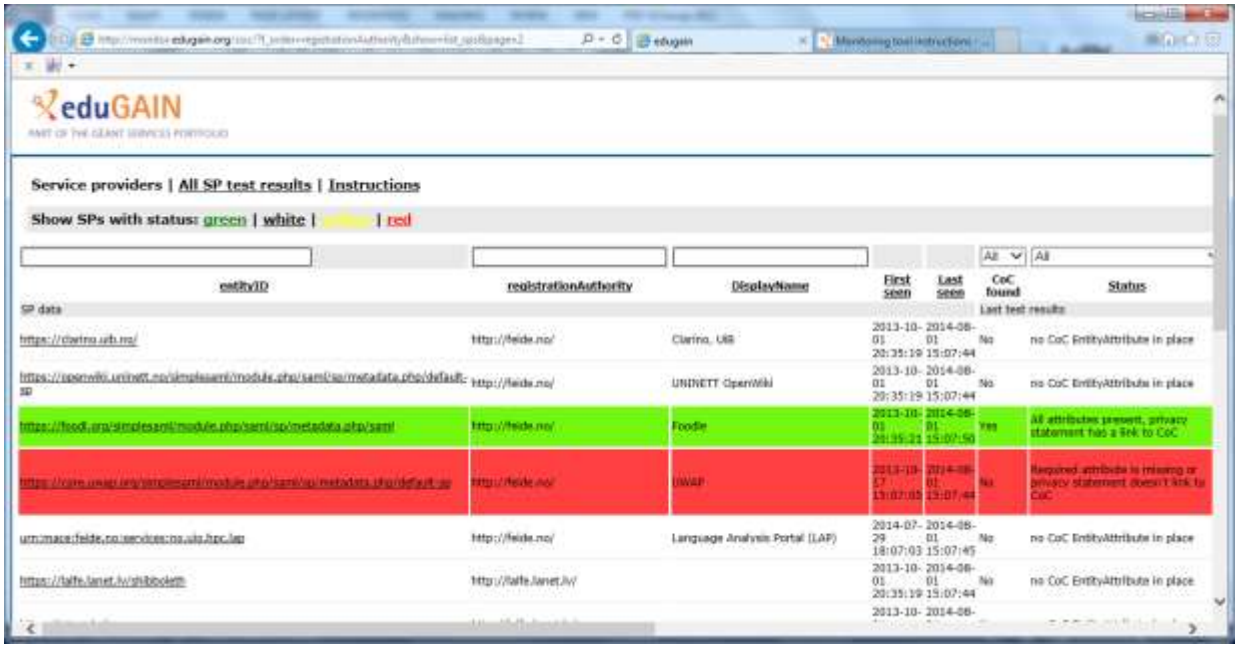



Figure 7. A Code of Conduct monitoring tool helps Service Provider administrators to ensure their technical conformance. Green line indicates the SPO is OK, red indicates a problem. White SPOs have not committed to the Code of Conduct.

The monitoring tool also periodically archives the monitoring results and the Services' Privacy notice documents for audit trail, in case of a later dispute.

[12] http://monitor.edugain.org/coco