Organising federated identity in Finnish higher education

Mikael Linden

CSC the Finnish IT Center for Science, P.O. box 405, FI-02101 Espoo, Finland
e-mail: mikael.linden@csc.fi

Abstract: Finnish higher education has been an early adopter of federated identity in Europe. The Finnish Haka federation is deploying Shibboleth, federating software by Internet2. This paper describes the federation as an organisational entity and explains how privacy issues are taken into account in its policy. Differences between the Haka federation and some other federations are pointed out. The main service areas for federated identity in Finnish higher education are also presented.

Key words: identity management, federated identity, privacy

1. INTRODUCTION

User administration means keeping track of an information system's users and their privileges. In an information system, a user identity is an abstraction of a person in the real world, and it is a collection of attributes describing her. Issues like management of user identities, authenticating users and authorising them to use services are all parts of user administration.

Traditionally, the maximum scope of a user identity has been only one organisation. The identity has not been shared with other organisations. If the user has used services outside her home organisation (for example, her employer or school), she has had separate usernames and passwords for each service. However, as the networking of organisations has become more common, it has become a subject of interest to share (*i.e.*, federate) user identities between organisations. In a federation, an end user only has the credentials (*e.g.*, username and password) her home organisation has given to her, and there is a specific middleware service that federates her attributes from the home organisation (called Identity Provider) to the service she is using (called Service Provider).

A federation is an association of organisations that come together to exchange information, as appropriate, about their users and resources in order to enable collaborations and transactions [1]. The federation, consisting of Identity Providers and Service Providers, has agreed on policies and practices necessary for carrying out the task. Some of these are of a mostly technical nature (such as the protocols used for communication and schemas for syntax and semantics of attribute exchange), some of them are more political (how to make the involved organisations trust each other) and some are legal (how the privacy of the end user is ensured as her personal data is disseminated between the organisations).

Shibboleth is a SAML-based middleware protocol specified by Internet2. Since the open source implementation became available in 2003, it has been deployed by higher education in several countries. In the United States, there are federations already using Shibboleth such as InCommon [1] and InQueue [2] and Australian higher education has shown interest in it [3]. In Europe, Swiss higher education has been the forerunner for Shibboleth. In the United Kingdom, projects funded by JISC are aiming at the deployment of a federation running Shibboleth. In addition, higher education in some other European countries has interest in Shibboleth.

In Finnish higher education, the development of the Haka federation has its origins in the year 2000, when the FEIDHE project focused on personal certificates on smart cards as a tool for strong authentication of end users. However, smart cards did not break through, and as an effect, the project recommended that its followers focus on organisational user administration rather than strong authentication [4]. Having identified the problems of LDAP in cross-organisational user administration, federating software was considered to be an interesting choice [5].

Several national research networks have developed architectures of their own, such as PAPI (Spain), Athens (UK) and FEIDE (Norway). In Finland, we had no resources to implement a protocol of our own. As designing and implementing a security protocol is difficult, we preferred adaptation of existing federating software. The Shibboleth architecture was sound and had the resources of Internet2 behind it. Therefore, it was easy to follow the direction chosen by SWITCH [6] and adopt Shibboleth as the federating software of Finnish higher education. The first Shibboleth pilots started in Spring 2003, and the pilot federation became operational in December 2003.

In February 2004, the Haka project ended and the proposed deployment of the Haka federation, which runs Shibboleth as the federating software [7]. CSC, the Finnish IT Center for Science, started to prepare the federation as a common infrastructure for universities and polytechnics in Finland. The production-level federation was formed in May 2005.

This paper focuses on the organisational elements of the Haka federation. The paper starts with the most important use-scenarios identified for federated identity. Chapter 3 discusses the organisational models for federations and presents the motivation for the choice made by Haka. Chapter 4 presents relevant parts of European data protection legislation from the federated identity point-of-view and how these regulations are taken into account in Haka. Chapter 5 discusses the quality of institutional identity-management and Chapter 6 concludes the paper.

2. USE SCENARIOS FOR FEDERATED IDENTITY IN FINNISH HIGHER EDUCATION

Different kinds of services can be identified as potential users of federated identity. To motivate the rest of this paper, this chapter introduces the four main service categories for federated identity in Finnish higher education.

2.1. Library services

Nowadays, researchers in institutions of higher education do not have to go to the premises of a university library to read scientific journals. Instead, the researchers use electronic services, such as electronic journals and databases, provided on the web by the journal publishers. University libraries pay licence fees to the publishers for making the journals available to the students and researchers in the institution. Typically, libraries intend to licence the journals for students, faculty and supporting staff in the institution and for other regular and registered users on-site [8]. At present, the access control of the journals is usually implemented by configuring the IP address space of the campus in the publisher's service.

IP address-based access control has known problems. It does not actually authenticate the end user; instead, the authorisation to use the service is based on the place where she is using the service. Legitimate users are not allowed to use the service outside the campus IP address space (*e.g.*, at home)[1]. On the other hand, illegitimate users, such as roaming users[2] or other users not considered as students or faculty members at the institution do have access, although, according to the licence terms, the material is not necessarily licensed for them. Furthermore, the authorisation is very coarse and there is no easy way to implement fine-grained access control. For example, the libraries might want to licence some more expensive material only to faculties in a certain department or to the participants of a certain course in the university.

For publishers, authorisation is not the only use for an identity federation. The publishers may like to develop their service further by providing end users with customisation. For example, computer science researchers would, perhaps, always like to see a list of the latest publications in the well-known LNCS publication series of Springer as they browse to SpringerLink. Thus, the publisher needs to get some persistent identifier of the user to which the user profile can be attached in the service[3]. In order to achieve this in its ScienceDirect portal, Elsevier Inc. has already joined the InCommon Federation.

In Finland, the libraries in higher education traditionally co-operate widely in licensing electronic journals. The Finnish Electronic Library consortium is the centralised organisation negotiating the licence agreements with publishers. Furthermore, the consortium has recently deployed a portal (Metalib, a product of Ex Libris Ltd.) that constitutes a common interface to the dozens of publishers with which the libraries have licence agreements. The portal uses services of the Haka federation to authenticate the user and provide her with customised services.

Furthermore, the Finnish libraries also have a common Library Management System (Voyager, a product of Endeavor Inc.), which, for instance, keeps track of library patrons' loans in a library. The web interface (WebVoyage), used by patrons for reviewing and renewing their loans, presently uses library card numbers for user identification. In an ongoing pilot project in Finland, Shibboleth is being integrated into WebVoyage to replace its current user identification system.

2.2. eLearning services

Utilising ICT for learning enhancement has been a subject, not only from the technical, but also from the pedagogical point of view. Several tools have been used, including video-conferencing, multimedia *etc*. The web has also become a commonly used environment for eLearning, and various web-based services have been developed, from simple web-based tools to fully-fledged learning management systems. Many of the eLearning services are interested in the identity and role of the end user.

There is a large number of commercial and open source learning management systems. In Finnish universities, the most widely used ones are WebCT, BlackBoard, Optima and R5 Vision [11]. The maintenance of learning management systems is not so well organised as the use of library services. In some institutions, the laboratories have their own installations of their learning management systems; in other institutions, there are some centrally-operated learning management systems that belong to the institutional IT infrastructure maintained by the university. Some initial discussion has been had about a national service centre for the

---

[1] VPN connections or dedicated proxy servers (such as EZproxy, http://www.usefulutilities.com/) are commonly used to circumvent the limitations of IP address based access control.

[2] VPN based roaming model is the only one giving a roaming user an IP ad-dress from her home institution [9].

[3] The eduPersonPrincipalName or eduPersonTargetedID attributes of the widely used eduPerson schema [10] can be used, for instance.

main-tenance of learning management systems for better efficiency. However, many lecturers consider the tools they use in teaching as part of their academic freedom.

Considering the aforementioned, it is not a surprise that the user administration of learning management systems is versatile. In some learning management systems, the students register to the system by themselves and get yet another username/password pair to remember. If the institution has a centrally operated learning management system, it is more likely to be coupled to the enterprise directory of the institution's IT department, allowing end users to use the same username/password pair they also use in other IT systems.

In Finnish higher education, it is possible to take courses from a neighbouring institution. Nowadays, the visiting students get local user accounts in the institution they are visiting, making cross-institutional user administration unnecessary. In other words, user administration of learning management systems is typically an institutional, not an inter-institutional issue, and there is little use in joining institutional learning management systems to the national Haka federation. Instead, in order to serve the user administration of learning management systems, IT departments are preparing to set up institutional light-weight federations, serving mostly laboratories inside the institution. These institutional federations may also use the Shibboleth technology, as it is easier to maintain only one middleware infrastructure[4]. In an institutional federation, the bureaucracy is easier because, for data protection, personal data is not disseminated between two organisations.

Having a national federation in place opens new business models for eLearning. The eLearning service need not be installed and maintained in the institution, and yet, it can utilise the user administration of the institution's IT department. In order to achieve economics of scale, there can be separate service centres that maintain learning management systems for several institutions. Furthermore, an institution can licence some specialised eLearning material for a small group of students; for example, to the participants of one individual course. For authorisation purposes, the participation of a student on a course can be expressed as a separate attribute that the institution's IT department provides to the eLearning service[5].

In Finland, several learning management systems have been, or are being integrated to Shibboleth, including WebCT (University of Helsinki), A&O (Tampere University of Technology), Moodle (University of Kuopio) and Optima (University of Oulu). First experiments are about to start on passing the students' course enrolment as an attribute to the learning management systems using Shibboleth.

---

[4] Shibboleth Identity Provider, version 1.3 is going to have multi-federation support in it.

[5] The CourseID working group
(http://middleware.internet2.edu/courseID/) of Internet2/MACE has specified how a person's role can be expressed as an attribute with respect to a given course offering.

2.3. National services for end users in institutions of higher education

In addition to eLearning and libraries, nationally centralised services are potential users of federated identity. Nationwide services are typically provided to a subset of end users that spans a large number of higher education institutions. Theend users can be, for example, students or researchers in any of the higher education institutions. As there has not been a national authentication and authorisation infrastructure in place, the services have either issued local usernames/passwords for end users or have not provided personal services to end users at all.

The Academy of Finland is a public body providing funding for research projects in universities. The funding application form has been made available electronically, and the Academy has issued usernames and passwords to the researchers for filling the applications. As the applicant has filed the application, it is circulated to specialists in other universities in order to get expert opinions on it. The use of the Haka federation instead of local usernames makes the application submission and circulation process easier for end users as well as for the Academy.

YTHS (Finnish Student Health Service) is a foundation serving all the masters degree students in Finnish universities. Presently, YTHS has no personal services on the web, as there has been no means to authenticate the 140 000 customers. YTHS would be interested in providing some basic services on the Internet. These services could include, for example, appointment reservation for the first-year-students' health examination or other functions that require no medical expertise.

2.4. Application service providers

Outsourcing applications is becoming common also in higher education. The universities in Finland are government agencies and are involved as the state corporation makes outsourcing decisions to reduce costs and increase administration efficiency. The first group-level outsourcing contracts made by the State Treasury cover electronic circulation of invoices and travel-expense administration. The application services are provided by large Finnish IT companies.

The Finnish state administration has 120 000 officers, 30 000 of which are staff and faculty in universities. Not all of these officers are involved in the circulation of invoices, but typically, most of them have travel expenses. Presently, user administration in the outsourced services is done manually, while some more advanced organisations have scripting in place to synchronise user databases with the enterprise directories. User authentication and authorisation in outsourced services is clearly a customer for identity federation, although it couples the identity federation of higher education to user administration issues in other Finnish government agencies.

3. THE ORGANISATION OF A FEDERATION

As defined in the first chapter, a federation is a set of organisations who have decided to co-operate in order to authenticate and authorise end users across organisational borders. In order to put the co-operation into practice, the organisations pick up and deploy some middleware technology, such as Shibboleth. In other words, a federation is an organisational, not a technical entity[6]. This chapter discusses how to organise a federation.

As the authentication and authorisation of users is an essential part of computer security[7] and the processing of personal data is regulated in the European Union, it is necessary to have written agreements between the federation participants defining related obligations and responsibilities. Furthermore, as the federation collects fees from the federation members to cover its costs, it also exists as an economic entity. There must be some kind of an organisation that signs the necessary agreements and deals with the accounting for incomes and expenses in the federation.

This paper identifies two ways to organise a federation. The InCommon and SWITCHaai federations have been organised as a service provided by a central organisation, such as InCommon LLC or SWITCH. The alternative would be to organise a federation as a consortium.

3.1. A federation as a service provided by an organisation

Having the federation organised as a service means that an organisation joining the federation signs a bilateral agreement with the operator of the federation (Fig. 1). In a way, the operator becomes a centre of a star, having bilateral agreements with all the organisations in the federation. In Switzerland, the operator is SWITCH, the maintainer of the national research and education network, a foundation of the governing bodies of
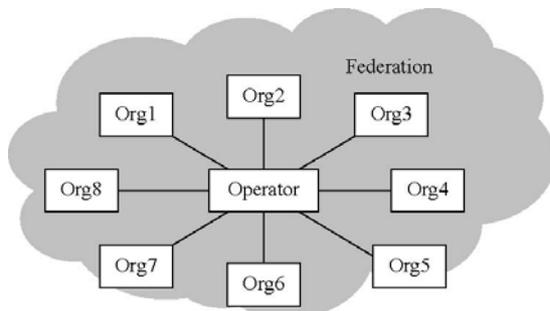
Swiss universities. In the US, InCommon is a limited liability company dedicated for the provision of federation services.

A benefit for organising the federation as a service is that no new organisation needs to be established for the federation. The joining organisations and the operator sign an agreement specifying the responsibilities of the two parties, and the federation is a collection of bilateral agreements between the operator and the participants. From a participant's point of view, all the other participants of the federation are subcontractors for the operator of the federation. If the participants of the federation have, for example, claims for each other, they have to discuss these with each other *via* the federation operator.

The downside is that organising the federation on top of bilateral agreements is not strictly consistent with the definition of a federation, which considers a federation as a set of organisations. As the centre of the star of agreements, the role of the operator becomes essential and demanding, for example, replacing the federation operator means, in practice, tearing down the federation and building a new one.

The business of the operator is inevitably to develop the federation service to make it more and more attractive and satisfying for the customers. The operator needs to deeply understand the requirements and, on the other hand, the limitations the federation participants. In higher education, the needs are typically driven by service providers like libraries, eLearning, etc. The limitations are set by the IT departments of the institutions and typically consist of issues like the quality of the institutional identity management systems or problems in linking organisational person registries to each other.

3.2. A federation as a consortium

Alternatively, a federation can be organised as a consortium (Fig. 2) that is, by definition, an agreement, combination or group (as of companies) formed to undertake an enterprise beyond the resources of any one member [13]. In that sense, a consortium is quite close to what we are looking for. In a consortium, organisations sign a multilateral agreement to become members. Having signed the consortium

Fig. 1. Federation as a service provided by the federation operator

---

[6] To distinguish the organisational and technical parts of federated identity, SWITCH has called the technical aspect (servers, configurations, *etc*.) Authentication and Authorisation Infrastructure (AAI).

[7] According to the definition by Gollmann [12], computer security deals with the prevention and detection of unauthorised actions by the users of a computer system.
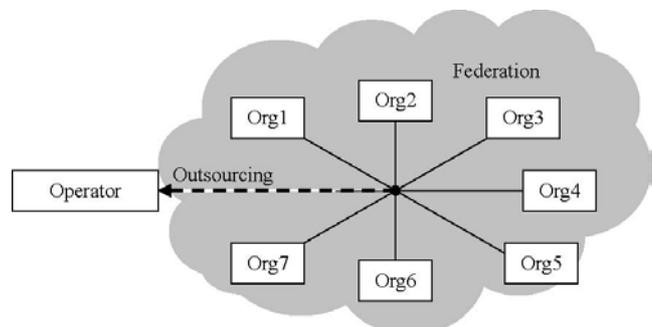

Fig. 2. A federation as a consortium

agreement, the organisations have direct contractual relationships and can make claims directly on each other.

The consortium would need a service centre that coordinates the federation. In higher education, it would probably make sense to place the service centre in some existing institution for higher education, for example, in its IT department. The secretaries of the consortium would be employed by the institution in question. At minimum, the consortium could be just an outsourcing organisation, buying the technical operations of the federation from commercial organisations.

3.3. The organisation of the Haka federation

In Finnish higher education, the two alternative ways to organise the federation were considered. Following the way SWITCHaai had chosen, the higher education institutions preferred organising the federation as another service provided by CSC, the maintainer of the national research and education network Funet. The most significant reason for that was minimising additional bureaucracy; there was no interest in forming yet another body for taking care of the common IT infrastructure in Finnish higher education. This was also the reason for not choosing to found a separate limited liability company, like InCommon in the United States.

CSC had been an active participant in developing the federation. As Funet was already a service provided by CSC, having the Haka federation as another service provided by CSC is not surprising. CSC, in turn, has the option to outsource some parts of the federation operations. For example, at the moment, CSC has no 24 hour support for servers such as WAYF (Where-Are-You-From, a Shibboleth server used by the end user for picking up her Identity Provider), which may become necessary as the use of the federation is increased.

Choosing the consortium would have meant that the institutions would have established the consortium and placed its administration in some existing IT department in a university. Most probably, the administration would have consisted of only one part or full time employee, who takes care of the consortium's administration and financing, of taking new members to the consortium and of outsourcing contracts for all technical issues in the federation. These would include issues like the maintenance of federation metadata and WAYF server, organising a helpdesk and courses for people in higher education institutions and so on. As there is little commercial supply for federated identity at the moment, the subcontractor would probably have been CSC, at least in the beginning.

The organisation of the federation is depicted in Fig. 3. As the federation is organised as a service operated by CSC, which is not a higher education institution itself, it becomes vital to set up mechanisms that make sure the operator has contacts to the daily life of federation users in institutions of higher education. To ensure that the requirements and limita-
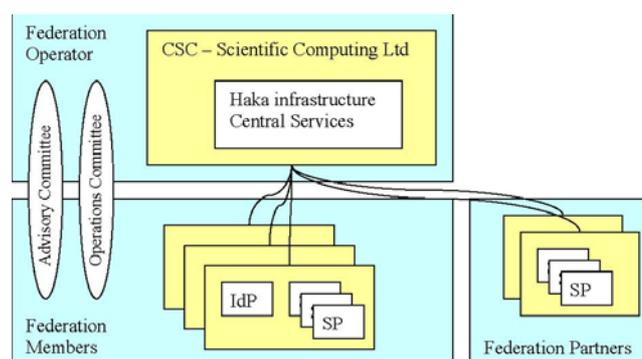

Fig. 3. Organisation of the Haka federation is similar to SWITCHaai

tions related to the federation are communicated to CSC, the federation has an Advisory Committee. The committee consists of representatives for the institutions' IT departments (4 persons), eLearning consortia (Finnish Virtual University and Virtual Polytechnic, 2 persons) and a library consortium (Finnish Electronic Library, 1 person) of Finnish higher education. CSC also has a representative on the committee. Participating in related events in higher education (such as gatherings of IT department employees, eLearning people, *etc*.) and personal contacts help the operator adjust to the needs of the customers.

3.4. The service agreement of the Haka federation

The Haka federation is a service provided by CSC as defined in the service agreement of the federation [14]. CSC, the operator of the federation, defines the terms of the service in the agreement's appendices. These can change over time. The Advisory Committee of the federation acts in an advisory capacity and represents the members of the federation. The meetings of the Advisory Committee are prepared and convened by the operator. In the service agreement, the Advisory Committee is defined as the authoritative body for a set of issues, such as accepting federation partners or members other than institutions of higher education. However, the committee's main role is advisory only, and the operator makes final decisions on the terms of service. If federation partners are not satisfied with the service, they always have the ultimate right to terminate the service agreement, or threaten to do so.

Like in SWITCHaai, the Haka federation has two categories for federation's participants; federation members and partners. Higher education and research institutions may join the federation as members and become both Identity Providers and Service Providers. Federation partners, such as library content providers, may only become Service Providers. As the service agreement of the Haka federation is signed between the federation operator and the participant, from the federation participants' point-of-view, the federation is a service provided by the operator and the other participants in the federation are subcontractors for the operator. In

the agreement, it is made explicit that the contents of the service agreements are equal for each federation member.

The section defining indemnification is modest. Neither party is liable for damages caused due to bad quality of the service such as its downtime or weak performance. Federation participants refrain from claims on each other. Other sanctions defined in the agreement were considered sufficient for all parties. If the operator has quality problems, the federation participants do not have to pay fees for the time - period in question. If a participant has a problem, the operator is allowed to stop providing the service to it. The ultimate consequence is the termination of the agreement.

It is clear that the service terms, including indemnification, are not very strict in the Haka federation. Having CSC, a non-profit company owned by the Ministry of Education, as the federation operator is far different from a commercial company. The operations of the Haka federation are based not only on the service agreement, but also on the trust higher education institutions have in CSC, which has been their partner for decades. A service agreement with a commercial company would be much stricter, as the nature of a commercial company is to try to minimise the costs and maximise the income from their services.

4. PRIVACY ISSUES IN A FEDERATION

As a member of the European Union, Finland has implemented the EU Data Protection Directive in the national legislation. The Finnish Personal Data Act restricts the way personal data may be processed by the Identity and Service Providers of a federation. This chapter points out the parts of the directive that affect especially on attribute release in a federation. The chapter also presents related means that have been implemented in the Haka federation policy.

The privacy related mechanisms in the Haka federation differ from SWITCHaai. In Switzerland, there is also cantonal privacy legislation in which not all details are similar. As Finland has consistent data protection legislation, the federation preferred to also cover detailed mechanisms for privacy in its procedures; centralising certain privacy-related check-ups in the federation reduces overlapping of work (which the technical staff usually considers boring). The other alternative would have been to leave the privacy issues uncovered and up to each federation participant to take care of.

Liberty Alliance has made an extensive study of European legislation and its effect on federated identity [15]. Although the study focuses on the Circles of Trust, *i.e.*, federations utilising Liberty technologies, the issues are, for the most part, applicable for Shibboleth-based federations as well.

Article 2 of the data protection directive defines personal data as information that relates to an identified or identifiable natural person. Processing of personal data is defined as any operation or set of operations which is performed upon personal data, such as collecting, storing, disseminating and so on. It is clear that user accounts in an Identity Provider are personal data, and, therefore, the Identity Provider processes personal data. The Service Provider processes personal data only if the attributes provided by the Identity Provider and other records collected by the Service Provider relate to an identified or identifiable individual[8]. As the attribute release takes place directly between the Identity and Service Provider, the operator, in turn, never processes end users' personal data in a federation running Shibboleth as the federating software.

4.1. The purpose of processing personal data

Dependency on the purpose of processing personal data is fundamental to privacy laws in Europe. According to the Data Protection Directive, *(Article 6) Member states shall provide that personal data must be (b) collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.* Liberty Alliance has not covered this aspect in its document [15].

In Finland, universities and polytechnics are public organisations as defined in the Universities Act and Polytechnics Act. The mission of universities and polytechnics is also specified in the acts; in short, it is research and education, with the polytechnics emphasising more applied aspects. Identity management is a supportive function in higher education institutions. Thus, according to the Universities and Polytechnics act, the purpose of processing personal data in institutional identity management systems is supporting research and education. Personal data may not be processed (for example, disseminated) in institutional identity management systems for purposes incompatible with that.

The Haka federation has addressed the purpose of processing personal data in its policy. The purpose of the federation is simply "to support higher education and research institutions". Only organisations having services compatible with this purpose are accepted to the federation. For institutions of higher education that act as Identity Providers or Service Providers this is not a problem. For organisations providing services to higher education, such as library content providers, this is not a problem either. On the other hand, services like Internet gambling that are clearly not supporting research and education and may not join the federation. Some organisations are partly compatible with the purpose; for example, the services related to applying for student loans at KELA (the Social Insurance Institution of Finland) can join the federation, but the services related to maternity allowance cannot. In borderline cases, it is up to the Ministry of Education to draw the line.

---

[8] The United Kingdom Information Commissioner emphasises identifiability as a contextual issue [16]. In the physical world, individuals are distinguished from others typically by names and addresses; in the on-line world, for example, by tracking cookies and pseudonyms.

Dependence on the purpose of the personal data processing makes European privacy legislation different, for example, from the legislation in the United States. In the United States, higher education is co-operating with the e-Authentication project of the Federal government in order to enable end users in higher education to use their credentials for authenticating to government services as well. According to the Data Protection Directive, this appears not to be possible in Europe. Government services, such as social security, taxation, etc, are not supporting research and education. This incompatibility can be seen as an obstacle when bridging the United States and European federations together in the future.

4.2. The relevance of attributes

According to the Data Protection Directive *(Article 6) Member states shall provide that personal data must be (c) adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed.*

In an identity federation, Identity Providers are not allowed to release and Service Providers are not allowed to collect attributes that are irrelevant for the service in question. The relevance of attributes depends on the service; for a student loan service, the Social Security Number is probably a relevant attribute, as the SSN is used for identifying individuals in government services. For a learning management system, the SSN is probably irrelevant.

From the data protection perspective, the optimum is that no personal data is processed at all. In higher education, there are several services (such as the article databases licensed by libraries or WLAN roaming access) which are typically not interested in the end user's identity but on her authorisation to the service. The authorisation may be derived from the end user's attributes (for example, faculty members are authorised to use the library database or WLAN network). If an individual cannot be identified, the Personal Data Act is not applied at all to the attribute release.

The Haka federation's policy documents define responsibilities for ensuring that only the relevant attributes are released to the service. The administrative contact of the federation participant signs a request and sends it to CSC before CSC adds the new service to the federation metadata. It is a responsibility of the federation participant's administrative contact to make sure that all the attributes in a service are relevant. In a higher education institution, the administrative contact is typically the information manager of the institution. He or she knows the local circumstances and is, unlike CSC, competent to deduce the relevance of attributes for the service in question.

4.3. Informed consent

According to the directive, an individual's consent is the basis for processing personal data. For Identity Providers,

among other things, release of attributes is considered as processing of personal data. For Service Providers, collecting attributes that identify an individual is processing personal data, no matter if the attributes are provided by an Identity Provider or by the end user herself. *(Article 7) Member States shall provide that personal data may be processed only if:*
(a) the data subject has unambiguously given his consent;
(b) processing is necessary for 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 data subject which require protection under Article 1(1).

Some activities in an identity federation could probably fall in a category other than (a). However, the Finnish Data Protection Ombudsman gave advice that the consent of an individual should always be considered as the primary way for making the release of personal data legitimate. Furthermore, according to Article 11, the subject of the data must, in any case, be informed about to whom and for what purposes his/her personal data is going to be released. This can be done conveniently when asking for his/her consent.

It is worth noting that the user's consent overrides neither the requirement for the compatibility of the purpose of processing personal data nor the requirement for the relevance of attributes released. Only relevant attributes may be released and only to services supporting higher education even if the end user has given her consent for the release of attributes.

The policy documents of the Haka federation mandate that the Identity Providers always ask the user when her personal data is released to a new Service Provider for the first time. The consent is asked after the Identity Provider authenticates the end user but before the end user's web browser is redirected back to the Service Provider. If the user denies the release of attributes, the Shibboleth message exchange does not continue.

A Privacy Policy is a document that the Service Provider maintains and that contains the information required by the Article 11. The federation operator gathers and distributes the Privacy Policies' links as a part of the federation metadata. To make the end user's consent an informed one, the Identity Provider is responsible for providing an end user with the link to the Privacy Policy of the Service Provider.

Thereby, the end user is able to read the Privacy Policy before he consents to the release of attributes.

4.4. How Shibboleth fulfils the privacy requirements

Shibboleth provides excellent tools for covering the three issues presented above. The Attribute Release Policies (ARP) provide the means for controlling to which services the attributes are released. In the Shibboleth implementation, there are two kinds of ARPs. Site ARPs are maintained by the Identity Provider and they permit or deny attribute release for any end user. Additionally, each end user may have her personal User ARP. The two ARPs are conjunctive; both the Site and the User ARP (if existing) have to permit attribute release to a certain Service Provider to make the attribute release take place.

Compatibility with the purpose of processing personal data (Chapter 4.1) can be ensured by making sure that the Site ARP does not permit the release of any personal data to a Service Provider incompatible with the purpose of the federation ("to support higher education and research institutions"). The site ARP can also be used to make sure that only relevant attributes are released to a given Service Provider (Chapter 4.2). In the Haka federation, Site ARPs are maintained by the federation operator and distributed to Identity Providers as part of the federation metadata.

The end user's consent (Chapter 4.3) is stored as a User ARP. When the user accesses a service for the first time, the Identity Provider asks her permission for attribute release and writes a relevant entry to her User ARP file. Having given her consent once, the user is not interrupted by the dialogue again when she uses the service the next time. However, the end user can be provided a separate tool for viewing and modifying the ARPs she has in force at any time.

As presented in Chapter 4.3, user consent does not override the requirement for compatibility and relevance of processing personal data. In Shibboleth, this is ensured by requiring that both Site and User ARP must permit the attribute release.

5. THE QUALITY OF INSTITUTIONAL IDENTITY MANAGEMENT

It has become evident that many institutions of higher education have problems with the quality of data in their institutional enterprise directories. User accounts are not systematically closed as students graduate. The links between the student registry, human resources registry and the enterprise directory are missing. The institutions of higher education that have gone through the project of improving the situation have found that it takes several years to fix an institutional user administration. In addition, the project is not only about technology but also about streamlining workflows in the organisation.

Previously, the quality of institutional identity management was an internal issue for each institution. However, in an identity federation, the user attributes, whether of good or bad quality, are visible not only to the Identity Provider itself but also to the Service Providers in the federation. From the Service Provider point-of-view, having Identity Providers with varying qualities of institutional identity management is a problem. Service Providers are questioning what the benefit is of the identity federation if they are not able to trust on the users' attributes provided by the Identity Providers.

Like the FEIDE federation in Norway, the Haka federation has made it a mandatory requirement for an institution joining the federation as an Identity Provider that its enterprise directory has high-quality data in it. Changes in the base registries (student and HR registry) have to be reflected to the enterprise directory. Releasing only high-quality-data to Service Providers has been considered as a high priority issue in the federation. As an Identity Provider joins the federation, it makes a self-audit in its identity management under the supervision of the federation operator. As an output, a doument describing the principles of the institutional identity management is published in the web.

In order to support institutions of higher education in the development of their institutional user administration, CSC has run a series of workshops called "the school in user administration". In the workshops, best practices have been introduced and new products presented. During the workshops, the participants have been asked to make an assessment of the present user administration system in their home organisation and to set the goal for its development.

6. CONCLUSIONS

Although driven by development of protocols such as Shibboleth, federated identity is not only about developing technology. An identity federation must be given an organisational shape as well. The policy documents of a federation have to be in place, defining requirements and best practices for organisations in the federation. Federation policy has to take into consideration the relevant privacy legislation and integrate the obligations to the organisation and procedures in the federation.

This document presented how the Haka federation, the identity federation of Finnish higher education, had come to the decision to organise the federation as a service provided by CSC, the Finnish IT Center for Science. The service agreement and controls over privacy and attribute quality in the federation were also introduced.

References

[1] InCommon Federation. InCommon glossary http://ww.incommonfederation.org/glossary.cfm Referenced 2.5.2005.

[2] InQueue Federation. http://inqueue.internet2.edu/ Referenced 2.5.2005.

[3] International Middleware Event. Australia Position Paper. http://www.jisc.ac.uk/uploaded_documents/Australia_Positio nPaper.doc Referenced 2.5.2005.

[4] M. Linden, P. Linna, M. Kivilompolo, J. Kanner, *Lessons Learned in PKI implementation in Higher Education*. Proceedings of EUNIS2002, the 8th International Conference of European University Information Systems, Portugal, 246-251, 2002.

[5] M. Linden, *Towards Cross-Organisational User Administration*. Informatica 27, 353-359 (2003).

[6] SWITCH Authentication and Authorization infrastructure. Architecture Evaluation. January 2003. http://www.switch.ch/aai/pilotdocs/Arch_Eval_v10.pdf

[7] The Finnish Haka Project. Conclusions and Proposals of Action. February 2004. http://www.csc.fi/suomi/funet/middleware/english/HAKA final_report.pdf

[8] The Ligue des Bibliothèques Européennes de Recherche. LIBER Licensing Principles for Electronic Information. http://www.kb.dk/liber/news/981116.htm Referenced 2.5.2005.

[9] *Trans-European Research and Education Networking Association*. Terena Technical Report TF-Mobility. Inter-NREN roaming. Final Report. 2004. http://www.terena.nl/tech/task-forces/tf-mobility/Deliverables/TF-MobilityfinalReport.pdf

[10] Educause. "eduPerson Object Class". http://www.educause.edu/eduperson Referenced 2.5.2005.

[11] K. Koivu, Learning management systems in use in Finnish universities, June 2004. Available in Finnish: http://www.uta.fi/itpeda/Raportit/koko_kartoitus101103.pdf

[12] D. Gollman, *Computer Security*. John Wiley & Sons, Inc, New York, 1999.

[13] Merriam-Webster's Online Dictionary. http://www.britannica.com/dictionary Referenced 2.5.2005.

[14] Haka Federation. Service Agreement for Federation Members. http://www.csc.fi/suomi/funet/middleware/english/ Referenced 2.5.2005.

[15] Liberty Alliance Project. Circles of Trust: The Implications of EU Data Protection and Privacy Law for Establishing a Legal Framework for Identity Federation. February 2005. http://www.projectliberty.org/specs/Circles_of_Trust_Lega lFramework_White Paper_322200522576.pdf

[16] The UK Information Commissioner. Data Protection Act 1998. Legal Guidance. http://www.informationcommissioner.gov.uk/cms/Docume ntUploads/Data%20Protection%20Act%201998%20Legal %20Guidance.pdf Referenced 2.5.2005.

LIC TECH MIKAEL LINDEN is a post-graduate student in Tampere University of Technology, focusing his studies on identity management. At CSC the Finnish IT Center for Science, he is coordinating the deployment and operations of the Haka federation of Finnish higher education.