

E-Learning in Shibboleth-based federations: The design rationale behind the German DFN-AAI E-Learning Profile

Wolfgang Hommel¹

¹ Leibniz Supercomputing Centre, Boltzmannstr. 1, D-85748 Garching n. Munich, hommel@lrz.de.

Keywords

LMS federation, Shibboleth, federation schema, DFN-AAI, federated identity management

1. EXECUTIVE SUMMARY

Federated identity management enables single sign-on and user data exchange across organizational borders. Higher education institutions (HEIs) typically use the Shibboleth open source software and are organized as national federations, whose technical and organizational core components are operated by the national research and education networks (NRENs). In Germany, the authentication and authorization infrastructure DFN-AAI is such a federation. As in several other countries, this federation was initially motivated by university libraries that supported Shibboleth as a new, privacy-preserving access control paradigm in order to replace archaic IP-address-based authorization. Based on the de-facto standard data schema eduPerson, the DFN-AAI federation allows authentication and authorization for quite a broad range of applications. Most of the inter-organizational IT service usage can be enabled by simply managing the user's affiliation (e.g. student or staff at the HEI) and entitlements (such as the famous "common-lib-terms" for basic access to the library resources).

1.1. Paving the road for federated Learning Management Systems

Learning Management Systems (LMS), however, failed to use the federation because it lacked a standard way to transfer service-specific information about users, such as a student's field of study. Several regional extension attempts were made, but the data's syntax and semantics were not understood by all E-Learning service providers and home institutions in the federation in a uniform way. Thus, to facilitate technical interoperability for E-Learning across HEIs' borders in Germany, an additional data scheme for Shibboleth has been designed. The newly specified user attributes are based on surveys and hands-on meetings, and the result was published in late 2008. It is commonly referred to as the DFN-AAI E-Learning Profile.

1.2. Federated E-Learning in Germany

With this presentation, we want to motivate the federations in other countries to consider the special needs of E-Learning and start discussions about European interoperability. We first outline the real world demand for inter-organizational E-Learning in Germany by taking the Virtual University of Bavaria as an example; it works as a broker, i.e. for one university's students it suggests online courses at other universities that fit into the student's curriculum. After using a proprietary data exchange protocol for several years, the infrastructure is now being migrated into the DFN-AAI federation. We shed light on the design criteria and challenges, evolution, and current adoption status of the DFN-AAI E-Learning Profile. While the results that have been achieved by TERENA SCHAC greatly influenced our design, we discuss why we sometimes had to deviate and complement the existing de-facto standards and prior work. Although interoperability is a major milestone for the German federation, it obviously can only be an intermediate step towards borderless, cross-federation E-Learning. Thus, we also discuss some of the most important challenges that lie ahead from our perspective, and we are eager to learn other federations' requirements and plans.

2. MOTIVATION FOR FEDERATED E-LEARNING AND ITS BENEFITS

Many of our colleagues pursued their own studies back in a time when there was no online learning and when fully-featured web-based LMS infrastructures – like we design, implement, deploy, and operate them successfully today – did not yet exist. Thus, connecting modern E-Learning paradigms to the evenly young, but also very successful federation technology is a major step whose motivation is certainly not obvious to everyone. For this reason, we outline our motivation and the benefits of LMS federation in this section.

2.1. Motivation and obstacles for inter-organizational E-Learning

The preparation of high-quality E-Learning material often is a time consuming and expensive process. To share the fruits of their labor with a larger audience, many authors and instructors want to make their courses available to interested learners from outside their own HEI. However, giving away the material completely, e.g. by sending copies to other HEIs, is often considered a suboptimal solution: The original authors and their home institutions might not be put in perspective appropriately, the distribution of content updates may become tedious, and the direct feedback link from learners would be cut.

Similar arguments are relevant for commercial 3rd party learning material providers: Due to intellectual property rights and potentially lost revenues, the unregulated circulation of restricted material must be avoided by using tight access control and authorization policies. Thus, a more central approach to hosting the learning content is often preferred.

Furthermore, small HEIs often dread the costs to set up and operate a fully-featured LMS infrastructure, especially when the number of offered online courses is quite low. In such cases, co-operations between multiple HEIs – again with a centrally hosted LMS – provide more cost-effective solutions, which also imply that the LMS users are external to the organization which locally operates the LMS.

Last but not least, LMS usage across institutions' borders is very attractive from the learners' perspective. It allows to choose from a much larger set of interesting courses, and in many cases the credits earned from participating in such courses' exams can already be applied to the learner's home university studies.

In the past, user management has turned out to be one of the major obstacles for LMS usage across institution's borders. In most cases, interested learners had to use a self registration web form to initially sign up for an LMS account at the other institution, and it quickly became annoying to get another username/password for almost each external course. In the course of time, the LMS-internal user data quality often suffered from outdated contact information as well as fake registrations and identity theft. For access to restricted learning material, such as hospital patient photos that are restricted to medical personnel and students, confirmations of enrollment had to be mailed via postal service. In sum, the lack of automation caused a significant overhead for instructors as well as learners that clearly lowered the user acceptance.

Several attempts were made to automatically transfer, for example, lists of learners from home universities to external LMS providers. However, they resulted in proprietary protocols, and by far not every HEI and LMS software producer was convinced to implement them. Due to the lack of consolidation, the deployment scope of these various regional efforts was limited and no interoperability was given.

2.2. Using Federated Identity Management for E-Learning

Meanwhile, Federated Identity Management (FIM) offers a standards-based way to exchange user data between organizations. It is not tailored towards a specific service and has already successfully been used for many services in the HEI environment, including electronic library media access, licensed software distribution, and Grid computing. As shown in figure 1, in FIM basically each user is assigned to her home organization, which is called the Identity Provider (IDP) for this user. Whenever the user wants to login to a service that is provided by an external Service Provider (SP), the SP can request three types of data from the user's IDP:

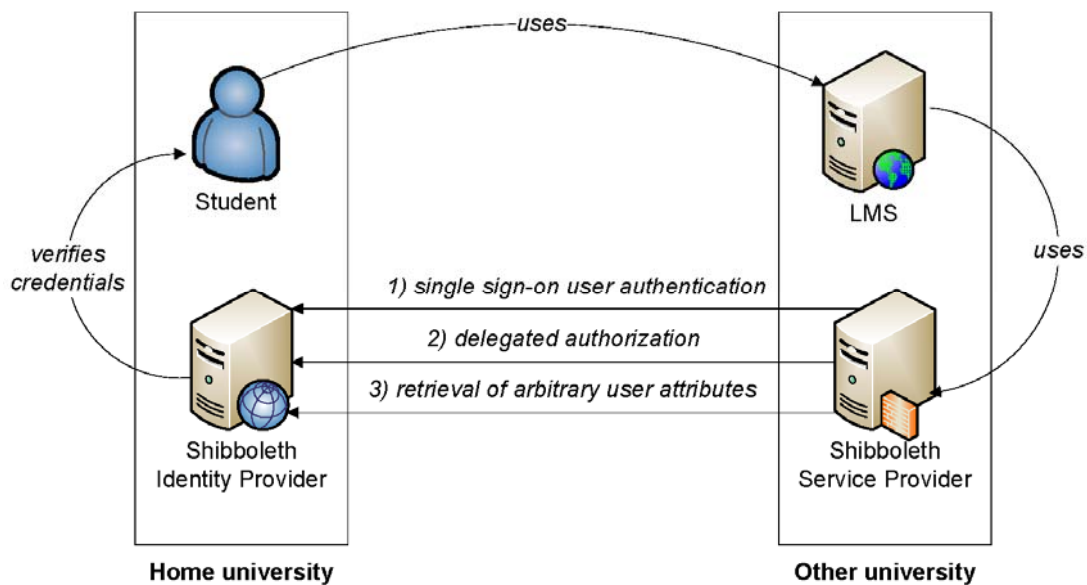


Figure 1: Roles and user data exchange using Federated Identity Management

1. Authentication assertions: Instead of a separate username and password for each service, any login is handled by the user's IDP. The SP trusts the IDP to properly and reliably handle the authentication process, and thus inter-organizational single sign-on can be achieved. This also eliminates the risk of password interception by compromised service providers.
2. Authorization data: The SP can optionally ask the IDP whether the user is authorized for the service from the home organization's perspective. For example, this can be used to restrict the user's participation to courses which fit her field of study according to the home university's regulations.
3. User attributes: Arbitrary additional data, such as the user's contact information, can be requested from the IDP. This obviously leads to data protection and privacy implications, which we discuss in section 4.

A federation is formed by an arbitrary number of IDPs and SPs that want to join forces. Besides the membership in the federation itself, service level agreements and other contractual agreements regulate whose users are authorized for which of the services that are available in the federation. Higher education federations have been established in most European countries as well as in USA and Australia; the constantly rising number of participating institutions shows that FIM is not just another short-lived hype.

The FIM user data exchange functionality solves the problems we discussed above: First, users do not need additional login credentials for each external service and can immediately start to use the services they are authorized for. Second, IDPs are reliable data sources for SPs, which reduces data quality problems such as outdated information and fake registrations. Finally, FIM is not a proprietary solution, but the IDP installations can be used for an arbitrary number of other services. Given these advantages, FIM is a prime candidate for inter-organizational E-Learning.

3. THE GERMAN DFN-AAI E-LEARNING PROFILE

The German HEI federation DFN-AAI is a service provided by the German NREN, Deutsches Forschungsnetz (DFN, see (DFN-Verein, 2009)). DFN operates the federation's technical core services, such as the federation metadata administration interface, and manages the various necessary federation policies and contracts, e.g. regarding the terms and conditions of IDPs' and SPs' federation membership.

DFN-AAI is in production since 2007 and was - similarly to many other federations - greatly influenced by the requirements of HEI libraries. Due to this heredity and as the term AAI (authentication and authorization infrastructure) implies, DFN-AAI primarily focuses on the data types (1) and (2) that we discussed above. The exchange of arbitrary user data - although technically possible - was barely used, since it was not required by the existing SPs, and minimizing the amount of data sent by the IDP to the SP is an essential privacy best practice.

3.1. The demand for a federation-wide LMS user data schema

In many of Germany's states, inter-institutional E-Learning infrastructures have been established. For example, the Virtual University of Bavaria (VHB, see (VHB, 2009)) is a network of more than 35 universities and universities of applied sciences that has been initiated to support and coordinate E-Learning efforts in Bavaria. Instead of operating a central LMS, VHB offers broker functionality: For one Bavarian university's students it suggests E-Learning classes at other Bavarian universities in accordance with the learners' study courses.

VHB had specified its own proprietary student record format and exchange protocol, both of which were implemented and deployed at about half of the participating HEIs. However, because several commercial as well as open source LMS products could not be adapted to this proprietary solution due to a lack of resources, VHB decided to evaluate standards-based approaches in 2006, and the upcoming DFN-AAI quickly turned out to be a very promising platform.

However, unlike the other services which DFN-AAI was designed for originally, LMS infrastructures offer a broad range of features, such as personalization and graded certificates of achievements that depend on quite detailed user records. As a consequence, the available basic authentication and authorization functionality as well as the eduPerson LDAP object class, which for example provides attributes for the user's name and email address, was insufficient to make full use of LMS across organizational boundaries.

When IDPs and SPs exchange user data in a federation, they obviously need a common understanding of the data's syntax and semantics, which is usually authoritatively specified in a federation schema document. While a VHB-specific extension to the official DFN-AAI federation schema could have been created and applied only by the Bavarian HEIs, there have been similar E-Learning federation movements in other German states, which were facing the same difficulties.

Thus, to avoid redundant work and to enable interoperability among IDPs and LMS SPs across state borders, DFN was approached with the request to extend the official DFN-AAI federation schema in order to include LMS-specific user attributes. However, because DFN wanted to avoid that the common federation schema gets cluttered up with E-Learning specific attributes, the idea of service-specific schema extensions - which were termed Profiles - was born, and the work on the DFN-AAI E-Learning Profile started.

3.2. User attributes specified in the DFN-AAI E-Learning Profile

A working group with representatives from most of the German states was established to elicit, discuss, and consolidate the schema requirements. For example, VHB conducted a thorough survey among its Bavarian member HEIs to determine and prioritize the list of user attributes that were already used in the local LMS deployments (independent of what the VHB's proprietary solution had provided in the past).

Afterwards, a core specification team was charged with drafting the E-Learning Profile specification, which was finally ratified and published in November 2008 (Deutschmann, Gietz, Hommel, Schroeder, Schwendel, & Thelen, 2008). It provides the following additional user attributes and is aligned to de-facto standard schemas like inetOrgPerson and to the results of TERENA TF-EMC2's work (SCHAC, (TERENA Task Force EMC2, 2009)):

1. Attributes for *LMS personalization* features:

- 1.1. *Gender*: For use in salutations ("Welcome back, Mr. John Doe"), the user's gender must be known. The *schacGender* attribute is used for this purpose.

- 1.2. *Personal title*: To properly address instructors and content authors, the *personalTitle* attribute as specified in RFC 4524 (e.g. "Prof. Dr.") has been included in the E-Learning Profile.
 - 1.3. *Preferred language*: For LMS deployments that support localization, which is especially important for HEIs that attract a lot of foreign students, the user's *preferredLanguage* can be specified as defined in the LDAP object class *inetOrgPerson* (RFC 2798).
2. Attributes for *authorization management* within an LMS:
- 2.1. *Study branch*: The user's study branch is essential for access control regarding restricted content. It is provided on up to three coarse- to fine-grained levels that are based on official numerical identifiers provided by the German Federal Statistical Office (FSO). First, the *subject group* allows distinguishing between, for example, natural sciences and medicine. Second, the *field of study* allows distinguishing between, for example, physics, biology, and computer science within the natural sciences subject group. Finally, there are fine-grained identifiers for *study courses*, such as geophysics. The official numerical values are readily available in the HEI administrations, since each HEI regularly has to create statistics and report them to its state's statistical office, which in turn submits them to the FSO. A new LDAP object class *dfnEduPerson* was created to contain these three study branch attributes, since neither *eduPerson* nor the *SCHAC* object classes offered an attribute that could have been used for this purpose.
 - 2.2. *Study course*: The learner's study course is not only provided using the FSO's numerical identifier, but also as a free text attribute that reflects the university's local nomenclature. This is useful for bilateral co-operations between HEIs, especially because it usually takes some time before an official identifier is created for a new study course. For example, no FSO identifier existed a priori when bio-informatics study courses were introduced. This attribute is also part of the *dfnEduPerson* LDAP object class.
 - 2.3. *Type of study*: The type of study specifies whether this is the learner's first course of studies, or, for example, her postgraduate course. It is required for LMS deployments that offer different courses for students of different types of study. Like the study branch attribute, it uses numerical identifiers provided by the FSO and is part of the *dfnEduPerson* object class.
 - 2.4. *Degree*: The degree attribute specifies whether the study course leads to, for example, a bachelor's or master's degree. It is used to determine whether the learner is allowed to take advanced online courses in the LMS, and is also based on official FSO identifiers.
 - 2.5. *Terms of study*: This attribute specifies the learner's semester for each subject. Similar to the distinction between bachelor and master students, access to certain material is restricted to upper semester students in several fields of study.
 - 2.6. *Cost center*: The *dfnCostCenter* attribute has been introduced to provide accounting and billing support for commercial online courses and on-the-job training.
3. Attributes for *issuing certificates of achievements*:
- 3.1. *Date and place of birth*: The *schacDateOfBirth* and *schacPlaceOfBirth* attributes are used to specify the user's date and place of birth. Both attributes are required by HEIs which issue classic paper-based certificates of achievements that can be submitted to the home university's examination office by the learner.
 - 3.2. *Matriculation number*: The student's matriculation number is used as an artificial unique identifier for the student record and typically printed on certificates of achievements along with the student's full name, study course, date of birth, and place of birth to ensure proper student record correlation in the examination office. The *schacPersonalUniqueCode* attribute has been used for this purpose.

The attributes used for the authorization of students (2.1 - 2.5 in the above list) can optionally be provided as a single concatenated attribute instead of multiple separate attributes; this reduces the implementation overhead for IDPs which do not already handle these user attributes and thus have to extend their student data import interface. While this provides additional flexibility for IDPs, it also requires that SPs are able to deal with two different syntactical ways how the authorization data can be transmitted. However, since LMS products need to be adapted to make full use of the DFN-AAI E-Learning Profile anyway, this burden was deliberately placed on the SPs in order to simplify matters for the IDPs; we will discuss the consequences in section 5. Furthermore, all the authorization attributes are indexed multi-valued attributes and thus can be used for students who are enrolled in multiple study courses at the same time. Interested readers are referred to the original E-Learning Profile specification document for details and examples.

Obviously, almost each of the attributes specified above contains personally identifiable information (PII); thus, we will discuss the security, data protection, and privacy challenges related to the DFN-AAI E-Learning Profile next.

4. SECURITY AND PRIVACY CHALLENGES IN FEDERATED LMS

The use of identity federations has several important consequences for the acquisition and protection of user data. These consequences are especially important for federated LMS infrastructures because much highly sensitive data is stored therein, e.g. each learner's full name, birth date, and matriculation number. The misuse of this data clearly must be prevented. Although many HEIs are either authorized by laws or acquire each student's consent to process PII in a campus-wide manner as a part of the enrollment process, the transfer of user data to 3rd parties, such as an external LMS SP, by an IDP is a new data processing purpose that needs to be dealt with appropriately.

In general, the user should give her informed consent to the transfer of her personal data to an SP. As a current best practice, IDPs use the business card metaphor for this purpose as shown in figure 2: Whenever the user is about to login to an SP for the first time, a digital business card is shown that lists all the user attributes that will be transmitted by the IDP to the SP. The user can then confirm that she agrees to the transfer of this data to the SP, or has the option to abort the login process; the latter action implies that the learner cannot use the service unless there are other ways to sign up, such as traditional self registration forms. However, first experiences clearly showed that the user acceptance of the business card based consent expression is next to 100%.

However, users should not give their consent unless they have read and acknowledge the SP's terms of use and online privacy statement: Once the IDP has sent the user's data to the SP, it cannot be held responsible for how the data is processed - or misused - by the SP. Thus, SPs should clearly state which purposes they need the personal data for and how long the retention time is. Usually, the learner's LMS account, which includes the PII, can safely be deleted after the courses and exams have finished and the results were sent to the examination offices. However, some SPs may want to keep the users' email addresses, e.g. to inform them about new and related courses in the following semesters; such data processing purposes generally require another explicit consent by the affected users.

Furthermore, for obvious reasons the PII should be encrypted at least while in transit, i.e. when sent by the IDP over the Internet to the LMS SP. Surprisingly enough, many commercial federation software packages still default to plain text user data transfer, but Shibboleth, which currently is the most widely deployed FIM software among HEIs worldwide, is pre-configured pleasantly well in this regard.

Finally, the existing LMS security policies must be updated to reflect the fact that external users' personal data is stored in the LMS. The access rights of privileged users, such as instructors, must be precisely specified and enforced in accordance with the privacy policy that the external users have consented to. This is especially important when learners, instructors, and the LMS provider are from three separate organizations, e.g. in central hosting scenarios.

This is the first time you use the service "https://lms-provider.example.com". The following user data is about to be sent to the service:

Digital ID Card	
Family name	Doe
First name	Jane
Gender	Female
Preferred language	English
Email address	jane.doe@idp.example.com
Study course	Computer science (master)
Term of study	6th semester
Matriculation number	1234567
Local account name	jdoo
Display name	Jane Doe
Home organization	Example university

I have read the service provider's terms of use and online privacy statement. I consent to the transmission of my personal data listed above to this service.

Abort Confirm

Figure 2: The business card metaphor facilitates informed user consent to PII transfer (the example shown is based on the SWITCHaai uApprove Shibboleth IDP plug-in)

5. ADAPTING LMS PRODUCTS AND IDENTITY PROVIDERS

Although about a dozen of LMS products has already been adapted for identity federation software like Shibboleth (see (Oishansky & Carmody, 2009)), many other LMS software packages are currently in use at German HEIs that still need to be extended to make use of the new federation features enabled by the DFN-AAI E-Learning Profile. However, the willingness to implement federation support has immediately risen after the country-wide concept was published, clearly opposed to the previous experiences when only regional requests were made to the vendors. For example, it took less than six months from the DFN-AAI E-Learning Profile publication to the release of a Shibboleth- and DFN-AAI-enabled new version of the commercial LMS imc CLIX (imc Advanced Learning Solutions, 2008), which also incites other vendors to implement federation support. In general, the following LMS components need to be adapted:

- The authentication and authorization modules must support the traditional local login workflow in parallel to federated single sign-on and IDP-enhanced authorization. LMS products that already support other types of single sign-on and identity management based authorization, e.g. via enterprise LDAP directory services, are most likely easier to adapt than monolithic LMS architectures.
- The data provided by the IDP, which complies with the DFN-AAI E-Learning Profile in our case, needs to be filtered and converted to the data schema that is used by the LMS internally. Again, not too much new functionality is required for LMS products that already support other types of data import, such as from LDAP servers.
- Federation membership must be supported, i.e. the LMS must become an SP in the federation and verify the federation membership of the IDPs, which the users login from. This generally requires LMS-internal workflows to automatically download, regularly update, and parse the XML-based federation metadata files.

- Support for parallel membership in multiple federations should be provided to facilitate the international use of the LMS. Consequently, parts of the discovery service functionality (also known as Shibboleth Where Are You From WAYF service) should be included in the LMS, i.e. the user should be able to choose her home federation / university from the LMS' login page.
- When a new user logs in via the federation for the first time, the terms of use and privacy policy should be shown along with the PII that is going to be stored in the LMS. At each subsequent login, the user data should be refreshed by retrieving it from the IDP again to keep it up-to-date and to avoid inconsistencies between IDP and SP.

Modifications are also necessary to those IDPs whose back-ends did not provide the new E-Learning Profile attributes yet. For example, many HEI identity management systems did not yet use the official FSO identifiers for fields of studies, and attributes like the place of birth were often not used for identity management at all yet. These IDPs need to extend their import interfaces to the student administration software, which in turn requires HEI-internal changes that usually have to be discussed with the local privacy officer. Consequently, putting the DFN-AAI E-Learning Profile into production use is not a trivial, quick task on either side, but certainly worth the efforts.

6. CONCLUSIONS AND OUTLOOK

The DFN-AAI E-Learning Profile is a milestone for the German HEI federation because it ensures interoperability and allows the fully-featured use of LMS across organizational borders. The federation-based transfer of user records in addition to the inter-organizational single sign-on functionality and the delegated authorization concepts is a very user-friendly and data quality enhancing solution. After several regional and proprietary attempts were made, the country-wide solution presented in this article has already motivated commercial LMS vendors to adapt their products, and the consensus reached in the DFN working group is a solid foundation for the future of inter-organizational E-Learning in Germany.

That said and given the historical development from chair- and faculty- to HEI-wide LMS deployments, which then moved on to regional co-operations and now provide country-wide interoperability, it becomes obvious that also the real-world demand for international, pan-European E-Learning will rise as has already been predicted for several years. Since most of the technical issues are successfully addressed by adapting LMS products for federation use and providing the relevant user attributes via the IDPs, the success of international E-Learning is within our grasp, but still largely depends on organizational issues. For example, common identifiers for fields of studies need to be specified and kept up-to-date. This clearly requires serious efforts and pragmatic solutions on the European scale, and we certainly want to encourage other federations to join these discussions and efforts.

7. ACKNOWLEDGMENTS

The work presented in this article has been carried out by the members of the DFN-AAI E-Learning Profile working group as well as E-Learning and federation enthusiasts at many German HEIs. I would especially like to express my gratitude to my fellow authors of the DFN-AAI E-Learning Profile specification: Jörg Deutschmann, Peter Gietz, Renate Schroeder, Jens Schwendel, and Tobias Thelen. I also thank Ulrich Kaehler at DFN for his major coordination efforts, and Stephan Graf and Armin Rubner for their significant contributions to get the work on the E-Learning Profile started as well as done.

The author also wishes to thank the members of the Munich Network Management (MNM) Team for helpful discussions and valuable comments on previous versions of this article. The MNM-Team, directed by Prof. Dr. Dieter Kranzlmüller, Prof. Dr. Heinz-Gerd Hegering, and Prof. Dr. Gabi Dreo-Rodosek, is a group of researchers of the University of Munich, the Munich University of Technology, the University of the Federal Armed Forces Munich, and the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences. The team's web-server is located at <http://www.mnm-team.org/>.

8. REFERENCES

- Deutschmann, J., Gietz, P., Hommel, W., Schroeder, R., Schwendel, J., & Thelen, T. (2008). *DFN-AAI E-Learning Profil (German)*. https://www.aai.dfn.de/fileadmin/documents/attributes/200811/DFN-AAI_E-Learning-Attribute_V.1.0.pdf: DFN-Verein.
- DFN-Verein. (2009). *Germany's HEI Authentication and Authorization Infrastructure*. Retrieved from DFN-AAI Homepage: <https://www.aai.dfn.de/>
- imc Advanced Learning Solutions. (2008, November 14th). *LDAP, Shibboleth & Co*. Retrieved from CLIX Webinar Part 3: <http://www.im-c.de/de/unternehmen/events/clix-webinar-reihe/>
- Olshansky, S., & Carmody, S. (2009). *Shibboleth Enabled Applications and Services*. Retrieved from <https://spaces.internet2.edu/display/SHIB2/ShibEnabled>
- TERENA Task Force EMC2. (2009). *SCHAC Attribute Definitions for Individual Data v1.4.0*. <http://www.terena.org/activities/tf-emc2/schacreleases.html>: TERENA.
- VHB. (2009). *The Concept of the Virtual University of Bavaria*. Retrieved from <http://www.vhb.org/en/students/concept/>