Privacy Issues when Rolling out an E-Learning Solution

Andreas Zinnen, Eicke Godehardt, Manuel Görtz, Andreas Faatz, Robert Lokaiczyk
SAP Research CEC Darmstadt
Bleichstr. 8, 64283 Darmstadt, Germany
{andreas.zinnen, manuel.goertz, andreas.faatz, robert.lokaiczyk, eicke.godehardt} @sap.com

Abstract. This paper describes privacy issues while rolling out an socio-technical solution into small or medium enterprises (SMEs) or departments of a global (European) company. We will use an e-Learning system as an exemplary system. The solution delivers relevant knowledge artefacts to workers so that they can learn from them effectively within their work processes. In order to select those knowledge artefacts, the system accesses and stores privacy critical information like users' competencies, performed tasks or social contacts. In consultation with several companies' departments of Works Council, Corporate Legal and Data Protection & Privacy Office a privacy policy was specified to fulfil the seven principles *Notice, Purpose, Consent, Security, Disclosure, Access and Accountability* as stated in OECD's recommendations for protection of personal data and in directive 95/46/EC on the protection of personal data. Instead of claiming to be complete, this paper addresses the privacy basics which might be slightly different in another use case depending on boundary conditions of the company or country. The integration of mechanisms to comply with the privacy rules is often neglected in research projects. However, these should be considered while designing research prototypes used for evaluation within real working environments.

Introduction

We have observed that within the research community there is a lack of considering privacy issues in the beginning of system dealing with personal data such as an e-Learning system. However, considering security or privacy requirements right from the start of system design is one of the finding of security engineering. Within this paper we describe how the EU regulations and privacy policies in different German companies are reflected in the design and rollout of the according prototypes developed.

In general, data privacy refers to the relationship between technology and the public expectation of privacy in the collection and sharing of data. Privacy concerns exist wherever uniquely identifiable data related to a subject or several subjects are collected and stored. The storage might be in digital form or otherwise. Improper or non-existent disclosure control can be the root cause for privacy issues. The challenge of data privacy in many IT systems is to both share data and protect personally identifiable information. The seven principles *Notice, Purpose, Consent, Security, Disclosure, Access and Accountability* as stated in OECD's recommendations for protection of personal data and in directive 95/46/EC (European Commission 2004) or the Directive 2002/58/EC of the European Parliament and of the Council (European Parliament 2002) on the protection of personal data demand additional privacy related steps have to be addressed by the system before rolling out an e-Learning solution within European companies. This paper summarizes the effort and outcome of specifying a privacy policy for an e-Learning system in consultation with several companies' Works Council, Corporate Legal and Data Protection & Privacy Office.

The goal of the e-Learning solution is to enhance knowledge worker productivity by supporting informal learning activities in the context of knowledge workers' everyday work processes and within their work environments. The approach tries to integrate seamlessly into the worker's environment and to support him in all his roles. Among others, a solution delivering relevant knowledge artefacts to workers so that they can learn from them effectively has to store privacy critical information like users' competencies, social network or performed tasks. This fact potentially gives rise to a misuse of the private data for employees' performance control.

We have identified three dimensions – as shown in Figure 1 – that have to be considered in order to address privacy issues in a socio-technical system in a proper way. All these dimension will be covered in this paper using the e-Learning solution as an exemplary system. However, the technical and legal dimension will be the main part.

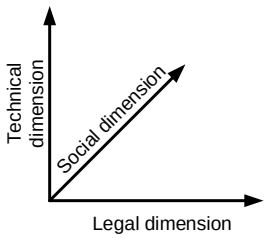

Figure 1 Three Dimension to Address the Privacy Issues

The rest of the paper is organized as follows. Section 2 summarizes foundations of privacy in the European Union. Overall, the directive 95/46/EC builds the base of the European privacy laws. Section 3 steps into data protection principles at the workplace of a German company. The three instances the Works Council, Corporate Legal and Data Protection & Privacy Office provide for data protection within global software companies. Section 4 introduces the basic concepts of the e-Learning system. Section 5 summarizes the elaborated policy and steps into technical details how privacy requirements are currently solved within the e-Learning solution.

Privacy in European Union

The Universal Declaration of Human Rights states in its §12 that:

> *No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.*

The right to privacy is a highly developed area of law in Europe. All the member states of the European Union (EU) are also signatories of the European Convention on Human Rights (ECHR). Article 8 of the ECHR (Council of Europe 1950) provides a right to respect for one's "private and family life, his home and his correspondence," subject to certain restrictions. The European Court of Human Rights has given this article a very broad interpretation in its jurisprudence.

In 1980, in an effort to create a comprehensive data protection system throughout Europe, the Organization for Economic Cooperation and Development (OECD) issued its “Recommendations of the Council Concerning Guidelines Governing the Protection of Privacy and Trans-Border Flows of Personal Data”. The seven principles governing the OECD’s recommendations for protection of personal data were:

- *Notice*: Data subjects should be given notice when their data is being collected;
- *Purpose*: Data should only be used for the purpose stated and not for any other purposes;
- *Consent*: Data should not be disclosed without the data subject’s consent;
- *Security*: Collected data should be kept secure from any potential abuses;
- *Disclosure*: Data subjects should be informed as to who is collecting their data;
- *Access*: Data subjects should be allowed to access their data and make corrections to any inaccurate data; and
- *Accountability*: Data subjects should have a method available to them to hold data collectors accountable for following the above principles

But the OECD Guidelines were nonbinding, and data privacy laws still varied widely across Europe. Therefore the European Commission decided to unify data protection regulation and submitted the Directive 95/46/EC on the protection of personal data.

In the directive 95/46/EC, personal data is defined as "any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;" (§ 2a). This definition is meant to be very broad. Data are "personal data" when someone is able to link the information to a person, even if the person holding the data cannot make this link. Some examples of "personal data" are address, credit card number, bank statements, criminal record, birth date etc.

The notion processing means "any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval,

consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;" (§ 2b). The responsibility for compliance rests on the shoulders of the "controller", meaning the natural or artificial person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; (§ 2d).

The data protection rules are applicable not only when the controller is established within the EU, but whenever the controller uses equipment situated within the EU in order to process data. (§ 4) Controllers from outside the EU, processing data in the EU, will have to follow data protection regulation. In principle, any on line shop trading with EU citizens will process some personal data and is using equipment in the EU to process the data (the customers' computer). As a consequence, the website operator would have to comply with the European data protection rules. The directive was written before the breakthrough of the Internet, and to date there is little jurisprudence on this subject.

Personal data should not be processed at all, except when certain conditions are met. These conditions fall into three categories:

- *Transparency*: The data subject has the right to be informed when his personal data are being processed. The controller must provide his name and address, the purpose of processing, the recipients of the data and all other information required to ensure the processing is fair. (§ 10 and 11). Data may be processed only when (§ 7):
  - The data subject has given his consent
  - The processing is necessary for the performance of or the entering into a contract
  - Processing is necessary for compliance with a legal obligation
  - Processing is necessary in order to protect the vital interests of the data subject
  - 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
  - Processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject

  The data subject has the right to access all data processed about him. The data subject even has the right to demand the rectification, deletion or blocking of data that is incomplete, inaccurate or isn't being processed in compliance with the data protection rules. (§ 12)
- *Legitimate purpose*: Personal data can only be processed for specified explicit and legitimate purposes and may not be processed further in a way incompatible with those purposes. (§ 6b)
- *Proportionality*: Personal data may be processed only insofar as it is adequate, relevant and not excessive in relation to the purposes for which they are collected and/or further processed. The data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected or for which they are further processed, are erased or rectified; The data shouldn't be kept in a form which permits identification of data subjects for longer than is necessary for the purposes for which the data were collected or for which they are further processed. Member States shall lay down appropriate safeguards for personal data stored for longer periods for historical, statistical or scientific use. (§ 6). When sensitive personal data (can be: religious beliefs, political opinions, health, sexual orientation, race, membership of past organisations) are being processed, extra restrictions apply. (§ 8) The data subject may object at any time to the processing of personal data for the purpose of direct marketing. (§ 14) A decision which produces legal effects or significantly affects the data subject may not be based solely on automated processing of data. (§ 15) A form of appeal should be provided when automatic decision making processes are used.

EU directives are addressed to the member states, and aren't legally binding for citizens in principle. The member states must transpose the directive into internal law. Directive 95/46/EC on the protection of personal data had to be transposed by the end of 1998. All member states have enacted their own data protection legislation.

Data Protection Principles@Workplace in a German Company

The EU Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data limits and regulates the collection of personal information on individuals of workers.

Firms that monitor employees' use of e-mail, the internet or phones as part of their business practice, and do not tell employees or have not obtained employee consent to do so, can in most cases be sued under Article 8 the European Convention on Human Rights which provides for the right to respect for his private and family life. On the other hand, although EU law is clear that e-mail interception is illegal, the law is not totally clear as to whether companies may prohibit employees from sending private e-mails. The figure below illustrates involved roles in the data protection process at the German company.

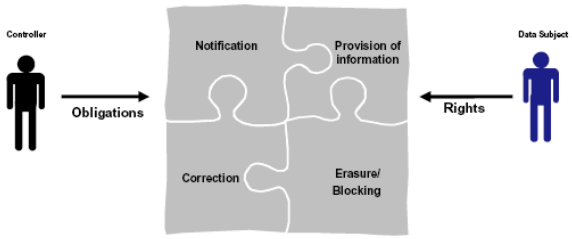

Figure 2 Data Protection Obligations and Rights

The controller is the legal party in charge of the processing of the data, e.g. the company in general, but also derived down to the single employee processing data on behalf of the company. As an employee of the global company or one of the subsidiaries you will get in touch with data protection in many ways:

1. You are working in areas where personal data is collected and processed
2. You have access to personal data, e.g. to the internal address book with e-mail addresses, phone numbers, positions of all employees in intranet
3. You are developing software for processing of personal data, which has to be compliant to data protection
4. You collaborate with customers, suppliers and other subsidiaries of the company and transfer or receive personal data
5. Your own personal data is processed by the company and/or the subsidiaries, you are also „Data Subject“ to this processing

Three organizational units supervise the rollout of privacy critical products within the company:

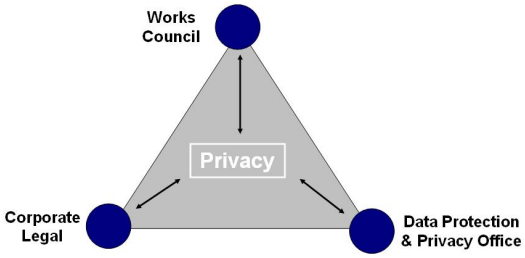

Figure 3 The Three Organisational Units Supervising the Rollout

*Corporate Legal r*epresents the company in proceedings before the labor court and provides advice to the Executive Board regarding the implementation of legal changes. Furthermore, the unit contributes to projects of the Executive Board and the HR department and guides the collaboration between the company's specialists for European law and legal regulations. The *Works Council (Betriebsrat)* represents workers and acts as local/firm-level complement to national labor negotiations. The *Data Protection & Privacy Office (DPPO)* is a supervisory authority ensuring the protection and privacy of all data collected, processed or used within the company's field of responsibility including employees' and clients' data. DPPO supports the company itself and its employees in following and meeting the security laws.

An e-Learning Solution as Showcase for Dealing with Privacy Issues

In the following section we will describe an e-Learning solution in more detail and show how our findings about privacy issues as described in the EU regulations as well as in the policy of German companies have been addressed by the system. The described solutions are considered mainly legally or technically according to the three dimensions in (Figure 1).

The System's Benefits for the user

The e-Learning solution enhances knowledge worker productivity by supporting informal learning activities in the context of knowledge workers' everyday work processes and within their work environments. Therefore, the system needs to know a number of information about its users to provide its full potential of functionalities to the user. The approach provides integrated ICT support for the three roles a knowledge worker fills at the professional workplace: the role of Learner, the role of Expert, and the role of Worker.

- *Work:* The e-Learning solution automatically identifies the knowledge worker's needs and provides context-sensitive support tailored to their specific learning goals and work situations. A crucial factor for a task-oriented e-Learning software is the user's context. To provide both suitable and helpful learning resources to the user, the application always has to consider the learner's actual work task, his environment, and history.
- *Learn*: The e-Learning solution helps knowledge workers explore, apply and reflect on knowledge in a self-directed manner: By considering their work context, the system ensures that learning and working are tightly integrated and learning is transferred to actual workplace tasks. Furthermore, the system considers a user's performed tasks (task profile) when proposing adequate learning resources.
- *Collaborate:* The e-Learning solution helps knowledge workers to convey and jointly create knowledge via their computational environment and embedded in their work context. The context of knowledge transfer and creation is captured in order to turn knowledge artefacts into learning resources. The system proposes collaboration partners with low social distance for specific problems. Both participants of collaboration have to agree on the collaboration.

The System & Collected Data

Chiefly, this support is provided within the work environment, and not in a separate learning environment. The e-Learning solution exploits synergies between learning and knowledge management by reusing content not originally intended for learning. It utilizes contextualized communication such as described in (Goertz et. al 2004) for knowledge transfer, and ease the burden on experts for these tasks. Finally, it is based on knowledge sources available within an organization – specifically business space, e-Learning systems, and knowledge management – and not require a switch to a new system but provide an integrated system.

The system is designed as a client/server architecture. The data protection needs to be in place on the client side where sensitive data is collected as well as on the server side where all data is stored. In (Figure 4) the system architecture from the client side is shown.

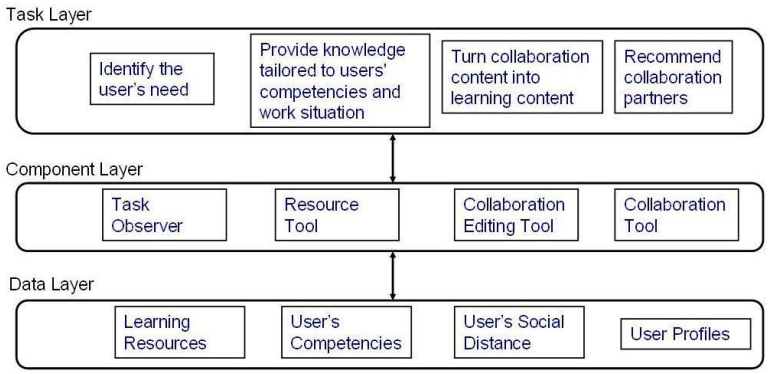

Figure 4 The System Architecture in a Schematic View.

The following data is collected and applied in the e-Learning solution deployed in several independent legal entities.

1. Activity Recognition: (Task Observer). To detect the current user task user interaction and system events are logged. These data is later applied to the models to users for activity recognition.
    - User Interaction: Keyboard and Mouse interaction (mouse position, entered words, letters) of several test users will be collected. The data is stored as hash values, not in clear text. The values will be used for statistical analysis. Results of the statistical analysis are activity models for each trained activity.

- System Status: The system status includes activities on the file system (access/delete/modify/rename/delete directories/files/documents), started/stopped applications, sending/receiving emails, visited or bookmarked web sites, content of documents, visited web sites or emails, printed documents, and finally the username. The data is stored as hash values, not in clear text. The hash values will be used for statistical analysis.

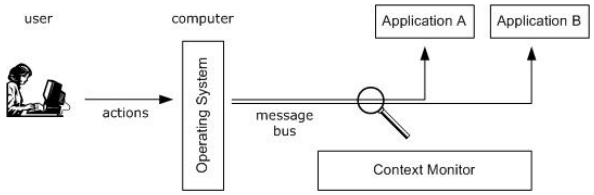

Figure 5 Collection User Interaction Events

2. Collaboration: To find adequate experts as described in (Lokaiczyk et.al. 2007), the system stores the following user data. This data can be seen by all users.
   - Starting a collaboration:
     - In-house Address/Email/Phone Number/Messenger IDs: Other participants can contact you via e-mail, mail, phone or messenger. You can choose not to be contacted at any time.
     - Organizational Unit/Job Role: The organizational unit and job role are considered in the social distance between participants while proposing collaboration partners.
     - Name/First Name/Picture: Name, first name and a picture are part of the general contact information. The picture mainly helps to visualize the collaboration partner. The user can choose to not publish his/her picture.
     - Performed Tasks, number of executions: For a selection of according experts, the system stores information about users' performed tasks and processed learning resources. The system also stores how often a task was performed. Assured by a low granularity of the collected information, conclusions about users' performance and overall learning goals in a whole work context cannot be drawn.
   - Extracting new learning artifacts from collaboration material like chat logs, phone calls etc.: Contents of the collaboration (e.g. chat contents) might be extracted and reused as future learning resources. Learning artifacts from collaboration will not be stored and published automatically. The e-Learning solution will provide an editing tool for the collaboration contents. Both parties must approve the publication before the contents will be accessible to third persons. A proper objectification is primitive before the publication. After making the artefacts public, they can be provided as learning material to all users of the system.
3. Retrieval of adequate learning material: For a selection of the according learning resources, the system stores information about users' performed tasks and processed learning resources. Assured by a low granularity of the collected information, conclusions about users' performance and overall competencies in a whole work context cannot be drawn.

A proper maintenance of the learning resources and personal learning goals in the system is a key for achieving the objective to provide an advanced process oriented self-directed learning environment. Every participating person is asked to create and maintain their personal performed task profile as well as knowledge artefacts. Participation is voluntary. It is suggested that you check and update the information in your personal performed task profile at least several times a year.

Administrators of the system will have access to all collected data. Users of the e-Learning solution will see your contact information and performed tasks as described in following details section. Access to collaboration contents will be granted to other persons after explicit agreement.

How the Privacy Regulations have been Addressed?

As motivated in the introductory section, a solution delivering relevant knowledge artefacts to workers so that they can learn from them effectively has to store privacy critical information like users' competencies, social network or

performed tasks. This fact potentially gives rise to a misuse of the private data for employees' performance control. This section steps into details how privacy can be ensured to a specific level considering all three dimensions.

On the social dimension the system visualizes the ongoing processes as much as possible to make them more scrutable to the user. Additionally, the system provides the user with a number of functions to control the system such as switching logging on and off. The users' personal data is collected and used only with her explicit consent, given by actively clicking the "I accept" button". In case that user declines to accept this privacy agreement, the user will not be able to use the e-Learning solution. The collected data within the system is used for the operation of the system itself and can be used for statistical analysis of research questions

In accord with several companies' Works Council, Corporate Legal and Data Protection & Privacy Office, the authors developed a privacy policy for the e-Learning solution consistent with current European privacy laws. Users of the system have to accept the policy before using the system. The policy will be shown while starting the system the first time. Users can withdraw their accordance at each time. Beyond signing the policy, the system has to fulfill following seven privacy requirements (see Section "Privacy in European Union ") based on the Directive. These have been fulfilled either using legal policies or technical methods:

- *Notice*: Data subjects should be given notice when their data is being collected. All users have to accept the privacy policy before being able to use the system. The policy contains a list of data that will be collected. Data of collaborations will only be collected with both collaboration partners' consensus. After revoking the agreement, all private data will be deleted or made anonymous. Additionally, the user will see when the data is collected. A visible signal (a red lamp) will show that the data recording is taking place. The user can at any time stop this process.
- *Purpose*: Data should only be used for the purpose stated and not for any other purposes. The policy contains the purpose of data collection. The data is not accessible by other persons or programs. The only purpose of data collection is to provide users the right learning resources and experts in time.
- *Consent*: Data should not be disclosed without the data subject's consent. The only disclosed data is listed in the privacy policy. If a user does not want this data to be shared, she should not accept the agreement or revoke it.
- *Security*: Collected data should be kept secure from any potential abuses. The system stores all critical data on a central file server. The data will be available to all users. All data is stored anonymous; therefore, drawing conclusions from the hash values or models to the test users' identities cannot be drawn. Furthermore, only the administrators will have access to all data. Low level data will be collected as hashed value (Lokaiczyk, Goertz & Faatz 2007) and not in clear texts. All data is stored anonymous; therefore, drawing conclusions from the hash values or models to the test users' identities cannot be drawn.

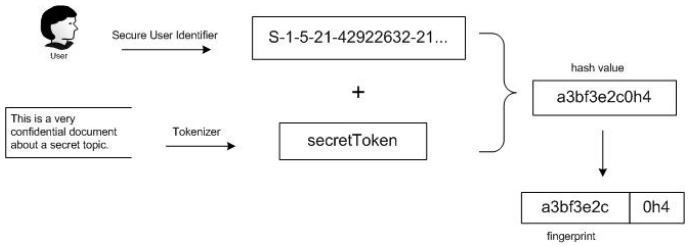

Figure 6 Using Hash Function to Ensure Security of Collected Data

- *Disclosure*: Data subjects should be informed as to who is collecting their data. Reading the privacy policy, the users know who is collecting which data.
- *Access*: Data subjects should be allowed to access their data and make corrections to any inaccurate data. Different tools offer the users to modify the collected data. The Competency tool supports users to manage the documentation about performed task of finished learning resources. A collaboration editing tool offers a way for editing collaboration content and stores it in an anonymous form. A user profile management tool is provided to administer private data like name, contact, etc. To restrict the access to the stored data on the server a privacy policy has to be implemented enforcing access policies. Several frameworks such as XACML (OASIS 2005) or Web Service Policy (W3C 2006) exist. These together with security mechanisms (IBM 2004) will ensure a secure data transfer and handling. The privacy enforcement will ensures that all requests to any of the methods within this service will be channeled through the policy engine.

- *Accountability*: Data subjects should have a method available to them to hold data collectors accountable for following the above principles

Summary

This paper describes privacy issues while rolling out an e-Learning solution into organizations. The solution delivers relevant knowledge artefacts to workers so that they can learn from them effectively within their work processes. In order to select those knowledge artefacts, the system accesses and stores privacy critical information like users' competencies, performed tasks or social contacts. In consultation with several companies' Works Council, Corporate Legal and Data Protection & Privacy Office, a privacy policy was specified to fulfil the seven principles Notice, Purpose, Consent, Security, Disclosure, Access and Accountability as stated in OECD's recommendations for protection of personal data and in directive 95/46/EC on the protection of personal data.

Although our findings are based on a research project, privacy is an important issue that cannot be ignored. It is not allowed to collect private data even for research purposes in a global European company. The users always must be informed and explicitly agree that the specified data can be collected. Our findings should raise awareness and provide some basic insights into the topic of privacy in socio-technical systems.

References

Council of Europe (1950) *Convention for the Protection of Human Rights and Fundamental Freedoms.* ETS 5; 213 UNTS 221. http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=005&CM=8&DF=4/25/2006&CL=ENG.

European Commission (2004) *EU Data Protection Page.* http://ec.europa.eu/justice_home/fsj/privacy/

European Parliament and of the Council (2002) *Directive 2002/58/EC*. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:HTML.

Goertz, M., Ackermann, R., Schmitt, J., & Steinmetz, R. (2004) Context-aware Communication Services: A Framework for Building Enhanced IP Telephony Services. *International Conference on Computer Communications and Networks (ICCCN)* 2004, p. 272--279, October 2004.

IBM (2004) *Web Services Security (WS-Security)*. http://www.ibm.com/developerworks/library/specification/ws-secure/

Lokaiczyk, R., Goertz, M., & Faatz, A. (2007) Towards Improving Privacy and Security in Context-Aware Workplace-Embedded e-Learning Environments through Data Obfuscation.. Flexibel integrierbares e-Learning - Nahe Zukunft oder Utopie? *Workshop on e-Learning* 2007 ISSN 1613-0073 187-196

Lokaiczyk, R., Godehardt, E., Faatz, A., Goertz, M., Kienle, A., Wessner, M., & Ulbrich, A. (2007) Exploiting Context Information for Identification of Relevant Experts in Collaborative Workplace-Embedded e-Learning Environments *EC-TEL 07*

OASIS (2005). Privacy policy profile of XACML v2.0. http://docs.oasis-open.org/xacml/2.0/access_control-xacml-2.0-privacy_profile-spec-os.pdf

W3C (2006). *Web Services Policy 1.2 - Framework (WS-Policy)*. http://www.w3.org/Submission/WS-Policy/