Aligning User Consent Management and Service Process Modeling

Nils Gruschka
nils.gruschka@fh-kiel.de

Meiko Jensen
meje@mmmi.sdu.dk

Abstract: With evolving functionality of Internet services, the management of user consents becomes a complex challenge. As consents are a common means for establishing a legal basis for processing privacy-relevant user data, a sound consent management approach is required.

In this paper, we outline an approach for semi-automated generation of letter of consent documents, based on existing service implementation documentation. We illustrate the challenges of consent management in relation to service evolution, and we outline an integration of consent management into model-based process development systems.

1 Introduction

In times of continuous evolution of the Internet, with new features being added to online services on a daily basis, it becomes more and more challenging to maintain oversight on the different versions of services a provider has offered in its history. However, when it comes to court rulings, a particular user's consent to data processing can easily become a key parameter for determining law compliance of an online service, e.g. in terms of data protection regulations. With the emerging set of side-effects of online services (e.g. regarding surveillance, trading of personal data, or data theft), the future Internet is going to see a lot of lawsuits discussing what service a human individual has consented to, what scope that consent had, and how the service evolved after the consent was given. Hence, it becomes necessary to keep track of all forms of consent a service user has given explicitly, along with the scope of that consent. This challenge of *consent management*, which already emerged on a larger scale within the healthcare domain (cf. [HS13, Bon13, JIS11]), is about to become one of the most important concepts in the Future Internet.

In this paper, we introduce general requirements of *consent management systems* (cf. [Con10, epS, Tex12]), and elaborate their parameters and conditions. We continue to discuss the integration of consent management systems into state-of-the-art e-commerce suites, and we analyze different integration options for common process description languages like UML or BPMN. We describe the two main aspects of consent management (which are *consent documentation* and *consent information*), and we explore real-world applicability and future research directions within the domain of consent management systems.

2 Consent to Processing of Personal Data

One of the key necessities for providing Internet services within Europe is legal compliance. Among the multiple regulatory norms that define the European legal basis, the regulations for privacy and data protection [The95, The15] are among the more important ones. Basically, for most of these norms, it is strictly required by law to adhere to a certain set of conditions whenever a service provider processes any sort of data that directly relates to a human individual, i.e. is *personally-identifiable information*.

In this case, the particular service provider is required to identify the legal basis for such processing, which e.g. can be specific laws of the country the service provider operates within. However, one of the most common means to form such a valid legal basis for service operation is the explicit *consent* to data processing, given by the affected individual herself. Within certain boundaries, explicit consent of all affected individuals allows for almost any sort of processing of personally-identifiable information, including the option to forward personal data to other organizations, such as business partners.

A key aspect of such a consent, however, is that it must fulfill a certain set of conditions itself, in order to become a valid legal basis. Any form of legally valid consent must be

- given *freely*,
- given by an *informed* citizen,
- *documented* explicitly, and
- *revocable*.

Therein, *free* means that the user of a service had a free choice to give or not to give consent. She must not have given consent involuntarily, i.e. may not be punished for rejecting to give consent. It must be the individual's own choice, and her own decision to give consent to having her personal data being processed by the particular service. Any consent given under pressure becomes invalid, and thereby invalidates the legal basis for processing of affected personal data.

The term *informed* reflects that a citizen must know in detail about the conditions under which her personal data is going to be processed. Most importantly, this directly involves knowledge of the following:

- Who is the legal entity responsible for this particular service?
- What is the purpose of processing of a user's personal data?
- Which sort of personal data is to be used within the service?
- If personal data is stored, for how long is it maintained?
- To which other legal entities is the personal data forwarded?
- For what purposes is the personal data forwarded to other legal entities?

As can be seen, the definition of informed consent pre-assumes a lot of information to be provided by the service operator towards the users a priori. Typically, such information can be given by means of a *terms of use* document, which the individual user explicitly has to agree to, e.g. by clicking some “OK” button. Another option is the explicit definition of a *privacy policy* that contains all relevant information.

Documentation of given consent is a key requirement in order to prove the existence of consent at a later stage. If, for instance, a Data Protection Authority challenges the legitimacy of a certain sort of data processing, the scope and validity of a given consent is of critical importance. Hence, a service provider is urged to document the exact conditions under which the consent was given, and the scope it covers.

On *revocation* of consent, i.e. if a user changes her mind and does not allow any further processing of her personal data at the service provider, the service provider is obliged to respect and follow that decision within reasonable means. For instance, this implies to stop any ongoing process instances that handle personal data of the particular individual, and to prevent subsequent processes from starting when they are processing such personal data. Also, deletion of personal data (within the boundaries defined by other legal norms) must be performed on user request.

3 Consent Management

As can be seen, the concept of consent to personal data processing is quite complex, and bears a lot of challenges, especially for digital services provided over the Internet. In face of growing number of users, continuous changes to the implementation of a service, and ongoing alterations in the set of collaborating business partners, keeping an overview of all events related to user consent becomes a management issue. Within this challenge, the following characteristics become important for *consent management*:

- Each given consent must be documented in a durable, standardized way that allows for subsequent recall.
- For each given consent, the exact conditions of the service implementation for exactly the moment the consent was given (i.e. the *scope* of the consent) have to be documented as well.
- Whenever a user revokes consent, this information has to be documented as well, and must be linkable to the documentation of the consent that actually was revoked.
- Changes in the service implementation might lead to changes in the scope of consent. Hence, each change in the implementation must trigger a re-validation of the information given to users prior to consent.
- Changes in the set of business partners involved in processing of personal data also must be reflected in the user consent information.

As can be seen, the management of consent has two sides. On the one hand, all user decisions regarding consent have to be documented explicitly. On the other hand, for each of these decision events, the scope of consent, i.e. the information provided to the individual user, and the state of the service implementation at the time of the decision event, have to be documented as well.

Derived from these observations, it becomes evident that consent management consists of two parts: *User Consent Information* deals with the informational text provided to (existing or upcoming) users of a service, in order to gather an informed consent, as obliged by regulatory norms. Hence, this part deals with gathering information on the state of a service's technical and organizational implementation, and transforming that information into a representation that fulfills the property of *informedness* as required for consent by law. *User Consent Documentation* covers all tasks that result from a user's decisions regarding giving or revoking consent. This includes documentation of user's decisions, documentation of the state of the service at the time of decision, and implementation of all technical and organizational tasks that result from each type of user consent decision.

4 Consent Management vs. Service Management

Taking the characteristics of consent management as a basis, it becomes immediately clear that the challenge of consent management is directly linked to the challenge of service management. Every event in service management, such as the roll-out of a service update, must be considered in consent management as well, as it might affect the scope of future consents. Vice versa, every user-generated event in consent management, i.e. giving or revoking consents, may affect the service implementation, as it must e.g. trigger deletion of personal data, or shutdown of affected process instances.

For the technical aspects of service management, the state of the art today already covers a broad spectrum of techniques, such as explicit process modeling (e.g. by means of UML [RJB04], BPMN [OMG13], or others) and well-defined implementation change management processes (cf. e.g. [PM08]). However, those techniques are typically more focused on technical aspects of a service, such as implementation details, security properties, organizational aspects, or external event handling. To the best of our knowledge, as of today, there exists no relevant technological approach to align service and process management with consent management. This gap is to be addressed within the remainder of the paper.

The upper part of Figure 1 shows the typical workflow of software development procedure. A software developer:

a) performs a requirement analysis (not shown in the picture),

b) creates the software or process design (using modelling techniques),

c) uses the model as a basis for implementing the software in a programming language,

d) compiles the software and deploys the resulting service.

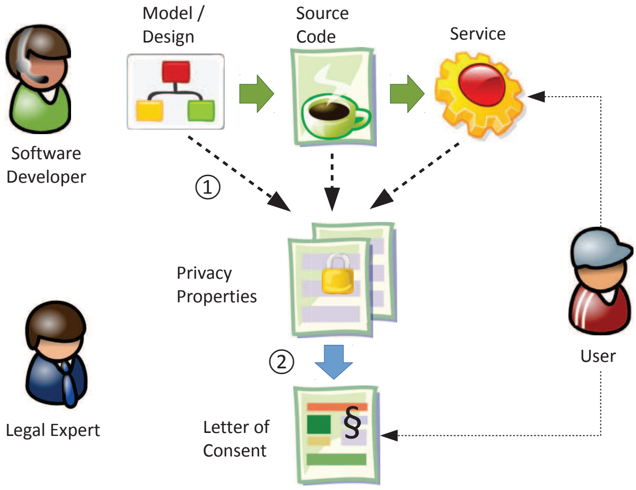


Figure 1: Architecture for consent management

In all these steps, information on processing of data is handled, including processing of private user information. Thus, we propose to extract all privacy-relevant information from the existing software design process (step 1). These privacy properties can be used as a basis for a legal expert to create a letter of consent (step 2). This letter is presented to the service user, who can read and agree to the privacy conditions of the service she is intending to use.

This combination of service management and consent management has a number of advantages (compared to current approaches), e.g. precise process information inside the letter of consent, better workflow for the legal experts, and fine-grained lifecycle management. Section 6 presents how privacy properties can be derived from the service design process. Before, we introduce an example service process, which will be used for illustrating our approach.

5 Example Process: Travel Agency

In the remainder of the paper we will use the example of a travel agency service. The service process is shown in Figure 2.

The user calls the travel agency service giving her name, address, date of birth and credit card number (plus the details on the travel booking like destination or car type which we will skip in the following). These values are stored locally in the travel agency's database. Further, the travel agency uses other services to fulfil the user's booking. After that (depending on the type of booking), either a flight is booked using an airline service or a car is booked using a car rental service or both in parallel. Finally, a billing service is called with billing address and credit card number as input parameters.

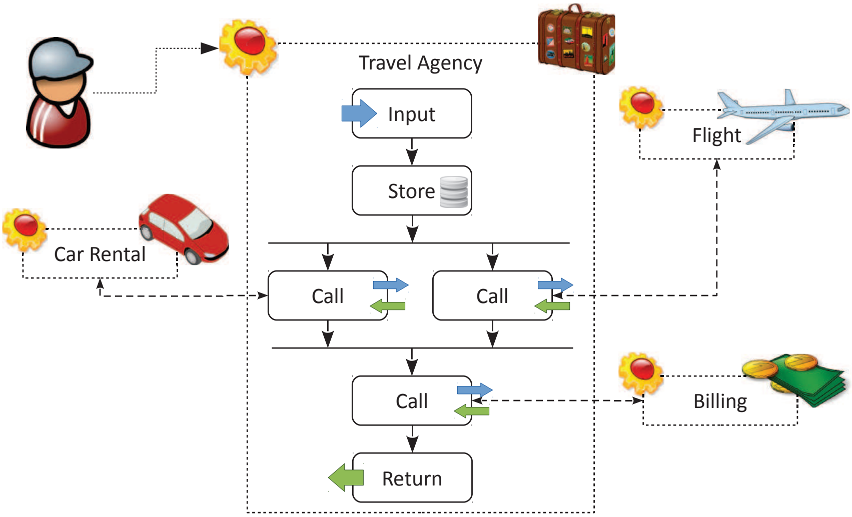


Figure 2: Process of a travel agency service

6 Generating Privacy Properties from Service Processes

6.1 Privacy Annotations

A letter of consent must include information on storage of personal data and its transfer to external institutions. In order to gather these information, the following steps inside a service process must be taken into account:

a) Which of the input parameters (i.e. data from the user) are privacy-relevant?

b) What processing steps are performed on the parameters? For what purpose?

c) Which data is stored by the service?

d) How long is the data stored?

e) Which data is used as input parameters when the service calls other services?

Item a) is obviously relevant. The letter of consent should only contain information on processing of personal information. However, it is not possible to perform this step automatically. For manual decision there are two possibilities. In the first approach, the generated privacy properties (from step 1 in Figure 1) contain processing information about all kind of data (personal and non-personal) and must be manually filtered by some legal expert when creating the letter of consent (step 2 in Figure 1).

The second possibility is to add *privacy metadata* to the service model, the source code, or other service descriptions, which holds additional information, e.g. about the criticality of parameters. Using this metadata, the generator can create more precise privacy properties. The metadata can be created by the software developer or a privacy officer together with the legal expert. This possibility has some advantages in comparison to the first one. First, the privacy relevance can be best decided when looking directly at the model or code. Second, once the persistent metadata has been added, the generation of the (precise) privacy properties can be repeated, e.g. for a new service version. Finally, this approach creates privacy properties (in contrast to the general processing properties of the first approach), which can be used for exchange of privacy properties, or for dedicated service lifecycle management. In the following, we will use this approach of metadata annotation. For the date of birth input parameter from the travel agency example before the metadata might look like this:

```
input_data {
   variable = birthdate
   type = date
   privacy_relevance true
}
```

Item b) is relevant to watch the flow of personal data during the service processing, and to decide if personal data is stored or transmitted later on. During service processing, data might be copied from one variable to another, or data might be converted (e.g. date-of-birth to age, address to geographic coordinates). For the first case, methods from the discipline of *information flow analysis* [BC85] operating on the source code can be used. However, these method are costly and too fine-grained for this application. For example, in the case of data conversion, information flow analysis would always mark the resulting data as personal. However, such converted data might not be privacy-relevant anymore, e.g. if anonymization is applied in the conversion. Thus, also for item b) we decide for adding metadata manually to variables, in order to mark if they include personal data or not. This marking is only required for data which is stored (item c) or transmitted (item e).

As mentioned before, all data that is stored by the service (in a local database, not via a remote service) is enriched with metadata in the process model. This metadata includes if the data is personal, and which original parameters have been used to calculate the information. Taking the example of a date of birth as user parameter and the resulting age as data to be stored, this might look like this:

```
stored_data {
   variable = age
   type = integer
   derived_from = birthdate
   purpose = age_verification
   privacy_relevance = true
}
```

When using such annotations, no explicit information flow analysis of the process is required (cmp. item b), just the knowledge of the developer regarding the service semantics. Further, as the user is informed about all possible data processing steps there is no need to distinguish if for a certain runtime condition the car rental or the flight booking service is called (or both in parallel). The overall privacy properties include the union of privacy properties from all process parts.

Item d) is very important, as the duration of data storage is a crucial factor of the informed consent. It is nearly impossible to derive this kind of information automatically from the service design or implementation. Thus, the only possibility is again to annotate the service. An extension to the previous example might look like this:

```
stored_data {
   ...
   storage_time = 3
   storage_time_unit = month
}
```

Item e) is the most relevant step inside a process, because the user needs to be informed if user data is not just processed inside the organisation, but also transmitted to an external partner. This information must include the relevant parameters, the original data, the parameters have been derived from, and information on the called service itself. If the called service is an internal service precise data processing information (like for the calling service) might exist and can be embedded or referenced. For example, if the privacy properties of an internal service shows, that it does not include any further service calls, the user knows, that his data does not leave the organisation. For external services this is usually not available. Thus, the service must be annotated with a least a semantic description of the service process. The letter of content at least contains the information, that the data leaves the organisation and the user can decide if she trusts the external organisation or not[1].

For the example of the travel agency service calling the (external) flight booking service the metadata might look like this (here: the country of residence is transmitted):

```
service {
   type = external
   description = Booking flights for various airlines
   privacy_info = N/A
   outgoing_data {
      variable = nationality
      type = string
      derived_from = address
      privacy_relevance = true
   }
}
```

[1] It must be noted that *internal* and *external* not necessarily reflect organisation borders. For example, internal might also be “same country” or “inside EU”.

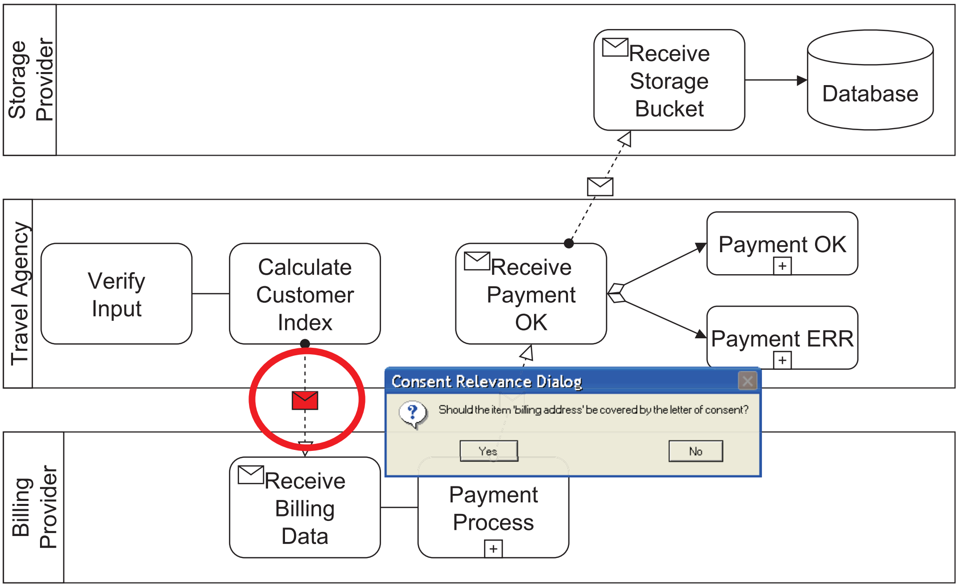
6.2 Model-Based Annotation of Privacy Properties

Most of today's services originate from software implementations that have been re-iterated and improved over the years. In that case, documentation of the service's inner workings may exist, but much of the service descriptions are purely based on existing source code documentation. For more modern, model-driven process development suites (cf. [Sof, Ecl]), however, the core elements of a service implementation are specified in a graphical modeling suite, and transformed into source code stubs—or even executable process implementations—subsequently. Such process modeling suites exist for a broad set of process modeling languages.

For the purpose of generating a letter of consent, each of these process modeling suites may be used as a basis for a dedicated *consent management extension*. Such an extension would then integrate into the modeling suite, e.g. as a dedicated consent management view (cf. [Fej09, SWF$^+$11]). On request, this extension can then ask for potential further input, and automatically generate a letter of consent document for subsequent utilization. The example in Figure 3 shows a sketch of a BPMN process model, as it could be created within any common BPMN modeling suite. By means of a dedicated extension, such a BPMN suite can then ask the model developer for input on the letter of consent. As is shown, this extension identified that a dedicated *billing address* data item is to be sent from the Travel Agency to the Billing Provider. Hence, it identifies this item as a potentially relevant privacy property of the overall process, and asks the model developer (in this case e.g. a lawyer) whether information on that data transfer is to be contained within the letter of consent or not. Similarly, by means of a sequence of dialogues, the extension can iterate over all data transfers, and also data storage points (as is shown in the Storage Provider's lane of Figure 3, represented by a database), and ask for inclusion into the letter of consent for each of these.

For subsequent re-use of the process model, it is recommended to store the lawyer's decisions along with the process model, so that each change in the process model can be aligned with a subsequent check of the set of stored privacy properties, to check whether new data transfers, storage locations, or other privacy-relevant parameters have been added. Then, the lawyer has to answer such dialogues only for the set of new and changed privacy properties, whereas all other information can be taken from the stored values. This way, the overall time to develop updates of a process (and of its letter of consent) can be minimized.

Obviously, the approach shown here for BPMN also works almost identically for other sources of process information, such as UML diagrams, WS-BPEL process descriptions, event-driven process chain models, even for semantic source code documentations, or any other sort of semantic service description.



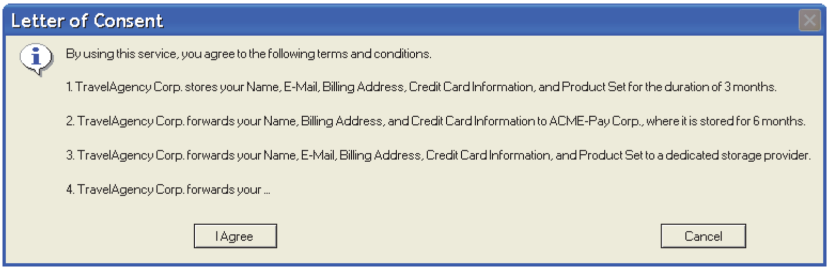
Should the item 'billing address' be covered by the letter of consent?

Figure 3: BPMN Process Model Example with Consent Relevance Dialog

6.3 Generating the Letter of Consent

Taking the data collected according to the techniques described above as a basis, it is now possible to automate (at least part of) the generation of a letter of consent document. Therefore, the deployment process iterates over the set of privacy properties, and identifies the information relevant for display towards the individual user. In the automated scenario, this task can completely be performed without human interaction, i.e. without consulting neither developers nor lawyers regarding validity of the generated letter of consent. However, it is not expected that such automation is mature enough yet to hold a thorough legal analysis. Therefore, we recommend the semi-automated setting (as described in Section 4), wherein a legal expert alters the automatically generated letter of consent stub according to specific legal requirements. Then, the user is presented with the manually altered version of the letter of consent. Nevertheless, both the automatically generated stub and the final version of the letter of consent are stored in the consent management system.

An example for how such a semi-automated letter of consent might be presented to the individual users is shown in Figure 4. As can be seen, each line in the letter of consent reflects one privacy property of the underlying process implementation. For example, the data forwarded to the external billing provider (as a separate organizational entity) is explicitly listed in the letter of consent, along with a detailed reference to the company's name (*ACME-Pay Corp.*), the exact data being forwarded (*name*, *billing address*, and *credit card information*), and the duration of storage at that external organization (*6 months*). Similarly, but more generic, the data forwarded to the external storage provider is covered by the 3rd line of the letter of consent. Here, the identity of the external or-



By using this service, you agree to the following terms and conditions.

1. TravelAgency Corp. stores your Name, E-Mail, Billing Address, Credit Card Information, and Product Set for the duration of 3 months.

2. TravelAgency Corp. forwards your Name, Billing Address, and Credit Card Information to ACME-Pay Corp., where it is stored for 6 months.

3. TravelAgency Corp. forwards your Name, E-Mail, Billing Address, Credit Card Information, and Product Set to a dedicated storage provider.

4. TravelAgency Corp. forwards your ...

Figure 4: Letter of Consent Dialog Example

ganization was hidden (*a dedicated storage provider*), and there is no information given regarding storage times. This may have resulted from a management's decision at the TravelAgency Corp. not to disclose this information publicly, which was then annotated within the consent management extension of the BPMN modeling suite. Thus, for all future versions of the letter of consent, this information will always remain hidden, unless this option becomes explicitly changed within the process modeling suite.

7 Conclusions and Future Work

In this paper, we outlined potential approaches for aligning service management with consent management. Therein, we identified the two parts of consent management, which are *User Consent Information* and *User Consent Documentation*. We discussed the link between consent management and service development, and we outlined a viable way to integrate them into state-of-the-art service modelling tools.

As many of the observations discussed in this paper require a more in-depth analysis of validity, future work obviously consists in verifying these observations for real-world use cases, e.g. in the healthcare domain. However, we are confident to find most of our observations to be accurate.

Another obvious future work consists in implementation of a consent modelling tool as outlined in Section 6.2. The plan is to extend a typical, existing, common-to-market process modelling suite with a dedicated consent management extension along the outlines given in this paper. Once such a dedicated consent management extension is prototyped, it can be tried in real-world use cases to verify our finding's validity, and to lay the ground for subsequent work in terms of optimizations and formalizations of consent management in complex services of the Future Internet.

References

[BC85] Jean-Francois Bergeretti und Bernard A. Carré. Information-flow and Data-flow Analysis of While-programs. *ACM Trans. Program. Lang. Syst.*, 7(1):37–61, Januar 1985.

[Bon13] C.J. Bonnici. An extended conceptual model of consent for information systems. In *Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on*, Seiten 149–154, June 2013.

[Con10] Joseph Conn. Govt., vendors show off consent-management tools. *Modern Healthcare, http://www.modernhealthcare.com/article/20100630/NEWS/100629932*, 2010.

[Ecl] Eclipse Foundation. Eclipse Modeling Framework. http://www.eclipse.org/modeling/emf/.

[epS] epSOS project. Technical solutions of patient consent. http://www.epsos.eu/technical-background/systems-standards/technical-solutions-of-patient-consent.html.

[Fej09] Sven Feja. An Approach for Semantic Checks of Process Models. In *BPSC*, Seiten 229–230, 2009.

[HS13] Susan D. Hosek und Susan G. Straus. *Patient Privacy, Consent, and Identity Management in Health Information Exchange*. RAND Corporation, 2013.

[JIS11] JISC Legal. Consent Management - Handling Personalisation Data Lawfully. *JISC Legal report*, 2011.

[OMG13] OMG Specification. Business Process Model and Notation (BPMN). Bericht, The Object Management Group, 2013.

[PM08] Robert A. Paton und James McCalman. *Change Management: A Guide to Effective Implementation*. SAGE Publications Ltd., 2008.

[RJB04] James Rumbaugh, Ivar Jacobson und Grady Booch. *Unified Modeling Language Reference Manual, The (2Nd Edition)*. Pearson Higher Education, 2004.

[Sof] Software AG. ARIS Business Process Platform. http://www.softwareag.com/de/products/aris/default.asp.

[SWF$^+$11] Andreas Speck, Sren Witt, Sven Feja, Aneta Lotyzc und Elke Pulvermüller. Framework for Business Process Verification. In Witold Abramowicz, Hrsg., *Business Information Systems*, Jgg. 87 of *Lecture Notes in Business Information Processing*, Seiten 50–61. Springer Berlin Heidelberg, 2011.

[Tex12] Texas Health Services Authority. Texas Consent Management Services Whitepaper. http://hietexas.org/component/docman/doc_download/387-thsa-consent-management-services-white-paper, 2012.

[The95] The European Parliament and the Council of the European Union. 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, 1995.

[The15] The European Parliament. Personal data protection: processing and free movement of data (General Data Protection Regulation, upcoming), Expected for 2015.