Privacy Assurance: Bridging the Gap Between Preference and Practice

Tariq Ehsan Elahi and Siani Pearson

Trusted Systems Laboratory, Hewlett Packard Research Labs, Filton Road, Stoke Gifford, Bristol, BS34 8QZ, UK
tariq.ee@gmail.com, siani.pearson@hp.com

Abstract. Personal identifying information is released without much control from the end user to service providers. We describe a system to scrutinize the stated claims of a service provider on safeguarding PII by interrogating their infrastructure. We attempt to empower end users by providing means to communicate their privacy concerns in a common language understood by the service provider, allowing them to set baseline privacy practices for service providers to adhere to, and providing a means of retrieving information from the service provider in the common language to base their PII release decisions.

1 Introduction

This paper will describe a system for providing privacy assurance information to end-users so that they can make an informed decision about releasing their PII to others, be they merchants, governments, or business partners. It hopes to be simple to use and deploy, give the end user more control over their PII, and be able to bridge the level of abstraction between high level privacy concerns and technical back-end implementation details.

1.1 Problems and Motivation

PII abuse can come in many shapes, like leaked credit card numbers, email addresses being sold to mailing lists, or search term histories [1]. Granted that the potential for abuse is always present the merchant can take steps to give consumers **assurance** that they can be **trusted** with private information**.**

Another compelling reason for businesses to take privacy seriously is regulations [2] and laws [3] concerning privacy of consumer records. The penalties are steep and the loss of reputation is unpalatable. Being compliant enhances the business's image with consumers since it shows awareness of privacy issues [4].

Efforts like Trust-e [5], BBBOnLine [6], and Platform for Privacy Preferences (P3P) [7] ─ amongst other privacy seal programs ─ help to provide assurance of merchants' willingness to take the issue of privacy seriously, but consumers still express dissatisfaction and want more safeguards for their Personally Identifiable Information (PII) [8,9].

The end user should also be allowed to choose how their PII should be handled [10]. To allow end user participation, unlike privacy seals which have no means of

C. Lambrinoudakis, G. Pernul, A M. Tjoa (Eds.): TrustBus 2007, LNCS 4657, pp. 65–74, 2007.

asking about the end user's choice, P3P, is an effort to give the end user some way of defining their own usage policies for their PII [11].

Unfortunately, neither of the above provide any means to interrogate the business and its processes to see if the promises being made can be fulfilled [12,13,14]. What is needed is for there to be some connection between what is stated on the privacy seal or P3P privacy policy and what really goes on within the business and its privacy capabilities [14].

This brings us to the problem that end users are not privacy experts Instead of discussing privacy at this mind-boggling level it is better to move the discussion to higher and more abstract levels where the business can express their privacy profile in terms that the end user can understand [15].

Another problem is how much information to provide. The right amount of information should be sufficient for end users' needs and also not be too much of a burden for the business in terms of volume and exposure.

1.2 Goals

What is needed is a solution that involves consumers more, is more transparent, and most of all simple [9]. We believe that a privacy assurance solution should allow communication between end-users and service providers in a common language, establish guidelines on levels of assurance information, provide mappings between privacy preferences and the back end, and above all provide trust in these mappings.

2 Our Solution

We will begin by examining how end users and businesses can communicate with each other in ways understandable to both. Then we will see how to reconcile each side's privacy concerns. Afterwards we'll look at how the high level expressions of privacy are mapped to back-end privacy technologies. Then we will consider how privacy information is provisioned on the business side which will lead us, finally, to a discussion of where and how trust fits into the solution.

It should be noted that the term “privacy policies or policy” as used in this paper is different from the typical definition used in privacy and security circles. It is usually used to define a formal means of capturing the privacy characteristics of system in terms of predicates involving rules on how to manipulate the data. In this paper the term is used to define a set of privacy preferences or practices that end users and service providers are interested in which are stated in natural language, and do not have strictly formal underlying semantics. This makes machine processing more difficult but in sections 2.1.1 and 2.2 we give an initial attempt of reconciling our privacy policies with processing systems.

2.1 Clauses: A Vocabulary for Expressing Privacy Policies

Both users and service providers will have the freedom to create policies to suit their needs. In order to bring the two together a common vocabulary is developed. This comes in the form of privacy statements or privacy clauses which are a basic primitive of our solution. A clause is succinct, clear, and unambiguous and clearly communicates

its intended purpose at a level that does not require expert knowledge of privacy systems or their implementation. It is expressed in natural language and it is hoped that both clients and services will be able to understand each other more clearly. This empowers an end-user, whom it is assumed has no technically advanced knowledge, to communicate their privacy preferences in a language they understand. In section 5 we discuss how our policies relate to previous work on policy definition.

An important aspect of clauses is that they are standardized. Since the same pool of statements are being used by both the users and service providers it is an easy matter to match up expected policies with actual ones and negotiate the mismatches. At least in this way the glaring omissions in service providers' policies will become obvious and in the same way unrealistic expectations from users will clear up. When there are deficiencies in specific clauses the totality of the policy must be looked at. The set of clauses that form the policy is a stronger indication of the suitability of a policy than the individual clauses of which it is made up. Even if there is disagreement between a user and the service provider at least both know where the other stands on privacy.

A policy is then just a collection of clauses, crafted for a particular purpose depending on the context of the interaction. For the user interacting with a bank they may invoke an “on-line banking” policy; for a service provider interacting with an online shopper they may invoke a “website customer” policy.

Templates for policies can provide a set of clauses that adhere to best practices or commonly held standards. To this a user can add or remove clauses depending on their preferences and needs. Templates are especially geared towards end users who may need help creating a privacy policy that would serve the purposes that the end user needed them for.

There is still a problem of where the clauses come from in the first place, and who provides guidance or establishes what is an appropriate policy for a particular purpose and what is not. In order to facilitate both problems it is important that there be some agreement about privacy in general and clauses and policies in particular. A way of doing this is through standardization. Trusted entities, such as governments or standardization bodies such as the W3C, who have experience in this field through efforts like P3P, can be called upon to provide a working pool of clauses and provide guidance on how to go about creating a privacy policy that is appropriate for a particular activity as a template.

We are aware that positive and negative clauses are subjective but it is hoped that through proactive efforts by privacy experts in concert with privacy groups we can arrive at a standard of privacy expectations and conduct.

2.1.1 Matching End User and Service Provider Privacy Policies Using Clauses

During a transaction where PII is to be divulged to the service provider, the end user can conduct a policy matching activity where the system can compare their privacy preferences (as stated in their privacy policy) to that of the service provider's policy.

The trivial case is when both the end user and service provider policies are identical. In this case there would be no warnings. When this is not the case then the system has two scenarios:

- *Missing Clauses*: This occurs when the service provider does not provide a clause(s) present in the end user's policy. This is flagged by the system and reported to the end user.
- *Excess Clause*: This occurs when the service provider's policy has a clause(s) not present in the end user's policy. This is not cause for alarm since all clauses are privacy positive and the additional clause will only strengthen the privacy policy.

After the matching phase the end user can make a decision on whether or not to divulge their PII. or they can then move to the next stage of the process which is validation of the clauses against capabilities of service provider's back-end systems. We talk about this in section 2.2.

2.2 Mapping and Capability Validation

Once a policy has been set by a service provider the onus is upon them to implement the measures to uphold those policies. The fact that clauses only talk about the "what" and not the "how" allows service providers flexibility in choosing the best solution for their particular infrastructure.

To tie together and bridge the "what" to the "how" there has to be some sort of mapping that facilitates this connection. The main job of this mapping is to communicate the back end privacy controls, processes, and other privacy enhancing features implemented by the service provider through the process of verification of clauses in privacy policies.

Our solution allows each clause to be composed of specific tests that query controls and system components on the back end. In this way a suite of tests can be created that inspects the system and reports back the results that can be used to verify clauses.

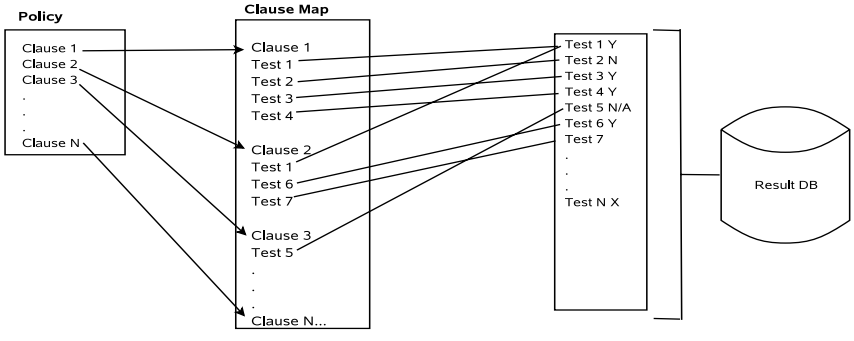

Fig. 1. Mapping clauses to the back end controls through tests

Figure 1 shows how each clause in the privacy policy is mapped to back end tests. A test only validates that the control or feature is in place and working in a known manner. There can be multiple tests on the same control to validate particular attributes, as long as they are relevant to the clause being verified.

Once the proper mapping between clauses and back end controls, via tests, has been established the service provider can now offer the end user a way to verify the

claims made on the service provider's privacy policy. This step, called capability checking, is crucial in affording assurance to the end user since it allows the user to see if the service provider is actually able to uphold their promises.

It is important to note that the knowledge of which tests are conducted is security sensitive and it would be a critical weakness against attacks because it would provide information about the nature of the systems on the back end. Therefore, detailed information is filtered out of the transmitted results to the end user. Also no information about the tests to be conducted leaves the service provider. The only information an attacker has is the privacy policy and the clauses. From that the attacker can only make inferences as to the nature of the service provider's back end. The end user does not suffer since all they require is for clauses to be fulfilled, how that is done is beyond their concern, they have the TTP to trust for that.

So far we have assumed that the correct back end controls are in place to ensure the privacy of end users' PII and only those clauses have been put into the privacy policy that are backed up by those controls. This is an obvious area of abuse and so trust has to be introduced here. In our solution trust comes in the form of third parties.

2.3 Trusted Third Parties and the Trust Chain

The missing trust has to come from entities that end users do trust such as trusted third parties (TTPs), like ISO, Trust-e and Verisign [16], or non-government consumer organizations. The way forward is to invite the TTP to scrutinize their back end systems, the mappings and their privacy policies in a compliance verification process similar to ISO 17799 and ISO 27001. If the TTP is satisfied it would issue a trust token that can be presented to the end user at the time of policy matching and verification, thus providing trust in the results and ultimately in the business.

The main concerns of the TTP are:

- Verifying that the controls and privacy enhancing technologies that are implemented by the service provider on their infrastructure are configured and functioning properly
- Verifying that the tests used to interrogate the proper configuration and function of are capturing and analyzing the correct data
- Verifying that the clause-to-test mapping is appropriate and complete
- Maintain the integrity, confidentiality, and availability of trust tokens and service provider data.

It is not the user who is responsible for validating the suitability or appropriateness of the privacy enhancing infrastructure of the service provider, but a trusted third party. The user will only be responsible for checking that third party seals are current and valid and accessing the trustworthiness of the vouching party.

In this way the end user can establish trust based on the reputation of the TTP, while the service provider can benefit from this trust relationship that has already been established, or has a better chance of growing stronger due to the fact that trust is a TTP's business and this shows the good intentions of the business to end users.

In a common usage scenario, the TTP performs its verification of the service provider's back end and how this translates to their privacy policies. It then transfers a trust token to the service provider to display along with their privacy policies as well

as with their policy validation results. The TTP will hold a copy of the model, or description, of the service provider's back-end and privacy policy for dispute resolution and as a means of recording the conditions under which the trust token was issued.

After that an end user can ask for privacy policies and/or verification results. The results and the trust token are transmitted back to the end user.

Finally, the end user must now verify that the trust token is valid and intended for this set of results and the privacy policy under scrutiny. The end user can do this via a privacy seal verification scheme, such as one described in [17]. Once the end user has checked the validity of the trust token they can then be assured that the results, whether positive or negative, are correct and worthy of trust.

Also worth noting is the fact that the TTP do not have exclusivity and that both the service provider and end user can utilize any number of TTPs. Situations can arise where no common TTPs are in use between the end user and service provider, at which point the end user can choose to add the TTP and complete or discontinue the transaction.

3 The Implemented System

Now we move on to discuss an implementation of the system described so far. This work is part of an ongoing effort funded by the EU called PRivacy and Identity Management for Europe (PRIME) [18]. This project is a multi-party endeavor with partners across Europe. As such our work is only one component of a large platform and we take for granted work being done by other partners, especially when it provides functionality we can utilize. The system presented has been fully implemented as part of the integrated prototype within PRIME, which is currently at version 2 [18].

The main functionalities provided by the Assurance Control component are to:

- **Compare** privacy polices of the service provider and the privacy preferences of the user and highlight similarities, differences and deficiencies
- **Conduct** capability tests to verify the statements made in the service policy and ensure the service side is capable of fulfilling the promises made in their policy
- **Provide** results of above in a way that allows a user to make informed decisions about releasing their private details, with some guidance built in

For a more in-depth discussion of the specific functions and how the module interacts within the PRIME frame work please refer to [18].

3.1 The End-User Experience

Presenting assurance information in a way that is simple, clutter free, and easily understood is still an area of research. To help direct our interface creation, we have worked in concert with a human computer interface team within PRIME and used their findings from usability tests conducted with end users. The preliminary findings have been published, for further details see [19]. As well, [20] discusses some general guidelines for indicators and their placement that has been incorporated as well in our GUI as well.

Most users claimed that the functionality provided by assurance control was useful and that assurance control features should appear just before release of PII [19].

Although the main purpose of PRIME is to empower individuals in protecting their privacy in customer to business (C2B) scenarios, our system is not limited to this type of usage. With the proper protocols, Businesses to business (B2B) and government to business use cases are also possible.

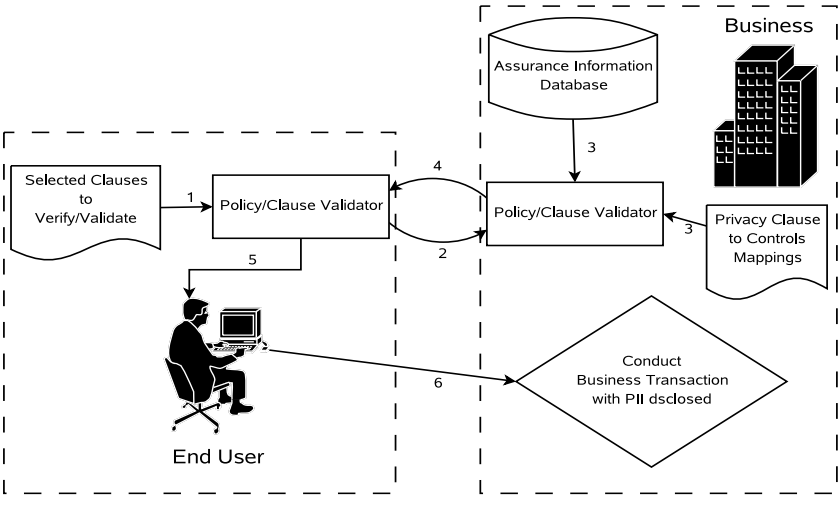

Fig. 2. Capability validation process

A simple walk through is shown in figure 2, corresponding to the following steps:

1. The user, having selecting which clauses they want verified, submits these to his or her capability checking, aka Policy Validator, module.
2. This module communicates this list to its counterpart on the service side and awaits its response.
3. The service-side Policy Validator searches for the clause to test mapping in the mapping file kept on the service side. It then queries the result database for these tests and retrieves their results. It can either aggregate the test results to a level that only verifies that the clause was fulfilled or it can send back more information. This is configurable and left up to service providers to choose how much detail they want to include in test result data.
4. The Policy Validator displays the results to the end user to allow them to make an informed decision about releasing their PII.
5. If the end user is satisfied then they can divulge their PII or if not they can provide feedback to the service provider so that it can make meet user demands in the future.

4 Related Work

There has been a great deal of work done on privacy polices [11,21,22,23,24]. In these policy frameworks the focus has been on access control based on conditional logic. Our polices are just collections or groupings of clauses that serve a particular purpose under a particular context. Since access control plays a big part in the control of PII our solution works in concert with the Access Control Decision Function (ACDF) and Identity Control (IDCTRL) in PRIME, to provide a total privacy package.

P3P is a W3C specification that allows websites and end users to specify their privacy practices and preferences respectively in a standardized way that are easy to retrieve and interpret by end users. There have been many critiques of P3P such as [12,13,21,25]. We shall focus on how our solution differs from P3P, the gaps it fills in, and how P3P could be used within the system we have implemented albeit with changes to its role.

Expressing privacy concerns in P3P is done by defining statements in a machine readable format written in XML [11]. Although there are editors [26] that help with this process, there are two problems that are not yet addressed.

First, end user must know what their privacy vulnerabilities are and how to check if a website will mitigate those risks. Most users are naïve and would not be competent enough to express privacy concerns beyond vague statements.

Second, even with the prerequisite privacy knowledge the definition of privacy polices must be in a language geared towards the facilitation of accessing PII based on conditions. This is a difficult task which our solution simplifies by introducing standardized privacy clauses and templates that are written in human readable form and are unambiguous, concise, and capture privacy concerns based on expert knowledge.

As is also the case with privacy seals, P3P can not link the privacy practices expressed by the website and anything tangible on the back-end. This gap is where our solution introduces mechanisms to check that policies and the technical realities of the website's infrastructure are coherent.

Although P3P has its limitations, its strength as a robust policy definition language and logic model allows it to translate complex privacy clauses into machine readable form. In fact, P3P's strengths could benefit our solution and could be incorporated under the clause layer as the gateway between human readable clauses and service provider result data bases and back end models.

Projects like Privacy Bird [27] from AT&T and Privacy Fox [28] try to bring a simplified and more useful solution to end user by providing a graphical face to P3P. Our solution differs in that instead of just a single aggregate representation embodied by the bird icon we opted to give a more granular output so that the end user could have more context as to exactly what went wrong.

In our solution once the end user divulges their information there is no way for them to sure that the service provider continues to adhere to their privacy practices. One way to combat this is to have a persistent service that monitors the end user's information and checks that the privacy practices are still in place. One such effort is Obligation Manager [29], which is also part of the PRIME framework. Working in concert, they can provide stronger evidence that the service provider is honoring its promises.

5 Future Directions

Since clauses are the central privacy vector they need to be developed further from the select set that are being implemented now. They need to be more complex and

recognize complex privacy needs of sophisticated users as well as laws and regulations that businesses must adhere to. They also need to be stated in such a way that is unambiguous in any language. The guidelines for TTP behavior are an open issue that requires research and reflection based on other established TTP standards and the outcome of discussions on privacy.

Presentation of privacy assurance information is an ongoing research effort in concert with the HCI team and efforts will reveal just how much trust can be conveyed between parties and identify the missing pieces in the puzzle.

At the moment the service provider depends on in-house security expertise or third party advice to implement and deploy privacy mechanisms. This dependence on security expertise could be avoided if the clauses themselves provided a set of tests that a service provider had to conduct. It could cut out the third party completely and move the reliance on to the PRIME system itself rather than third parties. The obstacles to resolving this are that service side topologies are not well understood and providing a generic yet robust enough set of tests that would be applicable everywhere is a difficult thing to do at present.

6 Conclusion

We have shown how a common standardized privacy clause pool would help communicate end user concerns as well as service provider promises. With the clauses forming policies we have designed a mapping framework that would allow high level clauses to be mapped to back end technology that would abstract the complexity away for the end user and at the same time allow the service provider flexibility in how they implement and manage their infrastructure. Finally we have shown how trust is injected into this system through trusted third parties and their role in establishing a trust chain. This allows end users to form their own trust relationships with TTPs independent of service providers depending on their preferences and experiences.

In summary, this paper reports work in progress to provide a simple and effective system for providing assurance information and building trust in privacy practices of businesses and other entities whilst being practical for deployment in current infrastructure.

Bibliography

1. AOL Search Data Scandal, http://en.wikipedia.org/wiki/AOL_search_data_scandal
2. Office of the Secretary, Standards for Privacy of Individually Identifiable Health Information, Federal Register, vol. 67–157 (August 2002), http://www.hhs.gov/ocr/ hipaa/ privrulepd.pdf
3. Data Protection Act (1998) - UK, http://www.opsi.gov.uk/ACTS/acts1998/19980029.htm
4. Kobsa, A.: Tailoring Privacy to Users' Needs. In: Bauer, M., Gmytrasiewicz, P.J., Vassileva, J. (eds.) UM 2001. LNCS (LNAI), vol. 2109, pp. 303–313. Springer, Heidelberg (2001)
5. Trust-e Privacy Seal Program, http://www.truste.org/
6. BBBOnLine Privacy Seal Program, http://www.bbbonline.org/privacy/
7. Cranor, L.F., Hogben, G., Langheinrich, M., Marchiori, M., Presler-Marshal, M., Reagle, J., Schunter, M.: The Platform for Privacy Preferences 1.1 (P3P1.1) Specification, W3C Working Draft 10 (February 2006)

8. Berlanger, F., Hiller, J.S., Smith, W.J.: Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. Journal of Strategic Information Systems 11, 245–270 (2002)
9. Shneiderman, B.: Designing Trust Into Online Experiences. Communications of the ACM 43-12, 57–59 (2000)
10. Shneiderman, B.: Designing Trust Into Online Experiences. Communications of the ACM~43-12, 57--59 (2000)
11. Leenes, R., Fischer-Hubner, S. (ed.): Prime Framework version 2, {https://www.prime-project.eu/prime_products/reports/fmwk/pub_del_D14.1.b_ec_wp14.1_V1_final.pdf}
12. Cranor, L.F.: Web Privacy with P3P, O'Reilly and Associates (2002)
13. Clarke, R.: Platform for Privacy Preferences: A Critique, http://www.anu.edu.au/ people/ Roger.Clarke/DV/P3PCrit.html
14. Ackerman, M.S.: Privacy in pervasive environments: next generation labelling protocols. In: Personal Ubiquitous Computing 2004, pp. 430–439, Springer, Heidelberg (2004)
15. Pearson, S.: Towards Automated Evaluation of Trust Constraints. In: Stølen, K., Winsborough, W.H., Martinelli, F., Massacci, F. (eds.) iTrust 2006. LNCS, vol. 3986, pp. 252–266. Springer, Heidelberg (2006)
16. PRIME principles, https://www.prime-project.eu/about/principles/
17. VeriSign Identity Protection, http://www.verisign.com/products-services/security-services/identity-protection/index.html
18. Moulinos, K., Iliadis, J., Tsoumas, V.: Towards secure sealing of privacy policies. Information Management & Computer Security 12-4, 350–361 (2004)
19. Sommer, D. (ed.): PRIME Architecture version 2, (will appear mid to late 2007) https://www.primeproject.eu/prime_products/reports/arch/
20. Petterson, J.S.: R1 - First report from the pilot study on privacy technology in the framework of consumer support infrastructure. Working Paper, Department of Information Systems and Centre for HumanIT, Karlstad University, Karlstad (December 2006)
21. Cranor, L.F.: What Do They "Indicate?": Evaluating Security and Privacy Indicators. Interactions, 45–47 (2006)
22. Hogben, G., Jackson, T., Wilikens, M.: A Fully Compliant Research Implementation of the P3P Standard for Privacy Protection: Experiences and Recommendations. In: Gollmann, D., Karjoth, G., Waidner, M. (eds.) ESORICS 2002. LNCS, vol. 2502, pp. 104–125. Springer, Heidelberg (2002)
23. Karjoth, G., Schunter, M., Waidner, M.: Privacy-enabled Services for Enterprise. In: DEXA 2002, IEEE, Los Alamitos (2002)
24. Karjoth, G., Schunter, M., Waidner, M.: Platform for Enterprise Privacy Practices: Privacy-Enabled management of Customer Data. In: Dingledine, R., Syverson, P.F. (eds.) PET 2002. LNCS, vol. 2482, pp. 69–84. Springer, Heidelberg (2003)
25. Backes, M., Pfitzmann, B., Schunter, M.: A Toolkit for Managing Enterprise Privacy Policies. In: Snekkenes, E., Gollmann, D. (eds.) ESORICS 2003. LNCS, vol. 2808, pp. 162–180. Springer, Heidelberg (2003)
26. Clarke, R.: P3P Re-visited, http://www.anu.edu.au/people/Roger.Clarke/DV/P3PRev.html
27. P3P 1.0 Implementation Report, http://www.w3.org/P3P/implementation-report.html
28. AT&T, AT&T Privacy Bird, http://www.privacybird.com
29. Arshad, F.: Privacy Fox – A JavaScript-based P3P Agent for Mozilla Firefox, http://privacyfox.mozdev.org/PaperFinal.pdf
30. Casassa Mont, M.: Dealing with privacy obligations: Important aspects and technical approaches. In: Katsikas, S.K., Lopez, J., Pernul, G. (eds.) TrustBus 2004. LNCS, vol. 3184,Springer, Heidelberg (2004)