PRIVACY-AWARENESS IN WEB APPLICATIONS AND SERVICES: A RESEARCH ROADMAP

Tania Basso[1], Regina Moraes[1], Nuno Antunes[2], Marco Vieira[2]

[1]University of Campinas (UNICAMP)
Limeira – SP – Brazil

[2]University of Coimbra (UC)
Coimbra – Portugal

{taniabasso,regina}@ft.unicamp.br, {nmsa,mvieira}@dei.uc.pt

Abstract. *Frequently users have to provide personal information for being able to use web applications and services. They are commonly confronted with a privacy policy that they must accept, implicitly trusting the provider organization to protect their privacy. The recent trend to develop frameworks for privacy policy definition has moved the state-of-the-art forward, but did not solve the main problems: allow users to express their privacy requirements and assure that these requirements will be enforced. This paper discusses the main challenges towards the development of privacy-aware web applications and services and proposes a research roadmap to tackle these challenges.*

Resumo. *Frequentemente os usuários proveem dados pessoais para que as aplicações e serviços web permitam o seu uso. Normalmente, os usuários são colocados frente a políticas de privacidade que devem ser aceitas, obrigando-os a acreditar que a organização provedora dos serviços irá proteger suas respectivas privacidades. Uma tendência no desenvolvimento de frameworks para definição de políticas de privacidade tem provocado uma evolução no estado da arte, mas ainda assim não resolveu os principais problemas: permitir que os usuários expressem seus requisitos de privacidade e assegurar que esses requisitos serão respeitados. Este trabalho discute os principais desafios para o desenvolvimento de aplicações e serviços web que tenham na privacidade um foco importante e propõe um roadmap de pesquisa para enfrentar esses desafios.*

1. Introduction

What is done with your personal information once you provide it to a website? Nowadays, web applications and services provide the users with a wide range of services, such as e-commerce and online banking. To use these services, users and customers need to provide personal information to the system they are using, but once they submit this information, they are unable to control the way it is used. This raises concerns about their privacy or, in other words, their right *"to be secure from unauthorized disclosure of sensible information that are contained in an electronic repository"* [Bertino et al. 2008]. The problem of privacy in web applications and services is multidimensional. First, the users are concerned with different types of data in different ways, ranging from data with low privacy requirements, such as address, to very critical private data like credit card numbers. Second, different people in the provider organization have different roles, as

administrators or vendors, which require different levels of access to the data. Finally, in multi-tier applications, personal information can be leaked at any application layer, ranging from storage to presentation. Figure 1 illustrates the complexity of privacy awareness in multi-tier application scenarios. Different users and third parties have distinct interaction with the system, providing and requesting private information. The computational environment is composed by several hardware and software, which should be managed to prevent attacks that can be done through web server and applications' users.

Due to regulations and in search for competitive advantages, new technologies started being developed to guarantee security and privacy of the information manipulated by web applications and services. Still, the current state of the practice consists in the presentation of a privacy policy with which the users must agree before using the service. These policies signal integrity commitment and are so important that can influence the credibility of the organization: *if the policies are clearly and explicitly stated, then the user perceives the organization as more trustworthy* [Han and Maclaurin 2002]. However, most of the times it is not possible for the users to express their own privacy preferences.

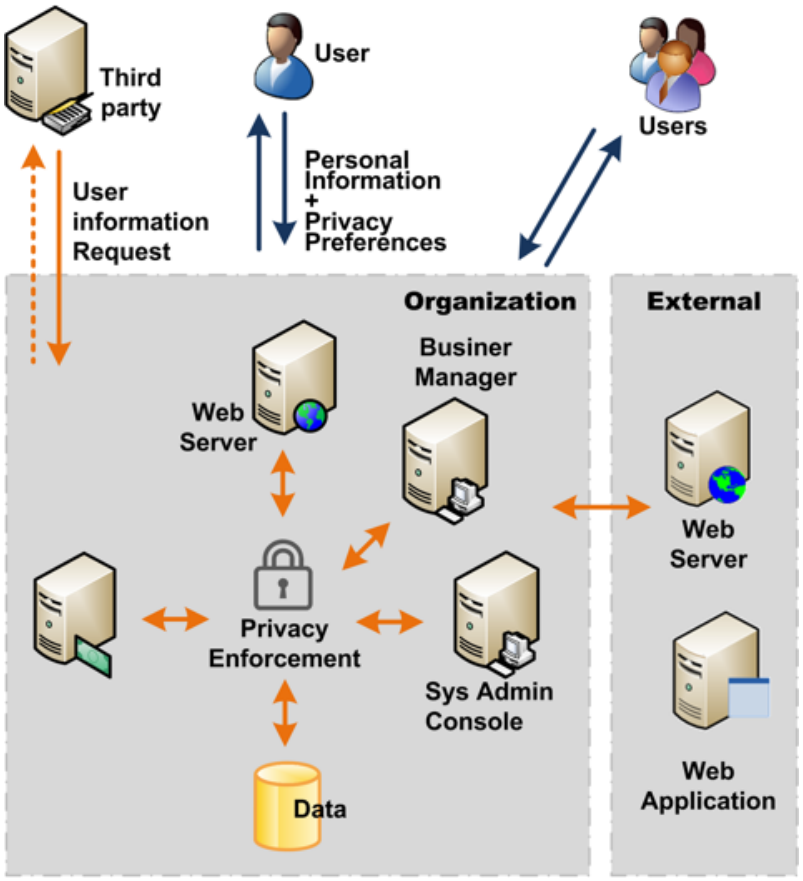

Figure 1. The complexity of privacy awareness in multi-tier application scenarios

Various privacy policy solutions try to help data owners and collectors in expressing their preferences have been proposed in the recent past, but these solutions are still far from satisfactory, and result in limited definitions that are insufficient and not customized. Such solutions also lack the mechanisms that enforce the application of the policies and

the expressed privacy preferences, a major shortcoming that might reduce the confidence of the users in the applications and thus, their willingness to use it. Web application providers need to move towards privacy-aware practices in order to increase the trust of their users and keep a competitive edge over other organizations. This evolution is only possible by adequately using the existing techniques and by developing new approaches that clearly focus the interest and individuality of the users.

The roadmap and the prioritization of activities were built on the privacy protection challenges we identified in existing privacy policies and applications. The main challenges are related to expressing and enforcing user's privacy preferences, which is discussed below.

2. On Expressing Personal Privacy Preference

Instead of the "broad audience" privacy policies generally applied nowadays, supporting custom policies that suit individual needs and preferences would benefit the users of web application and services. Several methodologies exist for describing privacy policies, but they are still limited when dealing with personal privacy preferences. We discuss the most widely spread technologies to express these preferences and their limitations, as well as the remaining challenges in this context. Also, recent achievements, which help this issue, are addressed.

2.1. Existing Technologies for Expressing Preferences

The Platform for Privacy Preferences Project[1] is a standard that allows websites to declare, in a standard format, the intended use of the information they collect about users, such as what data is collected, who can access those data and for what purposes, and for how long the data will be stored. Similarly, the Enterprise Privacy Authorization Language[2] allows enterprises to formalize their privacy practices into policies that define the categories of users and data, the actions performed on the data, the business purposes associated with the access requests, and obligations incurred on access. The P-RBAC (Privacy-Aware Role-Based Access Control) [Ni et al. 2010] extends the well-known RBAC model for organizational access control where permissions are assigned to roles and roles are assigned to users. P-RBAC incorporates notions of privacy policies, which are based on rules that need to be carried out by the organization after access is granted to the user. However, P-RBAC lacks algorithms for detecting redundancy in conditions and has limited support for relations among different permission assignments. The eXtensible Access Control Markup Language[3] standard describes both a policy language and an access control decision request/response language. The policy language is used to describe general access control requirements, while the request/response is responsible for enforcing them. Although the policy language is application independent, which makes it rather flexible, the language is limited when dealing with conflicts amongst rules and the request/response decisions are based on specific constructs that are limiting its adoption. The existing solutions are limited in two ways: (i) they are mostly centered on the priorities of the organization and less concerned with the preferences of the user. In fact, the user is not allowed to define his own policy requirements, and even the cited alternatives

[1] P3P, www.w3.org/P3P

[2] EPAL, www.zurich.ibm.com/security/enterprise-privacy/epal

[3] XACML, www.oasis-open.org/committees/xacml

present other major limitations, as we will discuss afterwards; and (ii) the definition of policies is very complex and the organizations tend to avoid using these methodologies because they are synonym of difficult and error prone tasks.

2.2. Challenges in Expressing Preferences

The use of the existing approaches is limited due to their incompleteness or because they are not scalable or portable enough. Also, such approaches are normally complex and lead to increased costs in terms of application development and maintenance. Thinking privacy in a user-centric way, giving him the tools to express his preferences, further exacerbates these issues. Access control policies can restrict unauthorized access to data and thus, protect data privacy. So, an obvious need is to allow the users to define the access control policies, or at least transpose the user preferences into policies. However, the common web users are not experts in policy definition and thus, may not be able to express correctly their preferences. Furthermore, the problem may not even be just technical (i.e. the format or the syntax to be used), but also of understanding what are the consequences of each policy defined, which leads to errors that may allow unduly access or incorrectly deny access to information. Another issue is the granularity of the data and, consequently, of the policies that need to be defined, which raises some key questions: is one policy enough to all the data inserted by the user or should he define one policy for each piece of information inserted? In a similar fashion, who is allowed to access the data? Is one policy enough for everyone or should different people on the organization behind the application have different types of access to the information? Offering usable and understandable interfaces that would allow users to express their preferences is a difficult task, especially to inexperienced users, as they are not familiar with the application's data handling practices and also with the privacy-related language used. A key challenge is how to explain to users the meaning/consequences of selecting a given privacy option. The challenge here is to provide solutions that can address the needs of users with different levels of experience. Users would also benefit from privacy recommendations, for instance based on learning processes. However, such processes require tools that accurately record individual privacy preferences in an understandable way, creating a history of privacy decisions (that may also be a subject for privacy concerns). This historical data can be used later as a basis for providing privacy recommendations to the user in new situations. Solutions in this direction could contribute to relieve users from the complex task of specifying every detail, thus simplifying and accelerating the overall process.

2.3. Relevant Achievements

We recently interviewed 22 IT professionals which, faced by a representative e-commerce scenario, reported their major concerns regarding privacy and what they would like to be able to define in order to protect the data handled by web applications. Based on the responses, we argue that using privacy policies for access control is of particular interest, but there is a large set of requirements that are not supported by the existing policy frameworks. A simple example is the definition of when the data can be accessed. The XML-based Policy Model [Basso et al. 2013] allows the easy definition of preferences in access control and data privacy protection. The model defines who can access certain information, when, from where, and how the required information can be accessed. The preferences are expressed through levels of criticality of the private information collected

by the application. Different policies can be enforced through a protection layer, as discussed later. The approach has the following strengths: a) the model provides completeness as it allows finer granular access control, combining database columns and rows; b) the policies and the mechanism are easy to be defined and based on portable technologies; and c) the model is simple, thus policies specifications can be kept under control and their integration with existing technology is quite easy. The P3P Server-Centric Architecture [Agrawal et al. 2003] proposes the use of server-centric technology for matching preferences against policies in relational databases. The policy is first flattened out into relations and then the privacy preference is converted into a SQL query over this database. If the preference matches against the policy, data is collected. This implementation works faster than a client-side implementation and is more extensible, but produces some performance impact that must be considered. In a lower granularity of preferences, the Hails framework [Giffin et al. 2012] allows users to choose which third-party applications will have access to their information. The user options are quite restricted, since he can only define the applications that have access to his information.

3. Enforcing Personal Privacy Preferences

Privacy policies are useless if they are not enforced. Several solutions provide means to protect the data from unauthorized use, enforcing the privacy preferences previously defined. In traditional multi-tier applications the personal information can leak at multiple layers, ranging from the data storage to the user interface, and policies may be enforced at any level. However, the closer to the data storage these policies are enforced, the better the data privacy is protected, while the closer to the user, the bigger are the chances of a data leak. Conversely, the closer to the data storage this enforcement is performed, the lower is the portability of the privacy protection system. This is particularly relevant if we consider implementation and maintaining costs.

3.1. Existing Technologies in Preferences Enforcement

P3P and EPAL provide standard means for privacy practices definition but neither provides mechanisms to enforce these practices. However, there are complementary mechanisms that support enforcement, as is the case of the relational database extension in [Agrawal et al. 2005], which is a server-centric architecture for matching preferences against policies at the database level, based in P3P. The purpose based access control [Byun and Li 2008] is an alternative based on P3P purposes (elements that define the intended use of data) for relational databases: accesses are granted if the requester's access purpose is among the ones previously defined by the user. In XACML, the access control request/response language used is responsible for querying a decision engine that evaluates requests against existing policies. P-RBAC considers the enforcement of policies on a per-user basis. However, in both cases, the enforcement of the policies is assumed as existing with its implementation being a responsibility of the organization. The Hails framework adds mandatory access control (MAC) and a policy language to the Model-View-Controller architecture, to ensure that calls to third-party Web applications have access to users' information without violating their privacy [Giffin et al. 2012]. Then, the MAC mechanism follows the data throughout the system enforcing policies when they pass between components with different privileges. The system supports only applications that were developed within the framework, resulting in a rather intrusive mechanism that simply cannot be generalized.

3.2. Challenges in Preferences Enforcement

The solutions based on P3P are frequently insufficient with more rules and elements being needed to describe access decisions. Also, the degree of protection provided by existing solutions is far from satisfactory and even the most complete and effective ones face several challenges. The first is the decision of where to place the enforcement components in the system, while balancing their effectiveness with the impact in performance, cost and maintainability. This decision is affected by many factors, ranging from the technologies to the types of data access inside the organization. The type of access relates to the issue of roles inside the organization. In a multi-tier application, we frequently assume (wrongly) that the web application interacting with the user is the only application using the data. However, other applications can be running inside the application server (e.g. administration application) or standalone (e.g. management console). They access data with different roles, requiring different types of access. In some cases we can even consider different roles for each application user, bringing the challenge to a whole new level. One can allow some of these applications to work around privacy enforcement, but we will be trusting blindly on the organization managers and admins. Another challenge is how to feed the enforcement system with the new data as the users input them, and at the same time, associate the policies with the correct data. This is also affected by the location selected for running the enforcement mechanisms. Due to this and other factors, solutions for enforcing privacy are difficult to implement, requiring further research in this topic. The people in charge of the business and legal departments of an organization are usually the ones that define privacy policies and this is, more often than not, done textually and in natural language. This means that policies are not defined in terms of low-level operations, but far from that. To enforce these policies using access control it is necessary to map the privacy rules into a lower level, which is not a trivial task because the mapping needs to infer the business purpose behind the access. Even assuming that this is feasible, the loss of semantics in this process can cause discrepancies from the original policy, raising legality issues in establishing equivalences between the natural language policies and the mapped/modeled ones. The Web is a prone environment for suspicions. Thus, besides enforcing the policies, it is of the utmost importance to provide guarantees to the user that his privacy preferences are being enforced, a challenge without a simple solution. Possible solutions may lie on the amount of trust the users deposit in the organization behind the application. Also, intermediary organizations, such as a certification authority or similar, may come into the play.

3.3. Relevant Achievements

To enforce the policies defined using the XML-based Policy Model [Basso et al. 2013], we developed two different mechanisms: the first, working at the application level and the second working at database level. Both have advantages and disadvantages. Our application level solution was implemented in such way that it can work (virtually) with any database management system. In practice, we instrument the data access drivers to intercept database responses and, based on the predefined policies, transparently filter the data allowing or masking (with static data) the information provided to the user. Experiments were conducted to evaluate the performance and scalability of the solution, and we found that, although the mechanism affects the system performance, it does so in a rational, close to linear fashion, which we considered acceptable. Also, the experiments

revealed that the solution is scalable both in terms of the number of users and the number of policies. On the other hand, the database level solution works inside the database guaranteeing the enforcement of the predefined policies, based on an independent set of interrelated tables that are added to the main database (these tables contain the necessary information to represent the privacy policies and user preferences). The policy enforcement is done through an access control algorithm that releases or masks the information before it leaves the database server. Once again we evaluated experimentally the scalability and performance, with the results showing that the solution is scalable in terms of the number of policies. However, we found that when the number of users grows, there is a performance price to pay, which we consider as acceptable in most cases, taking into account the data privacy protection provided. In addition to the higher performance impact, the database level solution is in disadvantage in terms of maintenance when it is necessary to change the DBMS. On the other hand, its advantage lies on filtering the data directly in the database, protecting against possible attacks to the web application or to the network.

4. A preliminary integrated solution for expressing and enforcing preferences

Overcoming these challenges of expressing and enforcing privacy preferences and thus improving privacy protection requires disruptive approaches. As a base step we proposed the development of a reference model [Basso et al. 2018] that provides a better understanding of the privacy domain and, consequently, facilitates research, modeling and development of privacy-aware technology. For sake of completeness, Figure 2 outlines the proposed reference model, which is composed by a privacy conceptual model, an UML profile, and a reference architecture. The privacy conceptual model [Basso et al. 2018] was composed by elements that represent privacy concepts and their relations, in an organized way. The goal was to specify how applications should handle privacy. In practice, the model represents the privacy policies and their statements, as well as the related services and the resources to be used for enforcing these statements. Also, it includes the use of privacy preferences, where users can agree or not with the statements. From the conceptual model we created the UML profile [Basso et al. 2015a], which allowed extending UML language to incorporate privacy concepts. The profile is useful to describe the privacy policies and how to enforce them, respecting the user's preferences. The description is done through UML diagrams that support the development process of privacy-aware applications and services. Also based on the conceptual model we defined a reference architecture [Basso et al. 2015b], which described the features and functionalities that must be addressed during the development to protect the privacy of the users. The goal is to derive concrete architecture models that can help better understanding of the privacy domain and, consequently, facilitate development of privacy-aware technology.

5. A Roadmap for Future Research

After reviewing the privacy-enhancing technologies that exist nowadays, we can highlight many challenges that keep us apart from a privacy-aware web environment that privileges the user's preferences and define the lines that must guide future research. Figures 3 to 9 present these challenges. The figures are organized in classes (labeled C in the first column) and, for each class we list the research questions that need to be answered. Also, for each topic we list a set of steps that will lead to answering one or more of the listed questions. Finally, we define a priority level (P) for each step level (High, Medium or

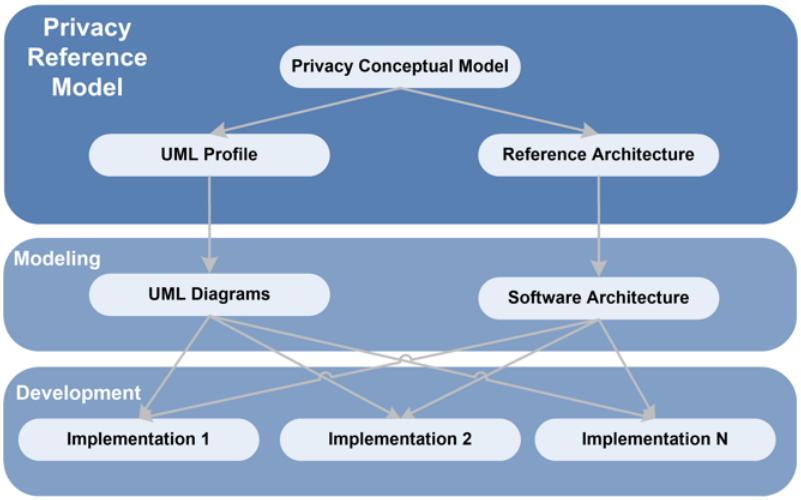

Figure 2. The privacy reference model and its application [Basso et al. 2018]

Low). We present, in Figure 3, the research challenges that deal with the user preferences, allowing users to express themselves. For example, to store and manipulate preferences, a tool that allow different actors to model their privacy preferences should be considered.

C	Research Questions	Research Steps	P
User preferences	*How to design the interface to collect user's privacy preferences?*	⇒ Define interfaces and policies considering that users have different levels of experience.	H
	How to explain the consequence of each option among existing preferences?	⇒ Define interfaces and policies considering that data has different levels of granularity.	H
	How to avoid that reading or defining complex privacy policies is skipped?	⇒ Research ways for explaining the impact of the decisions about the different privacy preferences users can choose.	M
	How to improve the process of collecting users preferences?	⇒ Create tools that accurately record individual privacy preferences in an understandable way.	H
	After the collection, how to deal with the preferences i.e., how to storage and manipulate these information?	⇒ Study the applicability of Pattern Recognition or other techniques to provide privacy recommendations.	M
		⇒ Study the use of privacy recommendations to easing the filling and avoid skipping this step.	H
	How to include new access control policies in the system without increasing the intrusiveness?	⇒ Develop tools that allow different actors to model their privacy preferences.	M
		⇒ Research non-intrusive techniques to add the new policies timely and with reduced perturbation of the web application functionalities (e.g. new data must be the minimum time unavailable because of the policies update).	L

Figure 3. Questions and steps to guide research towards privacy regarding User Preferences

We also present the management & outsourcing questions in Figure 4, which deal with the increasing use of third-party services. One important issue in this case is to be sure that practices will be enforced, which will increase the trustworthiness of the web application.

The language questions, in Figure 5, can contribute for the uniformization of policy definition and also help to the previous questions. Related to language the capacity to associate semantics and reduce subjectivity is very important in the writing of policies.

Afterwards, in Figure 6, we present the challenges related to the development

C	Research Questions	Research Steps	P
Management & Outsourcing	*How to guarantee that outsourced services apply compatible practices?* *How to be sure that the both parties (application requester and third-party service) will enforce their practices?*	⇒ Define ways for managing privacy policies in order to identify compatibility, especially at high level, to avoid loss of semantics. ⇒ Study the applicability of text mining resources for identifying similar text patterns in the application requester and third-party services. ⇒ Define processes to deal with policies incompatibilities as, for example, policy negotiation processes which can assure that the application requester have the priority in its privacy policy enforcement, i.e., the priority is the protection of its users information while using a third-party service. ⇒ Study techniques to measure the trustworthiness of web applications dealing with sensitive data (see ***Trustworthiness***).	M L M H

Figure 4. Questions and steps to guide research towards privacy regarding Management & Outsourcing

C	Research Questions	Research Steps	P
Language	*How to deal with privacy statements in machine-readable format, with the minimum of loss of semantics?* *How can the organizations write policies in a semi-formal and non-subjective fashion?*	⇒ Define natural language patterns that allow writing policy statements in more structured and less subjective way. ⇒ Define languages and formalisms that can better represent the policy high-level statements. ⇒ Define ways to associate semantics and reduce subjectivity.	M M H

Figure 5. Questions and steps to guide research towards privacy regarding Language

teams, which need tools for modeling and documenting the privacy of the system and also implementations to accelerate the development of applications with enforcement mechanisms, which is presented in Figure 7.

C	Research Questions	Research Steps	P
Modeling	*How to document privacy aspects of web applications and services?* *How to model different privacy views of the system?* *How to guide the development, standardization, and evolution of systems in terms of privacy?*	⇒ Extend existing modeling languages to better describe policy statements and the resources or technologies to enforce them. ⇒ Extend or create new reference models, architectures, frameworks and norms for privacy domain. ⇒ Extend or define privacy-related terminologies and taxonomies.	M M M

Figure 6. Questions and steps to guide research towards privacy regarding Modeling

Finally, we present the challenges to measure and improve the trustworthiness of the organization in Figure 8 and, in Figure 9, also the ones related to violation monitoring.

6. Extending the Previous Solution toward a Trustworthiness Benchmark

Although an integrated solution was emerged, it is far from complete. Since the first version was published, new technologies have arisen, as well as new researches were finished and published. More recent research works have been surveying and the findings have been compared and updating the existent reference model. Based on new privacy

C	Research Questions	Research Steps	P
Implementations	*How to enforce the predefined privacy policies?* *How to respect the users' preferences while enforcing the privacy policies?*	⇒ Study how privacy enforcement can co-habit with security enforcement techniques. ⇒ Develop enforcement mechanisms focused on the privacy of the data and different users preferences. ⇒ Create specific privacy-related domain enforcement mechanisms.	M H H

Figure 7. Questions and steps to guide research towards privacy regarding Implementation

C	Research Questions	Research Steps	P
Violation Monitoring	*How to detect privacy violation in a complete way?* *How to deal with potential false positives, which can lead to denial of service?*	⇒ Conduct studies to identify the more recent sources of privacy violation. ⇒ Develop tools that act in identifying privacy violation across the different sources. ⇒ Adopt auditing processes and tools to identify the source of violation (privacy policies can be mapped to auditing specifications). ⇒ Study improvements to existing monitoring and intrusion detection systems that can protect data privacy while dealing with multiple granularity and criticality of data.	M H H M

Figure 8. Questions and steps to guide research towards privacy regarding Violating Monitoring

C	Research Questions	Research Steps	P
Trustworthiness	*How to provide guarantees to the user that the privacy policy and his privacy preferences are being enforced?*	⇒ Research techniques to understand and measure the trustworthiness of web applications dealing with private data in order to establish trust relationships. ⇒ Study the applicability of digital certifications and certifications authorities. ⇒ Provide ways to measure trustworthiness of web applications. ⇒ Study the applicability of entities to state the trustworthiness of organizations. ⇒ Adopt privacy violation detection solutions (see below). ⇒ Extend security-related solutions to proceed towards a privacy-solution.	H M H M L M

Figure 9. Questions and steps to guide research towards privacy regarding Trustworthiness

policies, the reference model is acquiring new statements and enforcement elements. In consequence, the UML Profile is being updating, reflecting these novelties. Regarding the Reference Architecture, it will be investigated whether if needs some extensions to contemplate the changes performed in the reference model. These researches help to address the *Modeling* class of the roadmap, considering some steps from *User preferences* class.

Privacy protection is another concern that has been highlighted particularly in the legal scope of privacy. For example, General Data Protection Regulation (GDPR) indicates anonymization techniques (data suppression, generalization, masking, etc.) as a way to protect the privacy of individuals. GDPR is a regulation by which the European Union (EU) intends to strengthen and unify data protection for all its individuals, as well as addressing the export of personal data outside

the EU. It was approved (in 2016) and its enforcement is scheduled to May 2018 [EU General Data Protection Regulation (GDPR) 2017]. Our research group is investigating, as a means of privacy enforcement, how to better apply data anonymization techniques, mainly to deal with the trade-off between anonymization and data utility to mining Big Data (the more anonymized the data, the less utility to mining it has). More than that, we are concerned about the risk of re-identification (i.e., even if data is anonymized, an adversary is able to identify an individual using published data), researching anonymization models to improve data protection. Preliminary tests were performed to evaluate the impact of anonymization on the performance of data mining classifiers. In general, for these experiments, the classifiers presented small variations for different anonymization stages, i.e., the application of anonymization techniques did not cause relevant impacts on the performance of the classifiers. These solutions would address some steps from the *Implementation* class of the roadmap.

Regarding the *Trustworthiness* class, privacy measurement is the focus of more deeply study. Currently, we are studying how to place privacy concerns in a privacy quality model, defining metrics and computing scores to allow the benchmark of different systems, firstly focused on privacy and then composing with other properties to benchmark systems trustworthiness. The goal is to develop a platform to execute tests, monitoring the systems, collect trustworthiness measures and provide the trustworthiness assessment. Next step is to benchmark systems so that they can be classified as more or less trustworthy.

7. Final Considerations

Large steps must be taken for a privacy-aware information-technology society. It is frequent that the broad attention to security prevents us from looking at privacy in detail, which is particularly dangerous because many times privacy has priorities that are contradictory to other of the security properties, like availability. These issues will increase even further as the target audience of web applications broadens to virtually everyone, and such applications tend to use more and more of personal data.

Privacy awareness is much about the users, their duty of understanding the importance of protecting their personal data, and their right to it. The path towards it must be paved with innovative techniques and tools, but mainly with a rethinking of the privacy processes and priorities, that will lead to user-centric practices and more trustworthy applications. A trustworthiness benchmark could help in this direction, once it would allow users to assess and compare systems or applications according to specific privacy characteristics, and, also, select the ones that offer more guarantees in terms of privacy protection.

Acknowledgment

This work has been partially supported by the project **EUBra-BIGSEA**[4], funded by the Brazilian Ministry of Science, Technology and Innovation (Project 23614 - MCTI/RNP 3rd Coordinated Call) and by the European Commission under the Cooperation Programme, Horizon 2020 grant agreement no 690116.

[4]http://www.eubra-bigsea.eu

Also, it is supported by the project **ATMOSPHERE**[5], funded by the Brazilian Ministry of Science, Technology and Innovation (Project 51119 - MCTI/RNP 4th Coordinated Call) and by the European Commission under the Cooperation Programme, Horizon 2020 grant agreement no `777154`.

References

[Agrawal et al. 2005] Agrawal, R., Bird, P., Grandison, T., Kiernan, J., Logan, S., and Rjaibi, W. (2005). Extending relational database systems to automatically enforce privacy policies. In *Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on*, pages 1013–1022. IEEE.

[Agrawal et al. 2003] Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. (2003). Implementing p3p using database technology. In *Data Engineering, 2003. Proceedings. 19th International Conference on*, pages 595–606. IEEE.

[Basso et al. 2013] Basso, T., Antunes, N., Moraes, R., and Vieira, M. (2013). An xml-based policy model for access control in web applications. In *International Conference on Database and Expert Systems Applications*, pages 274–288. Springer.

[Basso et al. 2015a] Basso, T., Montecchi, L., Moraes, R., Jino, M., and Bondavalli, A. (2015a). Towards a uml profile for privacy-aware applications. In *Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), 2015 IEEE International Conference on*, pages 371–378. IEEE.

[Basso et al. 2018] Basso, T., Montecchi, L., Moraes, R., Jino, M., and Bondavalli, A. (2018). Privapp: An integrated approach for the design of privacy-aware applications. *Software: Practice and Experience*, 48(3):499–527.

[Basso et al. 2015b] Basso, T., Moraes, R., Jino, M., and Vieira, M. (2015b). Requirements, design and evaluation of a privacy reference architecture for web applications and services. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pages 1425–1432. ACM.

[Bertino et al. 2008] Bertino, E., Lin, D., and Jiang, W. (2008). A survey of quantification of privacy preserving data mining algorithms. In *Privacy-preserving data mining*, pages 183–205. Springer.

[Byun and Li 2008] Byun, J.-W. and Li, N. (2008). Purpose based access control for privacy protection in relational database systems. *The VLDB JournalâThe International Journal on Very Large Data Bases*, 17(4):603–619.

[EU General Data Protection Regulation (GDPR) 2017] EU General Data Protection Regulation (GDPR) (2017). Gdpr portal: Site overview. `http://www.eugdpr.org/`.

[Giffin et al. 2012] Giffin, D. B., Levy, A., Stefan, D., Terei, D., Mazières, D., Mitchell, J. C., and Russo, A. (2012). Hails: Protecting data privacy in untrusted web applications. In *OSDI*, pages 47–60.

[5] https://www.atmosphere-eubrazil.eu/

[Han and Maclaurin 2002] Han, P. and Maclaurin, A. (2002). Do consumers really care about online privacy? *Marketing Management*, 11(1):35.

[Ni et al. 2010] Ni, Q., Bertino, E., Lobo, J., Brodie, C., Karat, C.-M., Karat, J., and Trombeta, A. (2010). Privacy-aware role-based access control. *ACM Transactions on Information and System Security (TISSEC)*, 13(3):24.