A Holistic Approach for Privacy Protection in E-Government

Konstantinos Angelopoulos
University of Brighton
Brighton, UK
K.Angelopoulos@brighton.ac.uk

Vasiliki Diamantopoulou
University of Brighton
Brighton, UK
V.Diamantopoulou@brighton.ac.uk

Haralambos Mouratidis
University of Brighton
Brighton, UK
H.Mouratidis@brighton.ac.uk

Michalis Pavlidis
University of Brighton
Brighton, UK
M.Pavlidis@brighton.ac.uk

Mattia Salnitri
University of Trento
Trento, Italy
mattia.salnitri@unitn.it

Paolo Giorgini
University of Trento
Trento, Italy
paolo.giorgini@unitn.it

José F. Ruiz
Atos
Madrid, Spain
jose.ruizr@atos.net

ABSTRACT

Improving e-government services by using data more effectively is a major focus globally. It requires Public Administrations to be transparent, accountable and provide trustworthy services that improve citizen confidence. However, despite all the technological advantages on developing such services and analysing security and privacy concerns, the literature does not provide evidence of frameworks and platforms that enable privacy analysis, from multiple perspectives, and take into account citizens' needs with regards to transparency and usage of citizens information. This paper presents the VisiOn (Visual Privacy Management in User Centric Open Requirements) platform, an outcome of a H2020 European Project. Our objective is to enable Public Administrations to analyse privacy and security from different perspectives, including requirements, threats, trust and law compliance. Finally, our platform-supported approach introduces the concept of Privacy Level Agreement (PLA) which allows Public Administrations to customise their privacy policies based on the privacy preferences of each citizen.

CCS CONCEPTS

•**Security and privacy** →*Domain-specific security and privacy architectures;*

KEYWORDS

Privacy by Design, Privacy Level Agreement, Privacy Requirements, Privacy Enforcement

1 INTRODUCTION

An increasing number of government operations take advantage of new technological advances [17, 26] (e.g., Cloud and Big Data), moving toward e-government. This direction has provided new challenges for software engineers associated with information and data privacy management, technological complexity and restrictive laws and regulations [1]. From a societal perspective, citizens' lack trust for such services and their perception on how Public Administrations (PAs) store and deal with their data along with the lack of transparency, are a bottleneck to the wide adoption of e-government [10]. On the other hand, from a technical perspective, in e-government multiple organisations might require to process citizens' private data differently. This rises a major concern about the accountability of the PAs involved.

Existing privacy engineering frameworks, platforms and models [15, 16, 18, 28] do not support analysis of privacy issues from different perspectives (e.g., organisational, business-process, threat and mitigation), nor they allow public administration authorities to take into account citizens' needs in order to make their services transparent. Moreover, they fail to combine such analyses with trust analysis in order to better understand how trust influences the citizen needs and how it impacts potential privacy threats and mitigation strategies. A higher level of trust is very likely to increase the adoption of e-government by the society [3, 5].

This paper proposes a holistic, platform-supported approach for privacy protection in e-government that provides solutions to both the societal and the technical challenges discussed above. In particular, it contributes to improving privacy in e-government services through: a) the enhancement of user trust and confidence in e-government services, by combining existing software engineering approaches and modelling languages, in order to analyse trust relationships between citizens and PAs, and identify ways of strengthening such relationships, which could result in decreasing the number of users that are reluctant to use such services; b) the improvement of transparency, by imposing accountability to service providers (e.g., public authorities) with regard to privacy of citizen information; c) the empowerment of users (i.e. citizens and

ARES '17, Reggio Calabria, Italy
© 2017 ACM. 978-1-4503-5257-4/17/08…$15.00
DOI: 10.1145/3098954.3098960

Public Administrators), by providing a new type of Privacy Management system that allows them to take control over their data; d) the construction of personalised privacy agreements between governments and citizens, based on the individual preferences of the latter (e.g., choose which third party organisations will have access to their personal data) and through the provision of a privacy policy enforcement mechanism to guarantee that these agreements are respected.

The paper is structured as follows; Section 2 overviews the baseline of our work. Next, Section 3 presents the VisiOn Privacy Platform. Section 4 evaluates the platform with a real case scenario. Section 5 discusses related work while Section 6 concludes the paper.

2 BASELINE

In this section we present a set of research tools that constitute the baseline of our proposal. These tools support privacy analysis of socio-technical systems [7]. These tools allow capturing requirements for PA systems, systematically creating privacy policies and enforcing them.

2.1 Privacy by Design

Privacy by Design (PbD) is an approach to protecting privacy by embedding it into the design specifications of technologies, business practices, and physical infrastructures. PbD is currently included in the revised draft regulation for data protection in the EU [21], referred to as 'Data Protection by Design' in order to increase incentives to implement PbD for both suppliers of systems that process personal data and for government organisations that procure such systems. The following modelling tools adopt the concept of PbD by providing the means to systematically elicit and document privacy, security and trust requirements.

STS-Tool This tool supports the Socio-Technical Security Modelling Language (STS-ml) [8]. STS-ml allows modelling privacy and security aspects from social and organisational perspectives of PA systems. STS-ml enables its users to perform the following activities: a) capture the source and destination entities in a transmission of citizens' information and thus enhancing the transparency of the e-government services; b) identify privacy requirements and potential conflicts with social and organisational aspects of the system; c) specify who is authorised to access information of citizens. For example, STS-ml can be used to specify confidentiality requirements on medical record of patient's hospital, to specify how such information is transmitted and check whether the confidentiality is preserved or there are security leaks in the systems.

SecBPMN2 This tool uses an extension of BPMN 2.0 [20], the standard modelling notation for business processes, with security and privacy concepts. SecBPMN2 [24] allows to: a) specify secure business processes, i.e. business processes with annotated security aspects; b) define privacy policies, i.e. procedural constraints with privacy properties; c) verify privacy policies against secure business processes; d) verify if business processes are compliant with citizen requirements; e) generate a report that describes the secure business processes and the privacy policies specified by PAs. For example, SecBPMN2 allows PA administrators to specify business processes, implemented within a hospital, and capture how to store and distribute medical records of patients. Therefore, SecBPMN2 allows to verify that such processes do not violate security constraints, such as the restriction of distribution of medical record only to authorised doctors.

SecTro This tool supports the Secure Tropos [19], a security-aware software systems development methodology that shares several concepts with STS-ml on modelling socio-technical systems. SecTro allows the detection of security and privacy threats that could prevent the modelled system from fulfilling its goals or compromise the privacy of the data that handles. Moreover, SecTro allows the detection of vulnerabilities and the selection of security and privacy mechanisms, through a pattern library, to protect them. This library also guides the user on which pattern is most suitable, based on the threat that is addressed. The main purpose of this tool is to explicitly capture the security requirements of the analysed system and to facilitate the selection of a suitable security mechanism in order to mitigate potential attacks.

JTrust This tool supports a methodology for modelling and reasoning about trust relationships and for assessing the trustworthiness of a system under development. JTrust [22] enables PAs to model the privacy related trust relationships that affect the trustworthiness of their systems in terms of ensuring privacy. The tool also allows reasoning about these trust relationships in a structured way, facilitating the identification of technical or organisational controls in cases where there are gaps of trust, in order to ensure that data privacy is preserved. In the context of e-government services, capturing the level of trustworthiness between citizens and other organisations, contributes to eliciting questions about permitting of refusing access of their data to the latter.

CARiSMA The CompliAnce, Risk, and Security Model Analyzer (CARiSMA) tool allows modelling system architectures using UMLsec [13] diagrams. UMLsec is an extension of UML [4] in form of a UML profile that provides model driven development for secure information systems. The tool supports and implements UMLsec checks on compliance, risk or security. Tags and stereotypes are used to express security requirements and assumptions on system environments. CARiSMA is used for analysing the enforcement of security constraints at architectural level, as well as the enforcement of security requirements of citizens in the PA system. In particular, the latter will increase the security level of the system and therefore, in the long term, the trust of the citizens in the application.

2.2 Privacy Policies

Every PA, when dealing with data collected from the citizens, is required to apply certain privacy policies that must be compliant with the existing laws and guarantee that they will be respected. These policies are usually a result of explicitly stated preferences of the citizens provided through questionnaires. Hereby, we present two tools that facilitate the composition of questionnaires for eliciting privacy preferences and provide guidance to the citizens in answering them.

DAE The Dynamic Audit Engine (DAE) supports the elicitation of citizens' privacy needs. DAE allows PAs to easily create questionnaires that are used by the citizens in order to provide their preferences about privacy of their data (e.g., who can access, for which service, for how long, if the data can be shared, etc.). PAs use this tool for either creating a single questionnaire for their system or multiple ones that refer only to specific services they provide. This way they can link them to citizens according to the services they wish to use. The questionnaires can also be updated with new or additional information so citizens are always up to date with the new requirements of the PAs or express their preferences if new laws or policies are applied in the PA (e.g., new European law about data protection that mandates definition of a specific confirmation or definition by the user). The answered questionnaires form the privacy policies must be followed by the PAs.

DVT The Data Value Tool (DVT) uses simple questionnaires to capture the importance of citizens data and compares this with both the PAs' expectations and citizens' perspective. This tool calculates metrics, based on the answers of the questionnaires, and visualises the results to the users. This tool promotes the citizens' awareness about privacy, since it communicates to the citizens the relative value of their data and, consequently, increases the trust of the citizens to the PA.

2.3 Privacy Enforcement and Law Compliance

In this section we present a set of tools that support run time and monitoring.

LIONoso The machine Learning and Intelligent OptimizatioN (LIONoso) tool [2] performs data analytics that focus on history based assessment and law compliance. The former consists of an analysis of the authorisations request to access citizen's data and the generation of a prediction of the possible outcomes of subsequent requests. The latter part consists of a web-based component which permits: a) specification of how PAs use/manage citizen data; b) specification of constraints imposed by regulations and laws; c) verification of the conformance of the data management specified by the PAs with the constraints specified in laws and regulations. The main purpose of this tool is to guarantee to the citizens that a PA is compliant with the law.

PAE The Privacy Agreement Enforcer (PAE) tool [11] focuses on the protection of privacy of data by providing a policy and attribute-based access control functionality that is able to evaluate permissions for accessing confidential and private data. The protection of the privacy of the citizens' data is conducted according to privacy policies, which are created automatically by PAE, using as basis the privacy preferences defined by the citizens. Furthermore, this tool allows to automatically update and modify privacy preferences from computer-medium level (formatted) to computer-low level (policies). Additionally, the tool is able to evaluate requests for accessing private data against these policies, checking the different policies that apply to that specific data and enforce the result. The purpose of this tool in the context of e-government services is to secure and protect the access to the citizens' data by using as input the information of their privacy preferences and translating it to low level policies. Finally, PAE promotes accountability in e-government by processing and recording every access to the citizens' data and, if necessary, provides information on rejected data requests.

MANE The Media Aware Network Element (MANE) tool is responsible for monitoring and filtering the network traffic. It acts as an extra layer of data protection by applying access rules according to the data received from PAE. MANE provides additional protection of the privacy of citizens' data in the system as it monitors data exchange between a system storing sensitive data and a system that requests these data. MANE monitors exchanged packages by analysing, among other information, the data and the requester and then checks if the requester is in the white list of allowed accesses. If not then it prevents the host system from sending any data and registers and attempt of unauthorised access. The purpose of this tool is not only to control data requests to a system but also to monitor all data traffic in order to detect unauthorised transmissions. The control of the accesses is obtained from PAE, which, as we described previously, generates the privacy policies of the protected data. The automatic and continuous communication between these two tools guarantees that the data protection mechanism is always up to date.

3 THE VISION PRIVACY PLATFORM

In this section, we present the VisiOn Privacy Platform (VPP) which is designed to enable PAs, legal advisors, software and privacy engineers and domain experts to elicit privacy preferences from the citizens, identify privacy risks in the system-to-be and eventually propose countermeasures. Furthermore, our platform guides citizens to specify their privacy preferences and to increase their awareness about their personal data value. VPP also includes automated privacy protection mechanisms to guarantee that no personal information will leak by human error or malicious intention and therefore strengthening citizens' trust in e-government services.

3.1 VisiOn Architecture

The architecture of VPP, as shown in Figure 1, is composed of four major components namely, the Desktop Framework, the Web Framework, the VisiOn Database (VDB) and the Visualisation tool (ViTo).

The Desktop Framework is composed of the STS-Tool, SecBPMN2 SecTro, JTrust and CARiSMA. These tools, as described in the previous section, allow PAs to capture the privacy and security requirements for their systems in the diagrams of their respective modelling languages. These diagrams are stored in the VDB as well as additional information that is used to guide the privacy preferences elicitation from the citizens and the privacy policies that PAs will apply.

The Web Framework consists of LIONoso, PAE, MANE, DAE and DVT and offers four main functionalities: i) assists PAs to create privacy policies; ii) ensures that these policies conform to existing laws; iii) allows citizens to state and update their privacy preferences; iv) ensure that the privacy policies are respected. The main output of this framework is the Privacy Level Agreement (PLA), a bilateral contract between a citizen and a PA, which states how the latter shall handle the data of the first one, based on the provided privacy preferences and the guarantees offered by the system on security aspects. The PLA, an example of it is depicted in Figure 2, embodies the privacy policy that must be applied for each citizen.

For creating a PLA, the PA administrators provide an initial set of questions as input to the Web Framework which later on is enriched with metadata that support the automatic processing of

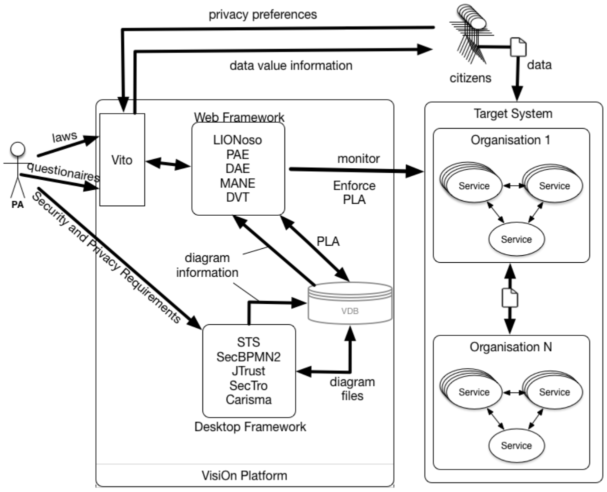


Figure 1: The VisiOn Privacy Platform Architecture

questions and their answers. These metadata provide information for the origin and the value of the data related to a question, who is responsible for handling, what operations is allowed to perform with these data, etc. The metadata generation is operated by the diagram information coming from the tools of the Desktop Framework and the DVT tool. Next, the citizens are requested to answer this questionnaire to state their privacy preferences. The answers are exported along with the questions and their metadata to the VDB in two different formats. One format contains questions, answers and metadata in a machine-readable document, the other format is a typeset textual representation of the filled questionnaire, excluding the metadata and intended to be displayed in the PLA document.

The PLA is populated with security and privacy reports in order to demonstrate the compliance level of the PA privacy policies with privacy laws and increase citizens' awareness on data valuation. Towards this direction, CARiSMA performs security and privacy checks on the PA systems and generates reports with the results, which will be contained in the PLA of each citizen. Then, LIONoso is responsible to check the compliance of the information treatment declaration of the PA system with EU and the PAs' country's privacy laws, in order to assess citizens' privacy requirements coverage, based on both historical values and monitoring results. Finally, DVT provides an indication to citizens regarding their perception of the value of their data, using enhanced visualisation elements.

At runtime, the purpose of the Web Framework and in particular PAE, MANE and LIONoso, is twofold. First, it allows to monitor events and traffic within the PA's system in order to provide to the citizens and the PAs the means of controlling who is requesting the data and to ensure that the privacy preferences set by the citizen are being fulfilled. Second, it enables the evaluation of requests, based on citizens' privacy preferences. The main goal is to ensure that the privacy preferences of the citizens control the accesses to their data. Therefore, these preferences will be taken into account by the platform for evaluating every received request of their data.

The Visualisation Tool (ViTo) was created specifically for VPP. The purpose of this tool is to provide a web interface for the tools of the Web Framework and, in the back-end, to generate the PLA documents for each citizen, which are stored in and retrieved from the VDB. The interface provided by ViTo allows PAs to submit and refine their questionnaire. Moreover, it allows citizens to answer the questionnaire and display information about the value of the data the questions refer to. ViTo is also responsible to download on demand the PLA when requested by a citizen. Such software is essential for the VPP since it eases the access of citizens to their data, increasing the transparency of a PA system. The generation of the PLA is central for increasing the trust the citizens have in the PAs.

3.2 The VisiOn Privacy Platform Process

VPP, when applied in a PA system, is operated in three phases. The first two phases, namely Requirements Specification and PLA Generation are executed at design time, whereas the third, named PLA enforcement is executed at runtime. Below the steps of each phase are described.

Requirements Specification Phase. In the first of the three phases the PA administrator along with the IT experts of the organisation use the Desktop Framework tools in order to capture the requirements and the structure of the system-to-be, by performing the following steps:

Step 1: The PA administrator uses the STS-tool to perform a privacy analysis of their system. More specifically, the PA graphically represents the organisational structure of the modelled system, i.e. which entities participate, what are their goals and how they interact with each other. The model includes privacy requirements (e.g., if a document is confidential or not) that are associated with the transmission of citizens' information among various entities within the system.

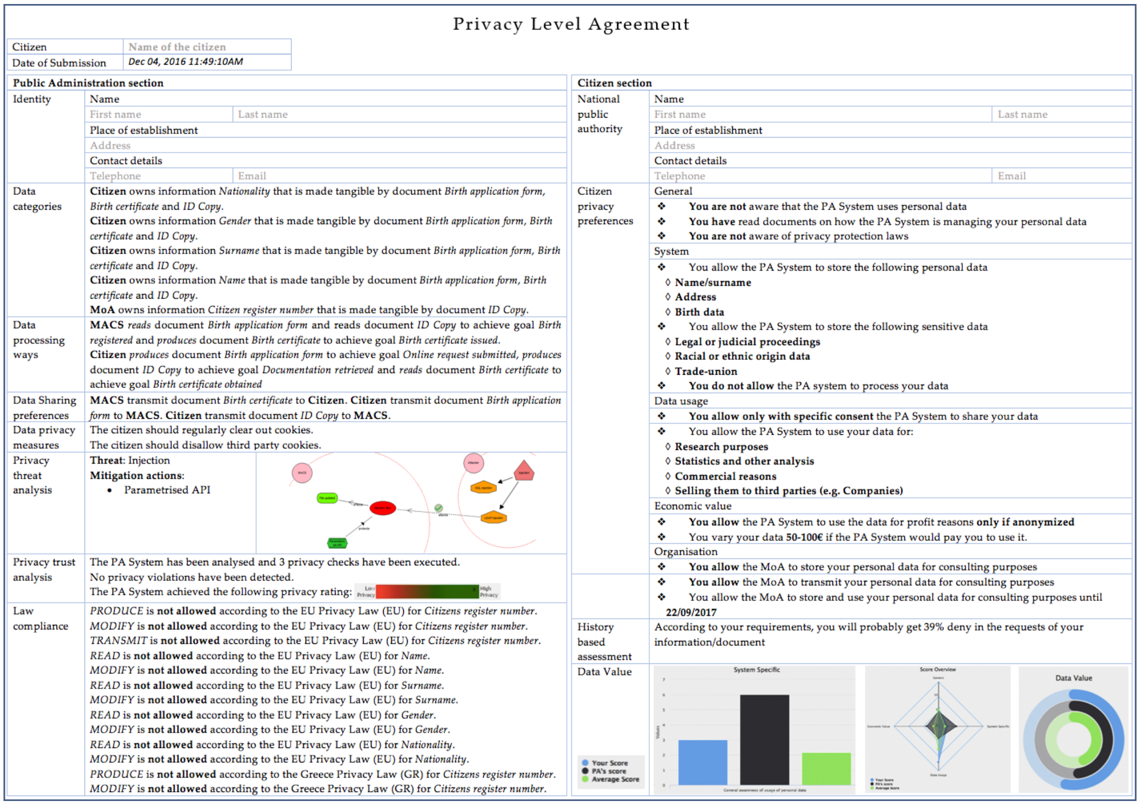
Step 2: With SecBPMN2 tool, the PA administrator models the business processes that are executed at the PA system and checks if security policies, derived from privacy requirements from Step 1, are satisfied by the business processes.

Step 3: The PA administrator uses the SecTro tool which imports the organisational structure and the privacy requirements, that are modelled in the STS-tool. The import is done through XSLT transformations[1] in order to convert the diagram from one modelling language to the other. In the converted model, the PA associates threats to the fulfilment of the privacy and security requirements and vulnerabilities of the PA system and uses a pattern library to propose mechanisms that will mitigate them.

Step 4: The PA administrator uses JTrust which imports the same model as in Step 3. In this step, the PA models and analyses the trust relationships among the entities that participate in the PA system. When an entity is not considered sufficiently trustworthy in order to fulfil the goals that are assigned with, control mechanisms are introduced to enhance the trust to the system. For example, if a system user cannot be trusted to change their password often, the PA administrator introduces a control mechanism that monitors the last time the password was changed and deactivates the account of the user, unless they update their password.

Step 5 The PA administrator uses CARiSMA to model the architecture of the PA system by using UMLsec. Then, the former performs

[1] https://www.w3.org/TR/xslt



Privacy Level Agreement

Citizen	Name of the citizen
Date of Submission	*Dec 04, 2016 11:49:10AM*

Public Administration section		
Identity	Name	
	First name	Last name
	Place of establishment	
	Address	
	Contact details	
	Telephone	Email
Data categories	**Citizen** owns information *Nationality* that is made tangible by document *Birth application form, Birth certificate* and *ID Copy*. **Citizen** owns information *Gender* that is made tangible by document *Birth application form, Birth certificate* and *ID Copy*. **Citizen** owns information *Surname* that is made tangible by document *Birth application form, Birth certificate* and *ID Copy*. **Citizen** owns information *Name* that is made tangible by document *Birth application form, Birth certificate* and *ID Copy*. **MoA** owns information *Citizen register number* that is made tangible by document *ID Copy*.	
Data processing ways	**MACS** *reads* document *Birth application form* and reads document *ID Copy* to achieve goal *Birth registered* and *produces* document *Birth certificate* to achieve goal *Birth certificate issued*. **Citizen** *produces* document *Birth application form* to achieve goal *Online request submitted*, *produces* document *ID Copy* to achieve goal *Documentation retrieved* and *reads* document *Birth certificate* to achieve goal *Birth certificate obtained*	
Data Sharing preferences	**MACS** transmit document *Birth certificate* to **Citizen**. **Citizen** transmit document *Birth application form* to **MACS**. **Citizen** transmit document *ID Copy* to **MACS**.	
Data privacy measures	The citizen should regularly clear out cookies. The citizen should disallow third party cookies.	
Privacy threat analysis	**Threat**: Injection **Mitigation actions**: • Parametrised API	
Privacy trust analysis	The PA System has been analysed and 3 privacy checks have been executed. No privacy violations have been detected. The PA System achieved the following privacy rating:	
Law compliance	*PRODUCE* is **not allowed** according to the EU Privacy Law (EU) for *Citizens register number*. *MODIFY* is **not allowed** according to the EU Privacy Law (EU) for *Citizens register number*. *TRANSMIT* is **not allowed** according to the EU Privacy Law (EU) for *Citizens register number*. *READ* is **not allowed** according to the EU Privacy Law (EU) for *Name*. *MODIFY* is **not allowed** according to the EU Privacy Law (EU) for *Name*. *READ* is **not allowed** according to the EU Privacy Law (EU) for *Surname*. *MODIFY* is **not allowed** according to the EU Privacy Law (EU) for *Surname*. *READ* is **not allowed** according to the EU Privacy Law (EU) for *Gender*. *MODIFY* is **not allowed** according to the EU Privacy Law (EU) for *Gender*. *READ* is **not allowed** according to the EU Privacy Law (EU) for *Nationality*. *MODIFY* is **not allowed** according to the EU Privacy Law (EU) for *Nationality*. *PRODUCE* is **not allowed** according to the Greece Privacy Law (GR) for *Citizens register number*. *MODIFY* is **not allowed** according to the Greece Privacy Law (GR) for *Citizens register number*.	

Citizen section		
National public authority	Name	
	First name	Last name
	Place of establishment	
	Address	
	Contact details	
	Telephone	Email
Citizen privacy preferences	General	
	❖ **You are not** aware that the PA System uses personal data ❖ **You have** read documents on how the PA System is managing your personal data ❖ **You are not** aware of privacy protection laws	
	System	
	❖ You allow the PA System to store the following personal data ◊ **Name/surname** ◊ **Address** ◊ **Birth data** ❖ You allow the PA System to store the following sensitive data ◊ **Legal or judicial proceedings** ◊ **Racial or ethnic origin data** ◊ **Trade-union** ❖ **You do not allow** the PA system to process your data	
	Data usage	
	❖ **You allow only with specific consent** the PA System to share your data	
	❖ You allow the PA System to use your data for: ◊ **Research purposes** ◊ **Statistics and other analysis** ◊ **Commercial reasons** ◊ **Selling them to third parties (e.g. Companies)**	
	Economic value	
	❖ **You allow** the PA System to use the data for profit reasons **only if anonymized** ❖ You vary your data **50-100€** if the PA System would pay you to use it.	
	Organisation	
	❖ **You allow** the MoA to store your personal data for consulting purposes	
	❖ **You allow** the MoA to transmit your personal data for consulting purposes ❖ You allow the MoA to store and use your personal data for consulting purposes until **22/09/2017**	
History based assessment	According to your requirements, you will probably get 39% deny in the requests of your information/document	
Data Value		

Figure 2: Example of a PLA

a security analysis in order to verify if the architecture satisfies the requirements imported from STS. The verification process includes checks on the security level of the communication among the components of the architecture.

Step 6: The PA administrator uses LIONoso to perform compliance check of the PA system against laws and regulations. More specifically, LIONoso permits to define the constraints specified by laws and regulation, and the operation executed by the PA system on citizens data. Then, the tool controls if the constraints are violated. If the model is found compliant, i.e. the laws are not violated, then the phase terminates, otherwise the PA must modify the system, starting from Step 1.

PLA Generation Phase. In the next phase, the PLA is generated by using the questionnaires provided by the PA and the answers given by the citizens.

Step 1: The PA administrator accesses DAE through ViTo, which guides the citizen to create questions to the citizens about the PA system and how they wish their data to be treated. DAE also collects the diagram information from the STS models and in particular what access rights are documented for each piece of data transmitted within the PA system. Then, the PA administrator forms questions accordingly to ask the citizens if they agree with these access rights.

Step 2: The PA administrator assigns scores to evaluate the sensitivity of the data mentioned in the questions. These values are inserted to DVT which presents in web diagram the value of the data mentioned in the PLA from a) the PA's perspective; b) the citizen's perspective; and c) an average of both.

Step 3: The PA administrator publishes the questionnaire.

Step 4: The citizens answer the questionnaire and provide their values for the sensitivity of their data.

Step 5: PAE collects the citizens' answers and creates privacy policies that will be continuously monitored in the next phase.

Step 6: For each citizen that answers a questionnaire, ViTo gathers the responses from the citizens, the related laws, the diagram information, including the security and trust analysis results and the data value of the citizen's data. This information composes the PLA for each citizen.

PLA Enforcement Phase. Every time there is a request for accessing data of a citizen, then the policies, dictated by the PLA, should be respected. This phase takes place at runtime and is implemented in the following steps:

Step 1: PAE receives a new request from an external entity to access data provided by a citizen and controls, based on the privacy policies created by the answers of the citizens to the questionnaires, if the access should be permitted or denied. Moreover, PAE sends notifications through ViTo to the citizen when someone attempts to access their data.

Step 2: MANE updates the network traffic rules based on PAE's feedback. Therefore, if the access by the entity is permitted or denied by PAE, MANE will form a rule to permit or deny respectively future requests from the same entity for the same data.

Step 3: The logs created by PAE and MANE that contain information about the amount of requests which were denied or permitted, are inserted in LIONoso. The latter, performs a history based assessment and provides in the PLA page of the citizen a value about how the percentage of the requests to their data will be denied.

4 CASE STUDY

To better illustrate the functionalities and the benefits of VPP we present how it is applied in a real-world system which is part of the pilot stage of our project. The system belongs to DAEM S.A., a government organisation that develops e-government services for the Municipality of Athens (MoA) and other local government organisations in Greece. More specifically, DAEM S.A. is responsible for the development and maintenance of the Municipality of Athens Computer Services (MACS), an information system of MoA that stores and manages personal data of Athenian residents.

4.1 Motivating Scenario

MACS is integrated with information systems of other organisations such as hospitals, banks, the fiscal office, sports facilities and many others. The main purpose of MACS is to interconnect such organisations, store and transmit information that belongs to a citizen upon request (e.g., birth certificates) without requiring the citizen's physical presence. In our scenario, the citizen wishes to buy a subscription at a local Swimming Pool facility. Athenian residents, i.e. Greek citizens whose permanent residence is registered in the city of Athens, have the right of a 10% discount. As proof of their residence, they are required to provide the facility with a birth certificate.

This certificate is provided directly by MACS to the information system of the swimming pool, as part of the e-government services in the MoA. The request will be triggered when the citizen visits the swimming pool for the first time to buy their subscription. It is also required that the citizen provides a medical certificate to prove that they do not have any skin condition. This certificate can be provided by the citizen, who needs first to visit their physician. In case the citizen has recently received a medical certificate which is stored in the clinic's database, MACS is able to retrieve it and forward it to the information system of the swimming pool facility.

The citizen, in order to access the area of the swimming pool in the facility, uses a badge. Each entrance is stored in the database of the information system and can be used to make personalised offers to the citizen. For example, such offers include higher discount for less popular hours or when the swimming pool is most busy.

In this scenario, there are three types of data of the system that will be handled by the PA system; a birth certificate, a medical certificate and the swimming pool access logs. Given that MACS is interconnected with numerous swimming pool facilities around Athens and other services, the citizen must provide their privacy preferences about which information wishes to be shared and how. Furthermore, the citizen, without knowing who is accessing their data might be reluctant to share it and eventually use MACS. This means that the PA administrators of MoA must guarantee the transparency of the data sharing process with other organisations. Finally, to avoid citizens blindly denying or permitting access to all their data, it is crucial to a) increase the awareness of the value of the shared data and b) inform the citizens about the mechanisms applied to protect their privacy.

4.2 Applying the VisiOn Privacy Platform

Hereby, we describe how VPP is integrated with MACS and illustrate execution instances of the three phases that we described in Section 3. Note that both MACS and VPP are installed in the premises of the MoA. Additionally, the data requested by each external service is always provided from or through MACS.

During the Requirements Specification phase, the administrators of MoA use the tools of the Desktop Framework to capture the organisational settings of the system and perform security and trust analysis. Given that multiple administrators might be working on the same models, and, therefore the Desktop Framework has more than one instances, a locking mechanism has been implemented in the VDB to avoid synchronisation issues.

First, the administrators of MoA design the STS-ml model of their system which describes the goals of each entity that participate in their system. In Figure 3, a partial STS-ml model is depicted which illustrates four actors, the Citizen, the Clinic, MoA which hosts the MACS system and the Swimming Pool (SP) Information System. The exchange of information that is captured in this model is annotated with privacy requirements. More specifically, the transmission of the medical certificate is associated with three requirements, i) confidentiality, i.e. the information is not disclosed to any unauthorised entity, ii) integrity, i.e. the information is not modified by any unauthorised entity and iii) undetectability, i.e. the inability for a third party to distinguish who is the user (among a set of potential users) using a service.

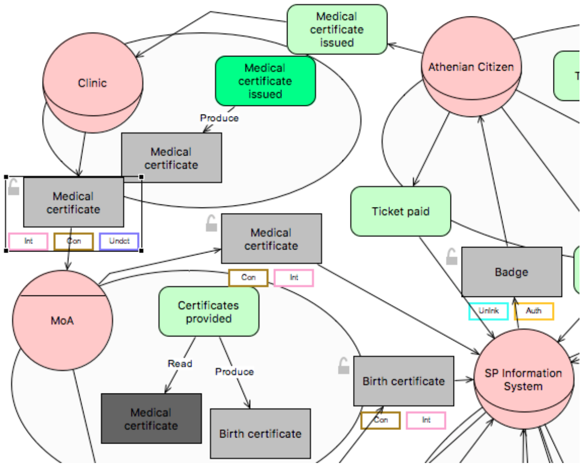


Figure 3: Example of a partial STS-ml diagram

Next, the MoA administrators design the business processes that are implemented by MACS and the interaction with other systems, by using SecBPMN2. Figure 4 shows portion of a SecBPMN2 diagram for our scenario. In particular, the tool specifies two activities executed by the SP information system. The first activity, Arrange ticket price, changes the price of the entrance ticket, by using the Visiting Time Record which is created by the Citizen. The activity Allow access reads the Medical certificate and stores it in the

Date	Resource	Subject	Role	Purpose	Action	Response
10 March 2017	MoACitizenXTA348065	SP_UserLT45675	SP_Administrative	read access log	Read	Deny

Table 1: Privacy policy from PAE

local database, allowing the citizens to access the SP information system. SecBPMN2 extends BPMN 2.0 with security and privacy annotations, which are represented with orange solid circles. In this case, the medical certificate is associated to an integrity annotation, which means that only authorised users can modify the document. The non-repudiation annotation is linked to the second activity, and it specifies that the SP information system must store a proof of the execution of that activity.

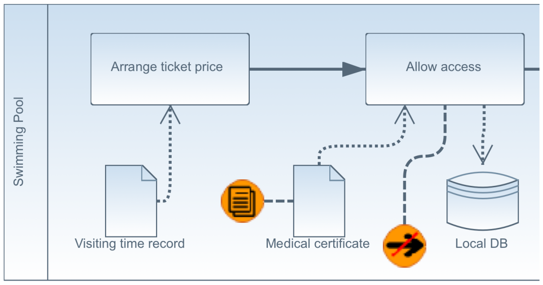


Figure 4: Example of a partial SecBPMN2 diagram

Then, the MoA administrators use the SecTro tool in order to enrich the STS-ml model with security requirements. Figure 5 depicts the privacy requirement Prevent unauthorised detection, identified previously with STS, which is modelled as a privacy constraint that restricts the use of the resource Medical certificate. Therefore, the privacy objective that satisfies the privacy constraint is to achieve undetectability with regards to the Medical certificate of the citizen. Privacy mechanisms are stored in a library of tools and specific ones can be selected according to the needs of each case. In this case, a set of administrative tools was selected, such as smart cards and permission management, along with a set of anonymizer mechanisms, such as Hordes, GAP, and Tor. This analysis enables the justification of why specific privacy and security mechanisms need to be placed and added to the PLA to assure the citizen about the level of their privacy protection and increase their trust in the service provided by MoA.

For the next step of this phase, the MoA administrators construct, with the use of JTrust tool, a trust model, as shown in Figure 6, where the citizen depends on the SP information system to have their visiting time kept confidential. This dependency implies a trust relationship between the Citizen and the SP information system. The trust relationship is justified with reported trust, i.e. MoA reports that the SP information system can be trusted to keep the visiting time confidential. Such information is elicited as part of the domain investigation and analysis. As a result, there is an underlying assumption that MoA can be trusted, represented with an ellipse symbol. Consequently, a new dependency is introduced on MoA for the validity of what is reported. The newly introduced dependency, and therefore trust relationship, is justified with normative trust, which is trust that is based on the norms of the system's environment. This information was elicited during the domain analysis. Likewise, there is an underlying assumption that the domain norm can be trusted. These two identified assumptions were further investigated and found to be valid. The developed model enabled the identification and justification of trust assumptions that underlie the analysis in order to be sound. In case there was lack of trust then control mechanisms would have to be added in order to enforce the fulfilment of goals such as visiting times to be kept confidential.

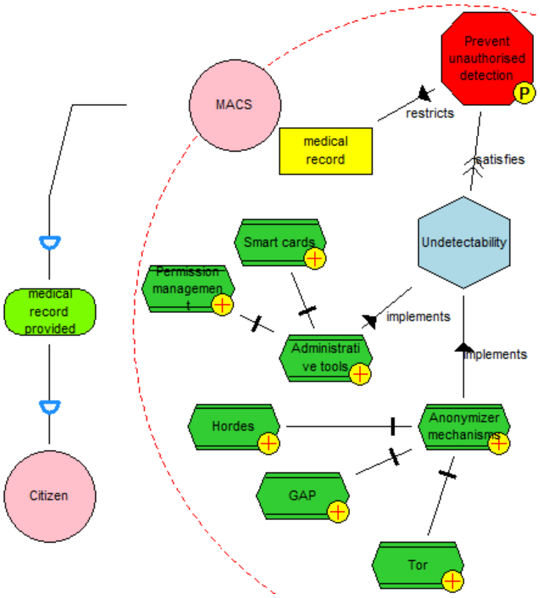


Figure 5: Example of a partial SecTro diagram

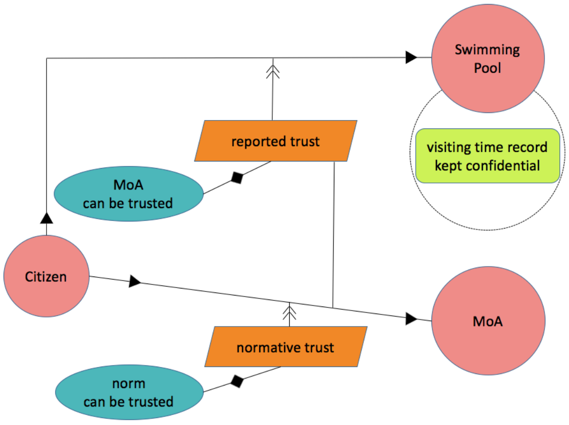


Figure 6: Example of a partial JTrust diagram

Date	Resource	Subject	Requested URL	Requesting IP	Response
10 March 2017	MoACitizenXTA348065	SP_UserLT45675	1.2.3.4:9999/log/179	5.6.7.8	Deny

Table 2: Network traffic rule from MANE

The use of the Desktop Framework is completed with CARiSMA, where the MoA administrators design the system's architecture by using UMLsec. The security annotations over the elements of the UML diagram are inserted from the STS-ml model. Then, the tool performs checks to validate if the architecture satisfies the security requirements. For instance, in our design the component that represents the information system of a citizen's clinic is connected with the component that represents MACS. Given that this transmission is annotated with a confidentiality requirement in STS-ml model on the side of the sender and the receiver, a security annotation must be added to both components, otherwise the checks performed by the tools will fail. This guarantees that the engineers of the system will implement all the necessary security measures to guarantee the protection of the citizen's data.

The Requirements Specification process concludes with the compliance check of LIONoso which receives the three types of data transmitted through the system from the STS-tool. The MoA administrators must insert through ViTo the Greek and EU privacy laws related to these types of data. Then, they specify through the same interface the operations that are applied over this data, e.g., MoA produces birth certificate. When all operations and the subjects who perform them are defined, LIONoso checks if the prescribed operations are compliant with the existing laws.

When the system is verified by LIONoso to be compliant with the existing privacy laws, the MoA administrators proceed with the creation of the questionnaires with the use of DAE through ViTo. The website guides the administrators on the creation of the questions by demonstrating information derived from the STS model. In particular, DAE suggests to the administrators to create questions related to the storage and the transmission of the data that are modelled in STS. For example, a few questions submitted are: i) 'Do you allow third parties to store your medical certificate?', ii) 'Do you allow sharing your birth certificate with third party organisations?', iii) 'Do you allow your access log to sport facilities to be used for commercial purposes?'.

When registering to MACS, the citizen has to answer a set of answers about which organisations that are integrated with MACS authorises to access their personal data and to handle them. For instance, they are asked if they wish the sport facilities that they are visiting to store and use for commercial use their visiting time record. DVT also provides them the information that this type of data is useful for the sport facilities organisations, in order to schedule their opening times and pricing policies. By sharing this information in the future, they could also receive personalised offers. Moreover, DVT informs that most citizens do not share this information, probably due to privacy concerns. After the questionnaire is published, every citizen when registering to MACS has to answer it. These answers formulate the privacy policies in PAE and ViTo generates and displays the PLA to each citizen.

After a few visits in the swimming pool, the SP information system will attempt to read the visiting time record of the citizen. PAE will receive this request and will check if this action is allowed, based on the PLA of the citizen. If the citizen forbids sport facilities to use their visiting time record for commercial purposes, then PAE has a policy, as the one shown in Table 1. The outcome of both PAE and MANE is forwarded as input to LIONoso which updates the number of requests to documents are being denied and displays it in the PLA.

After the request arrives at PAE, the result of the request, in this case is 'Deny', is forwarded to MANE, which creates network traffic rules, as the one shown in Table 2. This will cut automatically future requests to this piece of information. Moreover, a notification will be sent to the citizen about who tried to access their data and the result of the attempt.

If a malicious user or unauthorised entity attempts to access the citizen's data, VPP will reject the request and inform the citizen about who requested access to their data. Hence, the citizen is continuously updated about who is accessing their data. This increases the trust of citizens in the PA as they will see system protects and informs of all accesses, being this valid or not. For example, if an administrator who works for the Swimming pool tries to access their Visiting Time record or their Medical certificate, the citizen receives a notification (e.g., via SMS) about who accessed their data, when and for what purpose.

The process that takes place at system level is the following. First, the information system of the Swimming pool which is integrated with MACS and VPP, sends the request of data to the latter. Then, PAE processes this request and, using the privacy policies defined by the citizen's answers, detects the request is unauthorised. PAE registers this denied access and sends the information to LIONoso, which will process and display this information in the VPP as a new notification so the citizen, when accessing the system next time, can check it. Additionally, PAE will send the response of the request to the system informing about the denial of the request because of unauthorised permission. Additionally, VPP has a second line of protection of privacy of the data which is MANE. Therefore, if a malicious user would be able to retrieve the data from the database, using a third-party application that access directly to the data storage, MANE would check continuously the exchange of data in order to detect unauthorised accesses. If detected, it compiles the information and sends to LIONoso for notification to the citizen and the PA, as this illegal access could be a potential security weakness in the system. Consequently, the PA could use this information for revision of the security architecture of its system.

4.3 Discussion

Our platform with the use of the Desktop Framework, that we described earlier, allowed PA administrators to identify and model all the nodes (organizations, physical persons etc.) that the citizens' data pass through and define the access rights of each one of them based on their goals. Such models offer better understanding of how e-government systems operate and what are the risks related to private data. Additionally, by using SecTro and CARiSMA, the PA administrators are able to systematically identify the potential threats and vulnerabilities of their systems, propose defence mechanism and assess through various types of analyses the security level of the modelled system. The defence mechanisms and how they mitigate the identified threats are also presented in the PLA in order to increase the citizens' trust to the PA system. The analyses performed by JTrust also contributes towards this direction, by justifying why each node of the system can be trusted or not and provide this information through the PLA.

Despite the high level of maturity of the modelling languages of the Desktop Framework tools, training the pilot partners of the VisiOn project to use them has been a cumbersome task. The main reason was the reduced capacity of the pilots to provide multiple experts in the areas of software and security engineering along with domain experts in order to collaboratively produce the required diagrams,. To overcome this difficulty we organised workshops, tutorials and webinars which permitted our pilots to successfully use our tools and minimise the risk of human error. The user satisfaction has been confirmed during the evaluation process of VPP, where each pilot demonstrated the use of the platform, as it is integrated in their systems, and provided questionnaires[2].

Another contribution of our platform is the facilitation of the questionnaire creation for gathering privacy preferences from the citizens by DAE and ViTo. More specifically, the diagram information provided by the tools of the Desktop Framework guides the PA administrators to ask question about every data that is circulated within their system, ensuring the completeness of the privacy policies that will be composed from the answers. Moreover, by taking into account the opinion of the citizens on how they wish their data to be treated and customizing the privacy policies based on their preferences, VPP increases citizens' trust in e-government services. This can also be confirmed by the results of the aforementioned pilot evaluation. VPP also increases the awareness of the citizens about the value of their data and enables them to choose more carefully how they want their data to be managed in order to prevent future dissatisfaction.

Finally, the automated privacy enforcement provided by PAE and MANE relieves the PAs from the burden of performing manually regular audits on their systems to evaluate the level of privacy protection. These tools also promote transparency and accountability in e-government by capturing who requests what data, for what purpose and notify the citizen about these requests.

5 RELATED WORK

Various approaches have been proposed in the literature for systematically capturing privacy requirements. The Privacy Safeguard (PriS) [14] methodology enables the elicitation of privacy requirements in the software design phase, where privacy requirements are modelled as organisational goals. Next, in [25] the authors adopt the concepts of privacy-by-policy and privacy-by-architecture, and propose a three-sphere model of user privacy concerns, relating it to system operations (i.e. data transfer, storage and processing). Additionally, the Modelling and Analysis of Privacy-aware Systems (MAPaS) framework [6] is a framework for modelling requirements for privacy-aware systems. The ABC4Trust project [23] protects privacy in identity management systems. Differently than these works, VPP provides a start-to-end implementation of a privacy management approach that takes into account the PbD principles, since it starts with the elicitation of the user privacy needs and it ends with the provision of PA online services.

Trust analysis is yet another contributor to effective privacy management. A trust analysis method is proposed in [27] where the authors address the issues of trust at a requirements level and treat trustworthiness as an objective of the stakeholders. Next, in [12] the system analysis and design considers different domains in mobile communications. Additionally, in [9] the authors propose a method for discovering trade-offs that trust relationships bring between trust and control. Compared to these approaches, our work is applicable to PA organisations and the described platform facilitates the identification of organisational controls that will ensure privacy of citizens' data.

Recently, quite a few commercial products have been developed that highlight the importance of the individuals' data protection. The TRUSTe[3] platform focuses on Data Privacy Management (DPM), enabling users to take control of a set of technology-driven solutions for managing privacy challenges. Disconnect[4] is a software that facilitates users to easily understand the websites' privacy policies and realise how websites are handling their data. The common characteristic of these products is that they focus on the better analysis and comprehension of each privacy policy, protecting user from actions that will put their personal data in danger. Contrary to these products, VPP elicits from both sides (service providers and service consumers) their privacy preferences, developing personalised PLAs, according to them.

The Information Shield[5] provides a repository containing all the necessary material that can assist companies and organisations to formalise or update their privacy policies, maintaining them compliant with the relevant laws and regulations, at national and international level. Nymity[6] enables organisations to use an accountability approach to demonstrate data privacy compliance. 2B Advice[7] is a group of companies offering consulting services concerning data privacy advice, software solutions and certifications. Otris privacy[8] is a software for data protection management, focusing on the planning, setting-up, operation and decommissioning of data processing methods. OneTrust[9] platform ensures the data privacy compliance, helping service providers to guarantee to their service consumers that they are compliant with the laws and the privacy policies. As opposed to these works, VPP follows a holistic approach in order to create each PLA, conducting security and privacy analysis of the information systems of each service provider, and ensuring that their processes are law compliant and based on these results, it retrieves the privacy preferences of a service consumer.

[2] The statistical results of the answers to the questionnaire of one of the VisiOn pilots https://sense-cloud.brighton.ac.uk:5001/sharing/Fdkk9aCNl

[3] https://www.truste.com

[4] https://disconnect.me/icons

[5] https://informationshield.com/

[6] https://www.nymity.com

[7] https://www.2b-advice.com

[8] https://www.otris.com/products/data-protection-management/

[9] https://onetrust.com

6 CONCLUSIONS

In this paper we presented a novel platform that improves privacy in Public Administration. In particular, VPP provides PAs with the ability to create citizen's PLAs using citizens' privacy preferences, which are elicited through clear and non-technical questionnaires. Also, VPP enables citizens to understand the value of their data, using enhanced visualisation elements, and use that value to determine their privacy preferences. Moreover, VPP brings together a set of software engineering methodologies and tools across different levels, from privacy requirements to run-time, and different perspectives, from data evaluation to privacy assurance. Such integration provides a clear advantage over existing software engineering approaches and tools, since it enables a holistic analysis of privacy needs that includes both PAs and citizens. Moreover, VPP is the only platform in the literature, which we are aware of, that identifies and analyses privacy threats for PAs and it enables them to allow citizens to indicate their preference for the potential privacy mechanisms that can be used to countermeasure the identified threats.

The project is strongly linked to citizens and PA authorities, and therefore provides socially important impacts. In particular, VPP increases user trust and confidence in PA online services, therefore decreasing the number of users that are reluctant to use such services. VPP enables, on one hand, PAs to manage private data in an accountable and transparent way, and on the other hand, it provides citizens with the ability to control their privacy when they must share their personal data with PAs. Moreover, VPP makes transparency and accountability inherent characteristics of all activities related to citizens' data within PAs. Monitoring how this data is used after it has been given to PAs is one of the main functionalities of VPP provided by the Web Framework. This, along with the enforcement of PLA, plays a critical role in the maximisation of transparency and accountability.

ACKNOWLEDGMENT

The authors would like to thank DAEM S.A. for providing us with information about the e-government services of the MoA and the consortium of the VisiOn Project for their effort in the realisation of this work. This paper is supported by European Union Horizon 2020 research and innovation programme under grant agreement No 653642, project VisiOn (Visual Privacy Management in User Centric Open environments).

REFERENCES

[1] Data Protection Act. 2014. Conducting privacy impact assessments code of practice. (2014).

[2] Roberto Battiti and Mauro Brunato. 2014. *The LION way. Machine Learning* plus *Intelligent Optimization.* LIONlab, University of Trento, Italy.

[3] France Bélanger and Lemuria Carter. 2008. Trust and risk in e-government adoption. *The Journal of Strategic Information Systems* 17, 2 (2008), 165–176.

[4] Grady Booch, James Rumbaugh, and Ivar Jacobson. 2005. *Unified Modeling Language User Guide, The (2nd Edition).* Addison-Wesley Professional.

[5] Sofia Elena Colesca. 2009. Increasing e-trust: A solution to minimize risk in e-government adoption. *Journal of applied quantitative methods* 4, 1 (2009), 31–44.

[6] Pietro Colombo and Elena Ferrari. 2012. Towards a modeling and analysis framework for privacy-aware systems. In *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom).* IEEE, 81–90.

[7] Fabiano Dalpiaz, Paolo Giorgini, and John Mylopoulos. 2013. Adaptive Socio-Technical Systems: a Requirements-driven Approach. *Requirements Engineering* 18, 1 (2013), 1–24. Issue 4.

[8] F. Dalpiaz, E. Paja, and P. Giorgini. 2011. Security Requirements Engineering via Commitments. In *Proceedings of workshop on Socio-Technical Aspects in Security and Trust.* 1–8.

[9] Golnaz Elahi and Eric Yu. 2009. Trust trade-off analysis for security requirements engineering. In *2009 17th IEEE International Requirements Engineering Conference.* IEEE, 243–248.

[10] Rebecca Eynon. 2007. Breaking Barriers to eGovernment: Overcoming obstacles to improving European public services. *DG Information Society and Media. European Commission* 90 (2007).

[11] Mohamad Gharib, Mattia Salnitri, Elda Paja, Paolo Giorgini, Haralambos Mouratidis, Michalis Pavlidis, José F Ruiz, Sandra Fernandez, and Andrea Della Siria. 2016. Privacy Requirements: Findings and Lessons Learned in Developing a Privacy Platform. In *Proceedings of Requirement Engineering (RE) conference.*

[12] Janusz Górski, A Jarzebowicz, Rafal Leszczyna, Jakub Miler, and Marcin Olszewski. 2005. Trust case: Justifying trust in an IT solution. *Reliability Engineering and System Safety* 89, 1 (2005), 33–47.

[13] J. Jurjens. 2002. UMLsec: Extending UML for Secure Systems Development. In *Proc. of UML.* 412–425.

[14] Christos Kalloniatis, Evangelia Kavakli, and Stefanos Gritzalis. 2008. Addressing privacy requirements in system design: the PriS method. *Requirements Engineering* 13, 3 (2008), 241–255.

[15] Günter Karjoth and Matthias Schunter. 2002. A privacy policy model for enterprises. In *Computer Security Foundations Workshop, 2002. Proceedings. 15th IEEE.* IEEE, 271–281.

[16] Günter Karjoth, Matthias Schunter, and Michael Waidner. 2002. Platform for enterprise privacy practices: Privacy-enabled management of customer data. In *International Workshop on Privacy Enhancing Technologies.* Springer, 69–84.

[17] Gang-Hoon Kim, Silvana Trimi, and Ji-Hyong Chung. 2014. Big-data applications in the government sector. *Commun. ACM* 57, 3 (2014), 78–85.

[18] Marco Casassa Mont and Robert Thyne. 2006. A systemic approach to automate privacy policy enforcement in enterprises. In *International Workshop on Privacy Enhancing Technologies.* Springer, 118–134.

[19] Haralambos Mouratidis, Nikolaos Argyropoulos, and Shaun Shei. 2016. *Security Requirements Engineering for Cloud Computing: The Secure Tropos Approach.* Springer International Publishing.

[20] OMG. 2011. *BPMN 2.0.* Technical Report. OMG. http://www.omg.org/spec/BPMN/2.0

[21] European Parliament. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Coucil of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). (2016). http://eur-lex.europa.eu/legal-content/EN/TXT/

[22] Michalis Pavlidis, Shareeful Islam, Haralambos Mouratidis, and Paul Kearney. 2014. Modeling trust relationships for developing trustworthy information systems. *International Journal of Information System Modeling and Design (IJISMD)* 5, 1 (2014), 25–48.

[23] Ahmad Sabouri and Kai Rannenberg. 2015. *ABC4Trust: Protecting Privacy in Identity Management by Bringing Privacy-ABCs into Real-Life.* Springer International Publishing, Cham, 3–16. https://doi.org/10.1007/978-3-319-18621-4_1

[24] Mattia Salnitri, Elda Paja, and Paolo Giorgini. 2016. Maintaining secure business processes in light of socio-technical systems' evolution. In *Proceedings of Model Driven Requirement Engineering Workshop.*

[25] Sarah Spiekermann and Lorrie Faith Cranor. 2009. Engineering privacy. *IEEE Transactions on software engineering* 35, 1 (2009), 67–82.

[26] Aprna Tripathi and Bhawana Parihar. 2011. E-governance challenges and cloud benefits. In *Computer Science and Automation Engineering (CSAE), 2011 IEEE International Conference on*, Vol. 1. IEEE, 351–354.

[27] Eric Yu and Lin Liu. 2001. Modelling trust for system design using the i* strategic actors framework. In *Trust in Cyber-societies.* Springer, 175–194.

[28] Peng Yu, Jakub Sendor, Gabriel Serme, and Anderson Santana de Oliveira. 2013. Automating privacy enforcement in cloud platforms. In *Data Privacy Management and Autonomous Spontaneous Security.* Springer, 160–173.