Data Protection Risk Modeling into Business Process Analysis

António Gonçalves[1] [0000-0001-6091-2624] and
Anacleto Correia[2] [0000-0002-7248-4310] and
Luis Cavique[3] [0000-0002-5590-1493]
[1] INESC-ID, UL-IST, Lisboa,1048-001 Lisboa, Portugal
antonio.goncalves@inesc-id.pt
[2] CINAV, Alfeite, 2810-001 Almada, Portugal
cortez.correia@marinha.pt
[3] Univ. Aberta and MAS-BioISI, FCUL, Campo Grande
Lisboa, Portugal
luis.cavique@uab.pt

A**bstract.** We present a novel way to link business process model with data protection risk management. We use established body of knowledge regarding risk manager concepts and business process towards data protections. We try to contribute to the problems that today organizations should find a suitable data protection model that could be used in as a risk framework. The purpose of this document is to define a model to describe data protection in the context of risk. Our approach including the identification of the main concepts of data protection according to the scope of the with EU directive data protection regulation. We outline data protection model as a continuous way of protection valued organization information regarding personal identifiable information. Data protection encompass the preservation of personal data information from unauthorized access, use, modification, recording or destruction. Since this kind of service is offered in a continuous way, it is important to stablish a way to measure the effectiveness of awareness of data subject discloses regrading personal identifiable information.

1 Introduction

With the General Data Protection Regulation (GPDR), applicable from beginning 2018 on Europe [1], organizations must have a privacy configuration for their services by including in their operation the data protection capabilities necessary for regulatory compliance and attainment user belief, and therefore maintain organizations competitiveness.

The key to organizations adapt to the new GDRPR requirements is the ability to change the way it interacts with suppliers, partners, competitors, and customers to achieve with the new data and security protection requirements [1]. Improvement on the way organization operates should be done to confirm the new organization objectives imposed by regulators. The European Data Protection Legislation is a complex issue, whose techno-regulation transfers a bureaucratic overhead to system developers. However, GDPR is more linked with the data privacy requirements.

Business Processes [2] defined as a set of inter-related events, activities and decision points that involve several actors and objects which collectively pursue a business objective and policy goal, are a suitable way to capture the organization reality [3] and using a Business Processes model is possible to establish a Process discovery, i.e., gathering information about an existing process and organizing it in terms of an ‘as-is’ process model risk management to analyze security and data protection concerts of an organization [2].

In this paper, we discuss the foundations of quality assessment of data protection in business process models. Our main contribution is an approach considering human behavior aspects observed in process models to calculate a degree of possibility of data protection concerns. We validate the approach using some business process model. The outcomes highpoint which benefits organizations can have from artifacts used for data protection violation detection.

The remainder of the paper is organized as follows. Section 2 presents Event driven Process Chains (EPCs) [4], a modeling language to specify the temporal and logical relationships between activities of a business process that we use to exemplify our model. Section 3 presents our data protection concepts. Section 4 addresses the problem of data protection risk management. In Section 5, the previous techniques are combined resulting in the proposed risk data protection model. Finally, in Section 6, we draw some conclusions.

2 Business Process Modeling

We define a business process as a collection of inter-related events, functions, decision points, business objects and IT entities that involve several actors and that collectively lead to an outcome that is of value to at least one customer [2][5].

A function corresponds to a task which needs to be executed. Events define the state before and after a function is executed. Connectors can be used to connect functions and events. There are three types of connectors: and, exclusive or and or. Business objects can be input data serving as the basis for a function, or output data produced by a function and finally, IT Entities are used to describe IT input elements which are needed to perform the process.

Figure 1 depicts the ingredients of this definition and their relations.

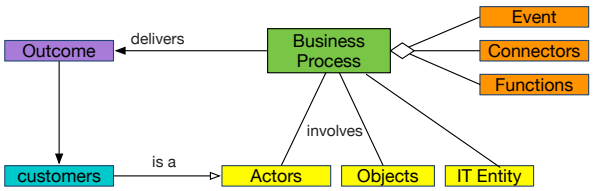

Figure 1. Elements of a business process.

To model a business process, we need some modelling language. The Event Driven Process Chains (EPC) is a business process modeling language that was presented in [6] and are used by us to describe organization process models [7].

The use of EPC as one its main purposes, to provide a notation understandable by different kinds of process modelers and users: (1) business analysts that sketch the initial documentation of business processes; (2) process developers which are responsible for implementing business processes; (3) business users which are accountable for business processes' instantiation and monitoring.

EPCs have some similarities with flowcharts but they differ from flowcharts in that they treat events as first-class citizens. EPCs specify the temporal and logical relationships between activities of a business process throw control flow [2].

EPCs offer offers the following element types: function type (i.e. activity that is executed in a process), event type (represent pre and post-conditions of functions) and connector type. All elements are linked via control flow arcs. In EPCs there are three distinctive kinds of connectors: AND, XOR, and OR. They may be used as either join connectors. Connectors have either multiple incoming and one outgoing arc (join connectors) or one incoming and multiple outgoing arcs (split connectors).

The semantics of an EPC connectors can be described as follows [6]. The AND-split activates all subsequent branches in a concurrent manner. The XOR-split represents a choice between one of several alternative branches based on conditions. The OR-split triggers one, two or up to all multiple branches based on conditions. For XOR-splits and OR-splits, the activation conditions are given in events after the connector. The AND-join waits for all incoming branches to complete, then it propagates control to the subsequent EPC element. The XOR-join merges alternative branches. The OR- join synchronizes all active incoming branches. The next description enacts EPC adapted from [5].

Definition 1 (EPC). An Event-driven Process Chain is a seven-tuple (*E, F, C, O,T, l, A*) such that:

-$\boldsymbol{E}$ is a finite, non-empty, set ($\boldsymbol{E \neq \varnothing}$) of event;

-$\boldsymbol{F}$ is a finite, non-empty, set ($\boldsymbol{F \neq \varnothing}$) of functions;

-$\boldsymbol{C}$ is finite set of connectors;

-O is finite set of business objects;

-$\boldsymbol{T}$ is finite set of IT entities;

-$\boldsymbol{l} \in C \rightarrow \{\wedge, XOR, \vee\}$ is a function which maps each connector onto a connector type;

-$A \subseteq (E \times F) \cup (F \times E) \cup (E \times C) \cup (C \times E) \cup (F \times C) \cup (C \times F) \cup (C \times C) \cup (F \times O) \cup (O \times F) \cup (T \times F)$ is a set of arcs.

Definition 1 shows that arcs of an EPC cannot connect two events or two functions directly, a well-formed EPC should satisfy other additional requirements.

Those expressions define that each event is at highest preceded by one input node and at highest succeeded by one output node. Every function has just one input and one output node. EPC needs at least one start event that is not preceded by any other node and one end event that is not succeeded by any other node. Connectors must have at least one input and one output node, but they can have numerous input nodes or numerous output nodes, with some restriction. In a well-formed EPC, there should be no paths connecting two events or two functions only via connector nodes in between.

Describe a Business process takes a significant part in an organization. It helps specify standard (as-is) and improved process of organization (to-be). Capture elements of business processes such us participants, their communications, resources contribute to organizational competiveness and could be a way to capture the data protection requirements. Thus, understanding and modelling of data protection becomes an important activity during a business process modeling.

3 Data Protection

The concept of data protection differs among different communities. We adopt the concept of data protection related with general legislation (EU directive 95/46/EC [8]) and privacy principles described by ISO/IEC 29100:2011 [9], which is mostly related with protecting personal data (i.e., Personal Identifiable Information or PII). The key concern of data protection is link to a person and related the protection of data that can be connected to an individual (i.e., data subject) [10].

Data can take on many forms. It can be printed or written on paper, stored electronically, transmitted by post or electronic means, shown on films, conveyed in conversation, etc. Normally organization try to use anonymized to avoid the require privacy protection. However, total anonymity is difficult, sometimes impossible [11].

The aim of data protection program at organization is to guarantee business continuity and minimize business damage by controlling the impact of data protection security incidents by implementing a framework according with a defined organization policy and consent compliance properties, according, for example, with EU directive [8].

Guarda et al. [12] describes the European Data Protection Directive in the following main principles: Fair and Lawful Processing; Data can only be collected and processed only if the data subject has given his explicit consent; lawful and legitimate use of personal data, data minimum necessary for achieving the specific purpose, Information Quality, Data Subject Control and Information Security.

A policy states general rules, determined by the stakeholders of the system, with respect to data protection. The policy and consent compliance property guarantees that the organization policy and the user consent are implemented and prescribed.

Several researcher emphasizes that data protection is not only a technology solution, but always encompass a process [13] [11]. For example, if we look at data security as a strictly technical issue, we also must take care of the process of securing these technical issues. Hence, it is necessary to evolve to extend beyond only the technical.

Data protection introduces an important set of definitions in terms of the properties that information manipulation should be concerned regarding privacy. These include personal identifiable information, PII (information which can be linked back to an individual), item of interest, IOI (information related to an individual) data subject (individual that is linked to the PII), unlinkability (not being able to distinguish whether two IOI are related), anonymity (not being able to identify the subject within a set of subjects), plausible deniability (being able to repudiate having performed an action), undetectability (not being able to distinguish whether an IOI exists), unobservability (undetectability against all subjects involved), confidentiality (authorized restrictions on information access and disclosure), awareness (being conscious about consequences of sharing PI information) and Compliance (following regulations and internal business policies) [14].

Regarding business process and data protection we can define, adapting from requirements engineering [15], data protection of business process (DPBP) as the elicitation, evaluation, specification, analysis and evolution of privacy objectives and constraints to be achieved by a business. A main concern of DPBP is to identifies potential threats and determine which threats are in fact applicable. To achieve that we should implement a Data Protection Risk Management to direct and control the risk.

4 Data Protection and Risk Management

The notion of risk is related to uncertainty from an expected organization objective and can be quantified as a positive or negative deviation. There are several ways to outline a risk. One of the possible ways is by linking it to the events that may happen, their consequences and the likelihood of the occurrence. The lack of information regarding the event occurrence, its consequence, or likelihood, is what drives to the state of uncertainty that underlies risk [16] [17].

Data Protection Risk management is an artefact that includes a set of coordinated activities performed to direct and control the risk of threads regarding properties that information manipulation should be concerned described in section 3. It includes a set of plans, relationships, accountabilities, resources, processes that provide the policy and objectives to manage data privacy risk. The risk management policy addresses the aims and strategy of the organization regarding risk management [16] [17].

Figure 2. Risk Management

A risk management framework can be understood as a system whose purpose is to ensure the fulfilment of the goal of risk management. It should also include a risk management process, and the resources and principles used in its implementation, as represented on Figure 1. These features can vary, however, the most important one in practice are grouped into three main stages (Figure 1) which can result in multiple solutions depending on the technical and technological support available to the risk management. The key concept of data protection risk, capture from ISO guide 73 [17], is present in table 1.

Table 1. Risk Management Concepts

Concept	Description
Risk	Effect of uncertainty on objectives.
Risk management	Coordinated activities to direct and control an organization about risk.
Risk management process	Systematic application of management policies, procedures and practices to the activities of communicating, consulting, establishing the context, and identifying, analyzing, evaluating, treating, monitoring and reviewing risk.
Risk management framework	Set of components that provide the foundations and organizational arrangements for designing, implementing, monitoring, reviewing and continually improving risk management throughout the organization.

5 Proposed Model for Data Protection Risk

In section 3, we analyze the privacy principles described by EU data protection regulation and ISO/IEC 29100:2011 and in section 4 we analyze the data protection risk management principles from the viewpoint of ISO guide 73 [17]. Both concepts are going to be used to arise the data protection model. The foremost goals with this approach is to contribute to minimize the following problems:

1. Despite the importance data protection for organizations, and therefore it enclosure on business processes, it is possible to find a set of common problems inside enterprises regarding protection, since security and data protection are integrate into organization in an ad-hoc way, often during the implementation process phase [18], or during the system administration phase of a business process management life-cycle [19];
2. There are not relevant artefacts used to protect data, made to business process. [20][19][21];
3. Data protection is not a group of principles and it cannot be reduced to the implementation of technological solution. It is a process involving various technological and organisational components, which implement privacy and data protection principles [22].

It is necessary to develop data security models that considers several organization security perspectives, such as static, about the processed information, functional, from the viewpoint of the system processes, dynamic, about the data security requirements from the life cycle of the objects involved in the business process, organizational, used to relate responsibilities to acting parties within the business process and the business processes perspective, that provides us with an integrated view of all perspectives with a high degree of abstraction.

Our approach is a model-based technique. That is relies on a representation of business process (BP). Since BP can be constructed to describe to viewpoint 'as-is' and 'to-be', our approach can be used to analysis the current data protection model or the future data protection model of a business and inforce the concept of privacy.

However, the definition of privacy varies depending on context, stakeholder interests. General privacy meanings comprise the right to informational self-determination and allowing individuals to control, edit, manage, and delete information about themselves and decide when, how and to what extent that information is communicated to others.

Our model, in Figure 3, integrated the main concepts related with information security applied to data protection and should be a baseline to implement and verify the stage of privacy.

A threat represents a possible violation of the security of a business asset with some negative impact [23] while vulnerability is a real security flaw which makes, an organization open to an attack. So, an attack is a use of a vulnerability to realize a threat. We can this combination an event. Threat modelling for data protection can support classify the threat, their attack surface and the entry or access points on business assets. A business asset is an element of business process that has value to the organization in terms of its business model and is necessary for achieving its objectives (e.g., function, business object, IT Entity). Data protection property on business assets characterizing their data protection requirements. Data protection property describe as a meter to measure the significance of risk. We adopt the taxonomy of privacy proposed by Pfitzmann [24] and from LINDDUN methodology [25], adapted as data protection property: i) Unlinkability means that all data processing is operated in such a way that the privacy-relevant data are unlikable to any other set of privacy-relevant data outside of the domain; ii) Transparency means

that all, privacy relevant, data processing, can be understood and reconstructed at any time. The information should be available before, during, and after the processing takes place; This allows that the data subject could have access to information requested from an organization; iii) Intervenability ensures intervention is possible concerning all ongoing or planned privacy-relevant data processing, by those persons whose data are processed. It allows the possibility of a data subject to request to rectification and erasure of data; iv) Anonymity refers to hiding the link between an identity and an action or a piece of information and v) Confidentiality refers to hiding the data content or controlled release of data content.

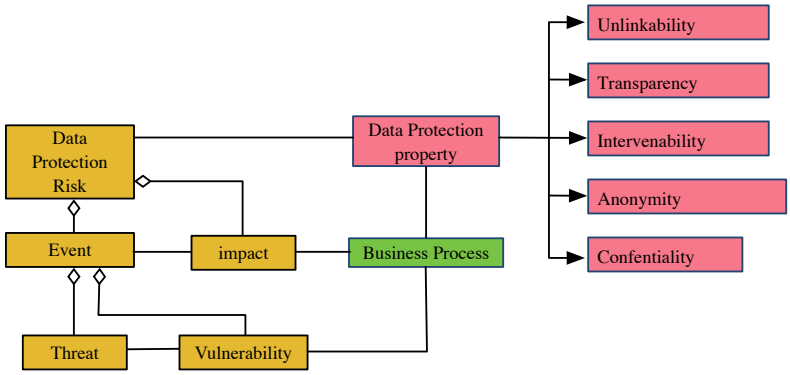

Figure 3. Data Protection Risk Model

We believe that from the business process perspective business analysts can integrate their view about business data security. Concerning the data protection requirements that can be modeled in business processes, it is necessary to consider that data protection requirements in any application at the highest level of abstraction will tend to have the same basic kinds of valuable and potentially vulnerable assets [29].

We can use risk data protection model in several stages of a risk management. risk data protection model centers on the potential weaknesses that might allow to someone cause the damage of a business asset. We can apply at diverse levels of abstraction, depending on the business assets considered. In other words, assets can be more abstract, such as in information objects our assets are more tangible, like IT components.

Besides, the development of a risk data protection model it is necessary to contemplate that apprehending the data protection of a Business process is a difficult work. One the benefices of using associated with a business process, is that it offers a structure view that can be used as a basis for the specification of data protection requirements. Business process model may present different levels of abstraction. Consequently, we believe that business analysts can integrate their view about business security into the business process perspective and in addition security requirements, since any application at the highest level of abstraction will tend to have the same basic kinds.

6 Conclusion and Future Work

Privacy can be defined as "the right of the individual to decide what information about himself should be communicated to others and under what circumstances" [26]. This description relates privacy to the right to control the information that is revealed to others. This is an issue that organization has concern with the new EU data protection regulation regarding data subject regrading personal identifiable information.

Data protection risks exist universally and can have costs every day, whether it is recognized by the organization affected by them. One of the main challenges that the organization must address is on the modelling of risk data protection using their context, i.e. business process model. Data protection modelling involves a highly heterogeneous set of business assets: events, methods, stakeholders and responsibilities, requiring adaptable methods and tools to support the exchange and interoperability of risk information, since risk management and risk assessment tend to be done by distinct teams with potential different views on the same risks.

Data protection as a key asset to today's organizations must be protected from increasing threats. Implementing data protection risk management compliant with ISO 31000:2009 [16] [17], EU directive 95/46/EC [8] and privacy principles described by ISO/IEC 29100:2011 [9], is the initial stage to ensuring data protection.

This work followed a model approach on the implementation of risk management and business process. Since the models are at a high level of abstraction, this approach contributes for bridging the gap within the information security community between domain analysts, who work with security at a domain level, and security implementers, who analyze the same issues at an architectural and design levels.

The research describe in this paper is driven by information security in any potential scenario that deals with information. A common way to model and address complex business systems is the use of model. Thus, we intend to bring the concepts and strategies of modelling into this subject. Modelling information risk is a complex task, especially in scenarios where protection of information is not a unique concern of the organization. To cope with information security risks, this there are need to have a cooperation between risk management and information security department, making it possible to align information security concerns with other, potentially relate, organizational concerns.

Future work intends to extend risk management and data protection models to include the dynamic perspective of information security, and conduct empirical studies for assessing the usability and efficacy of our approach in the risk management and information security domains.

References

1. The European Parliament, The European Council: General Data Protection Regulation. (2016).
2. Dumas, M., La Rosa, M., Mendling, J., Reijers, H. a.: Fundamentals of Business Process Management. (2013).
3. Becker, J., Kahn, D.: The process in focus. In: Process management. pp. 1–12. Springer (2003).
4. Scheer, A.-W., Thomas, O., Adam, O.: Process modeling using event-driven process chains. Process. Inf. Syst. 119–146 (2005).
5. Van Dongen, B., Dijkman, R., Mendling, J.: Measuring similarity between business process models. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 450–464 (2013).
6. Curran, T., Keller, G., Ladd, A.: SAP R/3 Business Blueprint: Understanding the Business Process Reference Model. Prentice Hall PTR (1998).
7. Becker, J., Kugeler, M., Rosemann, M.: Process management: a guide for the design of business processes. Springer Science & Business Media (2013).
8. Directive, E.U.: 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Off. J. EC. 23, (1995).
9. Drozd, O.: Privacy pattern catalogue: A tool for integrating privacy principles of ISO/IEC 29100 into the software development process. In: IFIP Advances in Information and Communication Technology. pp. 129–140 (2016).
10. Parlamento Europeu, Conselho da União Europeia: GDPR - EUR-Lex - 32016R0679 - EN. J. Of. da União Eur. 59, (2016).
11. Tucker, P.: Has Big Data Made Anonymity Impossible? MIT Rev. 116, 64–67 (2013).
12. Guarda, P., Zannone, N.: Towards the development of privacy-aware systems. Inf. Softw. Technol. 51, 337–350 (2009).
13. Laudon, K.C., Laudon, J.P.: Management Information Systems 13e. (2013).
14. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. (2010).
15. Sommerville, I., Kotonya, G.: Requirements engineering: processes and techniques. John Wiley & Sons, Inc. (1998).
16. ISO, I.: 31000: 2009 Risk management–Principles and guidelines. Int. Organ. Stand. Geneva, Switz. (2009).
17. Guide, I.S.O.: 73: 2009. Risk Manag. (2009).
18. Backes, M., Pfitzmann, B., Waidner, M.: Security in business process engineering. Bus. Process Manag. 168–183 (2003).
19. El-Attar, M., Luqman, H., Karpati, P., Sindre, G., Opdahl, A.L.: Extending the UML Statecharts Notation to Model Security Aspects. IEEE Trans. Softw. Eng. 41, 661–690 (2015).
20. Nunes, F.J.B., Belchior, A.D., Albuquerque, A.B.: Security Engineering Approach to Support Software Security. 6th World Congr. Serv. 48–55 (2010).
21. Abie, H., Aredo, D.B., Kristoffersen, T., Mazaher, S., Raguin, T.: Integrating a Security Requirement Language with UML. In: << UML>> 2004-The Unified Modeling Language. Modelling Languages and Applications. pp. 350–364 (2004).
22. Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Metayer, D. Le, Tirtea, R., Schiffner, S.: Privacy and Data Protection by Design-from policy to engineering. arXiv Prepr. arXiv1501.03726. (2015).
23. Oladimeji, E.A., Supakkul, S., Chung, L.: Security threat modeling and analysis: A goal-oriented approach. Proc 10th IASTED Int. Conf. Softw. Eng. Appl. SEA 2006. 13–15 (2006).
24. Pfitzmann, A., Kiel, U.L.D.: Pseudonymity , and Identity Management – A Consolidated Proposal for Terminology. Management. 1–83 (2008).
25. Wuyts, K., Scandariato, R., Joosen, W.: Empirical evaluation of a privacy-focused threat modeling methodology. J. Syst. Softw. 96, 122–138 (2014).