An approach to GDPR based on Object Role Modeling

António Gonçalves[1], Anacleto Correia[1], Luís Cavique[2]

[1] CINAV, Alfeite, 2810-001 Almada, Portugal
[3] Universidade Aberta, BioISI-MAS, 1269-01 Lisboa, Portugal

agoncalves@tecnico.ulisboa.pt, cortez.correia@marinha.pt, luis.cavique@uab.pt

Abstract. The General Data Protection Regulation 2016/679 (GDPR) is a set of legal rules to attain the privacy of people in the handling of their personal data and the movement of such data across countries. When those rules are considered in the operation of information systems, the one becomes attainable for legal approval within that scope. This paper presents a model we are developing to help enterprises do align their information system with the GDPR requirements. The model shall serve the purpose of analyzing the enterprises in what concerns the use of the subject's personal data, allowing to capture and improve data protection capabilities placed in the GDPR. The main issue of our approach is to set a baseline to define the requirements for establishing, implementing, maintaining and continually improving data protection management system on organizations.

Keywords: Personal Data Protection, Regulation (EU) 2016/679, GDPR.

1 Introduction and Motivation

The General Data Protection Regulation (GDPR) was stated by the European Union. It is a legal document consisting of a set of rules to achieve a high level of protection of natural persons in what personal data processing and the free movement of such data is concerned. The assurance of compliance with the GDPR, at the level of enterprise operation, demands an effort from people when analyzing and understanding such regulation since it is a mix of legal rules, organization rules, and technical rules.

We output that, an organization, to be compliance with GDPR, requires a dynamic approach where protection of personal data is achieved in a continuous way and personal data should be considered a valued organization information assets [1]. Personal data protection is part of a complex organization privacy process that encompasses the preservation of personal data from unauthorized access, use, modification, recording or destruction. Since this kind of process is offered in a

continuous way, it is important to measure the effectiveness of this process, i.e., the information security quality of process [2].

The goal of this paper is to propose a model that describes the concepts captured from the analyses of GDPR documentations. It is our goal to promote an improved understanding of the legal, organization and technical concepts present in the GDPR document.

From the breakdown of the GDPR document, we aim to elaborate a model of personal data protection, based on the capture of main concepts from the regulation.

We aim to use the model to help the operation of enterprise information systems, to show it possible for legal approval inside the scope of personal data protection, specifically within the requirements of the GDPR.

This paper also details some guidelines to assess the concept of personal data security risk assessment through the identification of threats and vulnerabilities that are carried out by the risk team. The goal is to operate the organization information system supported by a service based on risk management, in order to maximize the organization's output, while at the same time decreasing unexpected negative outputs generated by potential threads.

The analysis of the GDPR is a complex mission. It involved several tasks, namely reading, manual knowledge extraction, and characterization of many concepts and sentences expressed in that legal document.

The paper is structured as follows. Section two outline GDPR principles, that we used to present a methodology to analysis it. Section three related the GDPR data protection principles and information security principles. Section four explains our ontological model of analyzing the GDPR. Section five examines some challenge and future direction for our work.

2 GDPR Analysis

The GDPR is aimed at the protection of natural persons regarding the processing of their personal data. Across GDPR documentation is specified a set of definitions, such as personal data, sensitive data, and data processing. It also defines the actors involved in data processing such as: controller, processor, operator and subject, their roles, and responsibilities. Finally, GDPR portrays the obligations identifying with information controllers and processors with explicit reference to the legitimate motivation behind information handling and the reception of wellbeing safety efforts to control the risk unauthorized use of data.

According to the GDPR, a broad concept of the information system should be considered. For example, GDPR applies to a filing system if it is a set of structured personal data manageable via certain criteria. GDPR also mentions the possibility of data subjects getting direct access to their own personal data in order, for example to rectify personal data.

To implement GDPR on an organization a set of very important governance process may also be considered, such as process where controllers or processors could record and manage the way personal data is processed. This kind of process is very important since data subjects could request for the rectification of their own personal data, where controllers or processors could record processing activities and where data protection officers could monitor that registration.

GDPR, in the context of this paper, is a source to drive business process and also system requirements to support the analysis of information systems requirements, in order to capture the regulatory data protection capabilities disposed in the GDPR on the organization systems. The analysis of the GDPR involved a set of tasks as showed in Figure 1:

1. Task 1: Ontological Data Protection concepts. This task comprised of the familiarization of the organization with the domain of personal data protection. The output of this task is a set of GDPR domain terms definition. This task will allow the organization to understand regulatory systems, from a business perspective.
2. Task 2: Stakeholders Capture Model. This task comprised of identifying and typifying the parties interested in the development and operation of regulatory systems.
3. Task 3: Structure Model. This task comprised of an assembly composed of Data Protection concepts and stakeholders. This task consisted of modeling the relationship between the data protection concepts identified in the task1 and Stakeholders Capture identified in the task2, along with their attributes.
4. Task 4: Context Model. This task comprised of identifying each organization artifacts involved in the regulatory protection systems. The output of this task will be a list of in-scope artifacts and the characterization of its type.
5. Task 5: Regulatory protection business goals Model. This task comprised of identifying and characterizing the business processes, based on an ordering of tasks concerned with regulating the processing of personal data and identifying and characterizing the business goals to be achieved with the application of the provisions in the GDPR to regulatory systems.

The execution of these tasks is directed by a workflow that suggests the adopted GDPR analysis methodology.

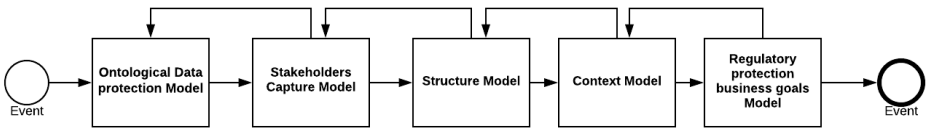

Figure 1. BPMN model of GDPR methodology.

3 The relation between Data Protection and Information Security models

Normally organization uses the term data protection, derived from the information security, to designate any artifact that has personal information with value for the organization, and for that reason, it needs to have adequate data protection [3].

While information security and data protection have a significant similarity, these two terms are not analogous. The general definition of information security comprises the CIA model of Availability, Integrity, and Confidentiality. Data protection includes other aspects that extend beyond the limits of information security, including the concept of data processing and genuine reason for information handling and the acceptance of security measures to limit information.

According to GDPR the aim of personal data protection is to guarantee business continuity and minimize business damage by controlling the impact of data protection incidents preserving the confidentiality, integrity, availability of personal data.

Additionally, it is also necessary to keep analyzing the legitimate purpose of personal data processing [4]. Data protection is not only technology but always encompass a process. For example, if we start information 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 in order to extend beyond only the technical [5].

Information security introduces an important definition in terms of the properties that data processing should have. These include the availability, confidentiality, and integrity of information. Availability refers to the fact that the personal data used by an organization remain accessible when they are needed. Thereafter, system failure is an organizational security issue [6].

Confidentiality refers mostly to restricting personal data access to those who are authorized. Organizations strive to control personal data access since technology offers developments aimed at making information accessible to the many [6].

Integrity refers to maintaining the values of the personal data stored and manipulated, such as maintaining the correct meaning. Personal data integrity is commonly assured by using cryptographic or/and information replication strategies. The cryptographic tools are indeed used to sign single pieces of data so that any faking information can be detected through signature validations [6].

According to ISO 31000, we can define personal data protection risk as the effect of uncertainty due to a deficiency of information that hinders achieving organizational objectives. Personal data protection risk management is a permanent challenging process which allows an understanding of the potential risks to the organization's valuable personal data assets and the tools to address them [7].

4 Proposed Ontological Data Protection Model

The integrated modeling of different aspects of an organization is only in the context of the emerging discipline of Organizational Engineering that concerns itself essentially with the modeling in three aspects:

- The core, essential model of an organization from the point of view of business (e.g., ontological model);
- The Business integration, information, and documents (e.g., the achievement of the Organization) and
- The model by which an organization is operated by all the people and information technologies (e.g., the implementation of the Organization).

An ontological data protection model is related to the construction and operation of an information system in the context of privacy. An ontological model has many advantages, such as it can help improve organizational consciousness; it enables sharing of knowledge between individuals through the representation of different organizational aspects, such as business processes, resources (technological artifacts, people, materials, etc.) and organizational structure.

From the lookout of the proposed ontological model of data protection, we support the proposal made by Dietz [8] where the ontological model abstracts from all realization and implementation issues.

The ontological model describes the essence of the business aspects of an organization. Realization model describes the detailed integration of all aspects of business, through the layered nesting, of information and document necessaries to the operation of the organization [8].

According to our suggestion, we can analyze the ontological of an organization with the use of fact model and use facts to find useful measures, viability norms, and dysfunction.

A Fact Model structures essential business knowledge about business concepts and business operations. It is occasionally named a business entity model. The main purpose of a fact model is to create a standard vocabulary by which all stakeholders can use to communicate clearly. Therefore, we can use facts that reproduces the real world.

The fact model focuses on the core business concepts, and the logical connections between them, which are named facts.

The facts are usually verbs which designate how one concept link to another. For example, the two concept Person and Personal Data may have a fact connecting them called have (a Person have Personal Data).

We agreed that the organization implementation is a result of an engineering process that can be analyzed from a fact model. This model can be used to understand technology (i.e., people, rules, a division of work and tools) that is part of organization operation. For that, we propose an ontological model to capture the essential structure of activities from ontological organization model. The model is present in fig. 2.

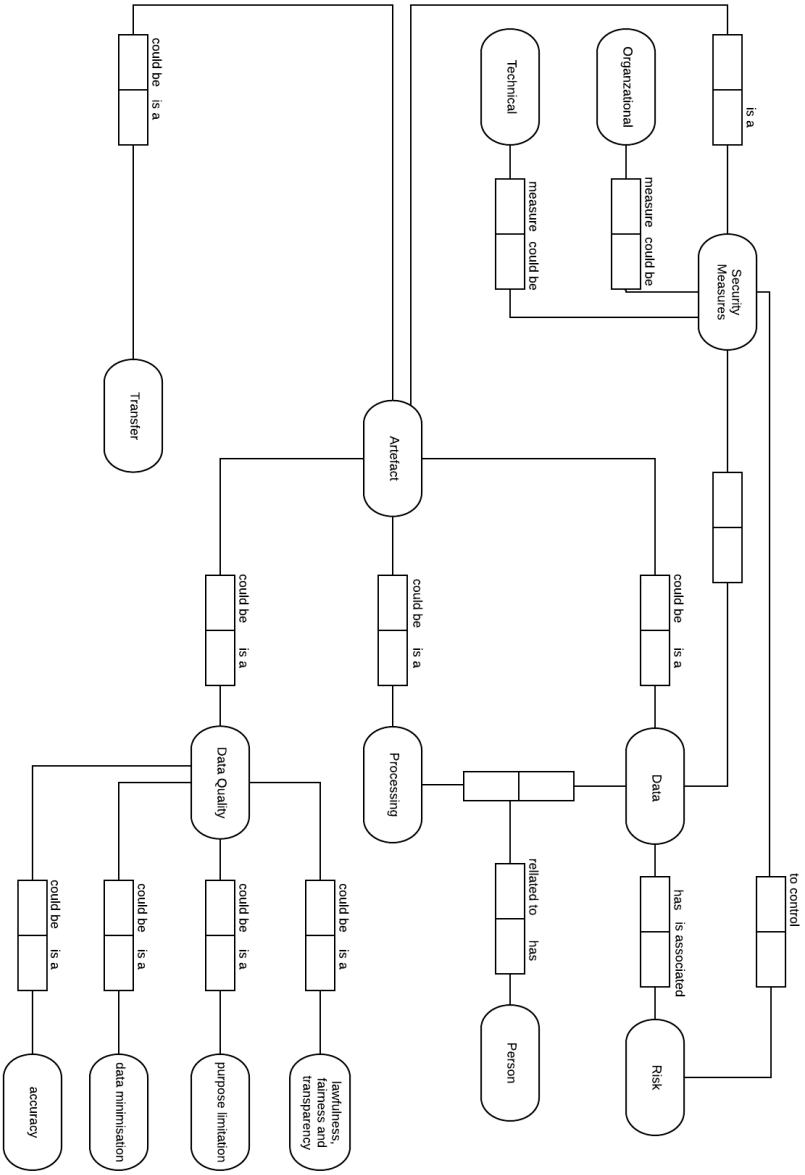

Figure 2. Ontological Model of Data Protection.

The model is composed of the following entity: E01 to E07. All entities are artifacts under this model. We identified the following entities: E0: data, E01: processing, E02: transfer, E03: data quality, E04: Risk, E05: Person, E06: lawfulness, fairness, and transparency, E06: purpose limitation, E07: data minimization, and E08: accuracy.

We identified the following facts that result in the relationship between two or more Entities: All relevant data we consider is data that is related to a person (i.e., that we can uniquely identify a person from that). Associated with personal data is the risk entity.

Risk entity is associated with uncertainty from an expected organization processing and can be quantified as a positive or negative deviation. There are several ways to output 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 [9] [10].

The definition of Data Quality is a multi-dimensional concept that varies depending on context, stakeholder interests. Data Quality depends on both subjective perceptions of the individuals involved with the data, and the objective measurements based on the data set in question. General Data Quality privacy is related with the following facts: comprise the lawfulness, fairness, and transparency of data processing, the purpose limitation that allows future processing if the new purpose if either is compatible with your original purpose. Data minimization creates the fact that data is acceptable, important and restricted for the end for which they are processed. Accuracy allows people to gain control about their personal data and choose when, how and to what degree the personal data is conveyed to other people.

A threat denotes a likely violation of the security of data privacy with some negative impact [11]. The vulnerability is a real security flaw which is an open door to an attack. So, an attack is a use of a vulnerability to realize a threat.

The Data protection facts model can be useful to measure the importance of risk. We use a taxonomy of privacy proposed by Pfitzmann [12], as follows: We categorize the entity unlikability, as a means to measure the relation between set of personal data outside of a specific domain; ii) Transparency, as a means that we can measure the likelihood that data, 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 a request to rectification and erasure of data; iv) Anonymity refers the measure of the set of subjects with potentially the same attributes and v) confidentiality refers to hiding the data content or controlled release of data content.

5 Conclusion and Future Work

This paper describes an ontological model on data protection and free movement according to European regulation 2016/679.

According to Dietz [8], it will only be possible to manage the complexity of an organization and reduce and manage its entropy through its ontological model. The ontological model, being coherent, comprehensive, consistent and concise, relies only on the essence of an organization and enables it to deal with the current and future problems of the business challenge. The assumption made in his proposal is that communication between the people of an organization provides the necessary and sufficient support to develop an organization theory.

In this context, an application of the holistic regulation is certainly useful and beneficial, in that its practice cannot be done in a disorganized and disenchanted way from the organization reality. However, to describe an organization is a complex task, since its representation must be done in an integrated way, which, if not supported in an organized approach, will translate into distinct and uncoordinated perceptions that manifest themselves internally and externally, so that the capture of relevant activities in an environment faces a high number of challenges, highlighting: the creation of methods that allow harmonious and cooperative apprehension by the people, the relevant activities and their respective articulation.

A number of approaches seek to capture activities for complex and dynamic domains such as teaching, information integration, the development of the Man-Machine interface, the description of system requirements. Unfortunately, however, uncertainty in these dynamic and complex domains prevents coherent understanding, in particular, because of the fact that there are sectarian and often inconsistent conspectations of their environment.

A starting point is an ontological approach, based on the description of the essence of the operation in an organization.

In this article, a fact-based model is proposed that describes and relates the main artifacts. To this end, interpretations of the main concepts of regulation and information security were suggested.

In the future it will be possible, from the proposed model, to advance to proposals for operation of treatment protection in organizations taking into account other socio-technical models, namely activity theory.

The theory of activity will allow the explicit description of the articulation of activities, based on the casual connection of the fundamental elements of the ontological model, enables to obtain a congruent model of the organization under the aspects of identification of the operations involved and the structuring of the activities in its nuclear elements

Acknowledgements

The work was funded by the Portuguese Ministry of Defense and by the Portuguese Navy/CINAV.

References

1. Peltier TR (2016) Information Security Policies, Procedures, and Standards: guidelines for effective information security management. CRC Press
2. Von Solms R, Van Niekerk J (2013) From information security to cyber security. Comput Secur 38:97–102
3. Cherdantseva Y, Hilton J (2013) A Reference Model of Information Assurance & Security. 2013 Int Conf Availability, Reliab Secur 546–555 . doi: 10.1109/ARES.2013.72
4. Whitman M, Mattord H (2011) Principles of Information Security. Cengage Learning
5. Laudon KC, Laudon JP (2013) Management Information Systems 13e
6. Andress J (2014) The basics of information security: understanding the fundamentals of InfoSec in theory and practice. Syngress
7. Purdy G (2010) ISO 31000:2009 - Setting a new standard for risk management: Perspective. Risk Anal. 30:881–886
8. Dietz J (2006) Enterprise Ontology: Theory and Methodology. Springer
9. ISO I (2009) 31000: 2009 Risk management–Principles and guidelines. Int Organ Stand Geneva, Switz
10. Guide ISO (2009) 73: 2009. Risk Manag
11. Oladimeji EA, Supakkul S, Chung L (2006) Security threat modeling and analysis: A goal-oriented approach. Proc 10th IASTED Int Conf Softw Eng Appl SEA 2006 13–15
12. Pfitzmann A, Kiel ULD (2008) Pseudonymity , and Identity Management – A Consolidated Proposal for Terminology. Management 1–83