





Sushant Agarwal and Sabrina Kirrane and Johannes Scharf

## Modelling the General Data Protection Regulation

Book Section (Published)

# MODELLING THE GENERAL DATA PROTECTION REGULATION

Sushant Agarwal / Sabrina Kirrane / Johannes Scharf

Ph.D. researcher, Institute for Management Information Systems, Vienna University of Economics and Business (WU)
Welthandelsplatz 1, 1020 Wien, AT
sushant.agarwal@wu.ac.at

Assistant Professor, Institute for Management Information Systems and Institute for Information Business, Vienna University of Economics and Business (WU)
Welthandelsplatz 1, 1020 Wien, AT
sabrina.kirrane@wu.ac.at

Associate, CMS Reich-Rohrwig Hainz Rechtsanwälte GmbH
Gauermanngasse 2, 1010 Wien
johannes.scharf@cms-rrh.com; https://cms.law/en/AUT/People/Johannes-Scharf

***Keywords:*** ***GDPR, Compliance, ODRL, Ontology, Data Protection***

***Abstract:*** *The new EU General Data Protection Regulation (GDPR) will soon replace the older data protection directive. Currently, the knowledge to comply with the regulation is only available in a human-readable format. If this knowledge is translated into machine-readable rules then computer based systems can ease data protection information retrieval as well as the process of checking GDPR compliance. In this paper, we model the obligations defined in the GDPR and then translate the model into a machine readable format by extending the Open Digital Rights Language (ODRL) ontology. The model is, in turn, used for a compliance checking tool.*

## 1. Introduction

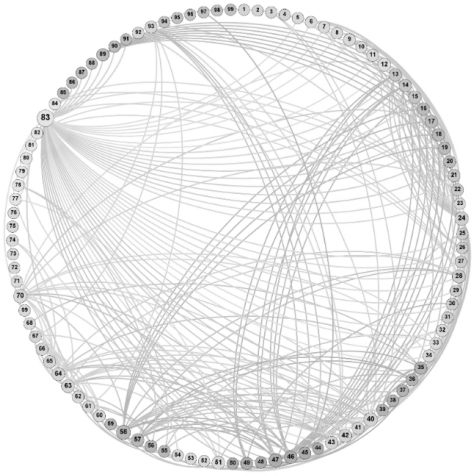

**Figure 1: Connections between the articles as defined in the GDPR text (approximately 350 connections), coloured based on the chapters**

The new EU General Data Protection Regulation (GDPR) has more than 80 pages of legal text with 99 articles. These 99 articles have many interconnections defined in the text, illustrated in figure 1. For instance, the obligation to ensure that information given to data subjects is transparent as defined in Article 12 para 1[1] cannot be checked in isolation as it is related to other obligations defined in Article 15-22 and 34. Thus articles cannot be analysed in solitude for ascertaining the compliance level. In total, there are approximately 350 interconnections defined, making it quite difficult and time-consuming to consider the applicable obligations as well as all the defined interconnections in order to check the compliance. Figure 1 shows a graph that illustrates all of the defined interconnections. Currently, a lot of human-readable reports have been prepared to provide a high-level list of requirements for the compliance checking process. However, to best of our knowledge no tool for filtering out applicable obligations defined in the interconnected mesh of the 99 articles of the GDPR exists. A software based tool can process all these interdependencies and can dynamically filter out the applicable obligations, easing the process of understanding the obligations as well as checking compliance. One of the primary objectives of our work is to translate the regulation's knowledge base (obligations defined in the articles) into machine readable rules such that the process for compliance checking can be assisted by the software tools. The rest of this paper is structured as follows: in section 2 we discuss the basic modelling of controller and processor obligations. Then in section 3, we translate the model into a machine readable form by extending the ODRL ontology. In section 4, we discuss related work which use the ODRL ontology and then lastly in section 5 we present our conclusions and discuss future work.

## 2. Modelling the obligations

The regulation has a total of 99 articles, however, not all of the 99 articles define obligations. For instance, there are articles which discuss other details like the objectives, definitions, GDPR's entry into force etc. Thus, such articles can be filtered out for our application. Also, as our objective is to provide a tool that can be used by controllers and processors for checking compliance with the GDPR, articles defining obligations for other parties like the Supervisory Authorities and European Data Protection Board can be neglected. After filtering out these articles, 31 articles that discuss the **obligations** for the controllers and processors were shortlisted. For the selected articles, due to the interconnections, not all obligations are at the same level, for instance, some can be defined independently, like *obtaining consent* (Art 6), whereas, some obligations like *being able to demonstrate consent* (Art 7 para 1) are dependent on the obligation of *obtaining consent* and thus cannot be considered independently. We refer to the dependent obligations as **sub-obligations**. After analysing these dependencies, we extracted a total of 48 obligations and 105 dependent sub-obligations.

As the sub-obligations cannot exist independently, they are always associated with an obligation. An obligation may have more than one associated sub-obligation and as such compliance is dependent on the satisfaction of all the sub-obligations. Similarly, some sub-obligations are associated with more than one obligation. For instance, *transparency* (coded as a sub-obligation) is related to the obligation of providing info, providing several rights etc. Considering all these relations between obligations and sub-obligations we identified a total of 310 such relations.

[1] Art. 12 para. 1:«*The controller shall take appropriate measures to provide any information referred to in **Articles 13 and 14** and any communication under **Articles 15 to 22 and 34** relating to processing to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language*»

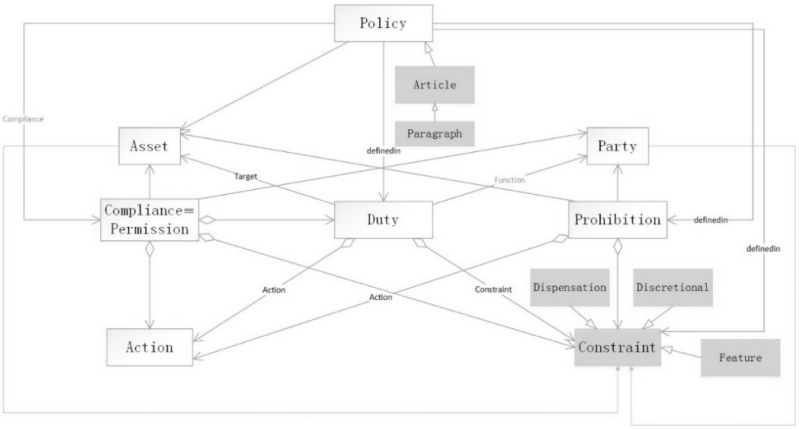


**Figure 2: The modified ODRL model for the GDPR**

Unfortunately some of the paragraphs of these articles cannot be modelled as simple obligations or rules. The normative text also discusses recommended actions which are not obligatory as well as scenarios which are exempted from certain obligations. As such, we refer to the optional actions as Discretional and exemptions as Dispensation. Using *Privacy icons* (Art 12 para 7) is an example of discretional actions and *processing for legitimate interests* can be considered as a dispensation for the obligation of *obtaining consent*. Thus, use of privacy icons is not compulsory and if processing is done for legitimate interests then there is no obligation to obtain consent from data subjects.

## 3. Translating the model in a machine-readable format

After modelling, we translate the model into a machine readable format. For this task we chose an ontology proposed by the World Wide Web Consortium's (W3C) ODRL working group (W3C 2015) that has been designed to define rules for the publishing, distribution, and consumption of digital media. The ODRL ontology is modular i.e. defines a core ontology (abstract model) and provides options to develop profiles (instances) for specific applications, both of which are described below.

The *Abstract model* has the following components – a Policy is the central entity which describes the policy expression. It can either grant Permissions or Prohibitions. Permission may have a Duty as a condition which must be fulfilled in order to be permitted to perform Action on an Asset. A Party is an entity which participates in policy related transactions. It can either be an assigner or an assignee. Also, an Asset is anything which can be subject to a policy. The asset can have a Constraint that is associated with restricting or limiting the scope of the defined rule.

We define the *GDPR instance* as follows: The GDPR is represented as an instance for Policy, while personal data is an instance of the Asset class. Similarly, the controllers, processors, data subjects, supervisory authorities can be expressed as instances of the Party class and data processing as an instance of the Action class whereas GDPR obligations are defined as instances of the Duty class. Using the existing ODRL model, it is difficult to express: 1) the discretional (optional) duties, 2) dispensation scenarios, 3) Sub-obligations which are not independent and 4) constraints related to Action, Asset and Party. As such, we propose minor modifications in the core model, as shown in figure 2 (highlighted in yellow). Addition of the following subclasses for the Constraint class would tackle the first three issues – 1) Discretional for the duties which are recommended but not obligatory, 2) Dispensation for modelling the exception scenarios, and Feature for the sub-obligations. For the fourth issue, we define additional relations between Asset-Constraint, Party-Constraint and Action-Constraint. For the ODRL vocabulary, if all duties are fulfilled then a Permission is granted. For our scenario,

if duties are fulfilled then compliance is attained, thus, we define the Compliance class equivalent to the Permission class. Also, for clarity reasons we add Article and Paragraph subclasses to the Policy class.

## 4. Related Work

ODRL in the literature has been mainly utilised for applications like access control and licensing where Llorente et al. have proposing a solution which uses the ODRL ontology to make data on social media more secure [Llorente/Rodriguez/Delgado 2010] and Çelikbaş demonstrating the licensing of ebooks for use in digital libraries [Çelikbaş 2011]. Dahlem et al. in their paper on Application Rights Management architecture extended ODRL by defining fine-grained application rights in the field of software licensing [Dahlem/Dusparic/Dowling 2004]. Similarly, Steyskal et al. discuss access policies for linked data [Steyskal/Polleres 2014] and Arnab et al. define protocols for Digital Rights Management (DRM) license negotiation using ODRL [Arnab/Hutchison 2007]. ODRL, therefore has been utilised extensively in the field of digital media for access control related applications. In the field of data protection, Korba et al. have proposed a Privacy Rights Management (PRM) model based on the concept of DRM [Korba/Kenny 2002]. Except their paper, to best of our knowledge, no attempt has been made to use ODRL ontology for modelling data protection related legal knowledge.

Our model currently focuses on defining machine readable obligations for controllers and processors that can be utilised for various applications. For instance, systems for checking compliance can benefit from such a model as they would not have to recreate the knowledge based on the GDPR. The model also has applications in the field of legal answering systems like CLIME [Winkels et al. 1998] where, for example, the users of the application can query the applicable duties that need to be fulfilled by their IT application. Additionally, as the model can also be used with other regulations, the applications are not restricted to the GDPR but can incorporate other data protection guidelines as well.

## 5. Conclusions and Future Work

By using the ODRL ontology we model the obligations defined in the regulation and represent it in a machine readable format. Our model based on 31 articles has a total of 313 instances or nodes and 810 defined edges or relations. Using this model for obligations, we are currently implementing a compliance checking tool. For future work, we plan to extend the model further in order to model the consequences and fines in the case of non-compliance in order to use the model for other applications like Privacy Impact Assessments.

## 6. References

Alapan Arnab/Andrew Hutchison (2007), DRM use license negotiation using ODRL v2.0, in: Proceedings of the 5th International Workshop for Technical, Economic and Legal Aspects of Business Models for Virtual Goods.

Zeki Çelikbaş (2011), EPUB use in digital libraries: Developing an online epub creator application, in: Digital Publishing and Mobile Technologies, 120.

Dominik Dahlem/Ivana Dusparic/Jim Dowling (2004), A Pervasive Application Rights Management Architecture (PARMA) based on ODRL, in: ODRL Workshop, 45–63.

W3C ODRL Group (2015), ODRL Version 2.1 Core Model. www.w3.org/community/odrl/two/model/.

Larry Korba/Steve Kenny (2002), Towards Meeting the Privacy Challenge: Adapting DRM, in: Digital Rights Management: ACM CCS-9 Workshop, DRM 2002.

Silvia Llorente/Eva Rodriguez/Jaime Delgado (2010), Secure Management of Social Networks Applications Data, in: Proceedings of the 8th International Workshop for Technical, Economic and Legal Aspects of Business Models for Virtual Goods.

Simon Steyskal/Axel Polleres (2014), Defining Expressive Access Policies for Linked Data Using ODRL Ontology 2.0, in: Proceedings of the 10th International Conference on Semantic Systems, 20–23. SEM'14.

Radboud Winkels/Alexander W.F. Boer/Joost Breuker/Doeko Bosscher (eds.) (1998), Assessment based legal information serving and cooperative dialogue, in: CLIME vol 98.