

# DPMF: A Modeling Framework for Data Protection by Design

Laurens Sion<sup>\*,a</sup>, Pierre Dewitte<sup>b</sup>, Dimitri Van Landuyt<sup>a</sup>, Kim Wuyts<sup>a</sup>, Peggy Valcke<sup>b</sup>, Wouter Joosen<sup>a</sup>

<sup>a</sup> imec-DistriNet, KU Leuven, Belgium, [firstname.lastname@cs.kuleuven.be](mailto:firstname.lastname@cs.kuleuven.be)

<sup>b</sup> imec-CiTiP, KU Leuven, Belgium, [firstname.lastname@kuleuven.be](mailto:firstname.lastname@kuleuven.be)

**Abstract.** *Building software-intensive systems that respect the fundamental rights to privacy and data protection requires explicitly addressing data protection issues at the early development stages. Data Protection by Design (DPbD) – as coined by Article 25(1) of the General Data Protection Regulation (GDPR) – therefore calls for an iterative approach based on (i) the notion of risk to data subjects, (ii) a close collaboration between the involved stakeholders, and (iii) accountable decision-making.*

*In practice, however, the legal reasoning behind DPbD is often conducted on the basis of informal system descriptions that lack systematicity and reproducibility. This affects the quality of Data Protection Impact Assessments (DPIA) – i. e. the concrete manifestation of DPbD at the organizational level. This is a major stumbling block when it comes to conducting a comprehensive and durable assessment of the risks that takes both the legal and technical complexities into account.*

*In this article, we present DPMF, a data protection modeling framework that allows for a comprehensive and accurate description of the data processing operations in terms of the key concepts used in the GDPR. The proposed modeling approach accommodates a number of legal reasonings and assessments that are commonly addressed in a DPIA exercise (e. g., the compatibility of purposes). The DPMF is supported in a prototype modeling tool and its practical applicability is validated in the context of a realistic eHealth system for a number of complementary development scenarios.*

**Keywords.** GDPR • data protection by design • privacy by design • data protection impact assessment • privacy impact assessment • accountability • compliance • architecture viewpoint

Communicated by Judith Michael. Received 2019-12-30. Accepted after 3 revisions on 2020-09-22.

## 1 Introduction

Addressing privacy and data protection issues at the software design stage – rather than adding a clunky layer of legal compliance to a near-final system – is increasingly recognized as the right approach to ensure durable, efficient conformity with data protection law. This has recently been acknowledged by the EU General Data Protection Regulation (GDPR) (European Union 2016)

\* Corresponding author.

E-mail. [laurens.sion@cs.kuleuven.be](mailto:laurens.sion@cs.kuleuven.be)

This research is partially funded by the Research Fund KU Leuven and the PRiSE KU Leuven-C2 research project.

Note: This paper revises and extends Sion et al. (2019a).

which obliges controllers to adopt a proactive stance when both designing (*by Design*) and setting up (*by Default*) their processing operations. Indeed, Art. 24(1) requires controllers to “*implement appropriate technical and organisational measures to ensure and demonstrate compliance with the Regulation*” while Art. 25(1) requires them to do so “*both at the time of the determination of the means for processing and at the time of the processing itself*”.

An efficient Data Protection by Design (DPbD) approach inherently involves assessing the compliance of a given set of processing activities, and this commonly takes on the form of a Data Protection

Impact Assessment (DPIA). Bieker et al. (2016) describe that such a DPIA typically consists of: (i) describing the processing operations, (ii) identifying and documenting the risks to data subjects' rights and freedoms, (iii) implementing appropriate mitigations, and (iv) ensuring accountability by documenting this process. In that sense, DPbD and DPIA essentially share the same approach in that they always start with the description of the system at stake and involve the identification and mitigation of non-compliance issues based on the risks posed by the processing operations.

The execution of a DPIA suffers from the following three practical issues: (i) the description of the system which lays the groundwork for the risk analysis is usually built using an unharmonized legal lexicon; (ii) there are no guidelines or best practices in terms of soundness or completeness of these system descriptions; and (iii) they are usually performed manually which requires tremendous effort, can lead to human errors, and is hard to keep up-to-date with changes to the system.

Many methodologies and tools have emerged to support DPIA, ranging from simple template- and questionnaire-based approaches, to dedicated modeling frameworks. However, in their experience-based report on the adoption of model-based approaches, Torre et al. (2019) highlight the lack of support for constrained modeling and compliance checking as two major impediments to the further adoption of model-based approaches. Additionally, Ferrá et al. (2020) identify the need for a common and unambiguous language and support for DPIAs as "living documents". Although significant efforts have been made to support the systematic modeling of data processing activities in terms of GDPR-related concepts (Blanco-Lainé et al. 2019; Palmirani et al. 2018; Tom et al. 2018), these attempts fall short of providing support for (i) the relevant criteria that are common in DPIAs and (ii) the inclusion of legal rationale and the construction of legal arguments.

In this paper, we present a systematic, model-driven approach called "Data Protection Modeling Framework" (DPMF), which is based on a

meta-model developed through intensive interdisciplinary collaboration between legal and model-driven engineering experts, and therefore provides extensive coverage of the key concepts and abstractions outlined in the Article 29 Working Party (2017) Guidelines on DPIA. As a result, it relies on a conceptual framework that is explicitly defined in the text of the Regulation. This is complemented by a comprehensive set of automated and semi-automated compliance assessments that are implemented on top of the model elements.

The proposed approach does not replace nor eradicate the necessity to deploy proper legal reasoning, but rather provides a streamlined way to inject such argumentation (e. g., the compatibility assessment required under the purpose limitation principle) in a model-based representation of the processing activities. As a result, they make out an integral part of the model throughout its refinement and evolution over time. Consequently, they will also explicitly be included in the different outputs (e. g., records of processing operations).

The DPMF is supported in a tool prototype that builds upon the Eclipse Modeling Framework (EMF) and leverages model-based pattern matching (VIATRA) to implement the legal assessments. We validate the proposed modeling framework in the context of a real-world, IoT-based eHealth application for monitoring patients diagnosed with cardiovascular diseases.

In summary, the main contributions are:<sup>1</sup> (i) a comprehensive set of model and soundness constraints and legal assessments drawing upon the requirements stemming from the GDPR and in support of performing DPIAs; (ii) the explicit formulation and inclusion of legal rationale in the model itself; and (iii) the validation of these contributions by applying the prototype tool to a

<sup>1</sup> This paper is based on Sion et al. (2019a) which is extended with: (i) a more elaborate motivation, based on an in-depth analysis of the state of the art; (ii) an improved version of the meta-model to address a number of challenging cases; (iii) a comprehensive overview of the supported model-based assessments; (iv) more information about the tool prototype of the modeling framework; and (v) an in-depth scenario-based validation and evaluation of the approach.

real-world application case and demonstrating the value of tool-supported construction and model verification, as well as document generation.

The remainder of this paper is structured as follows. First, Sect. 2 introduces the necessary GDPR concepts and motivates the paper with an extensive assessment of the state of the art. Next, Sect. 3 presents the meta-model for constructing Data Protection Models (DPMs) as well as the model and soundness constraints. Sect. 4 then discusses the model-driven implementation of the various legal assessments that can be performed on these DPMs. Then, Sect. 5 details the proposed methodology to build DPMs and apply the said constraints and legal assessments. Subsequently, Sect. 6 presents the prototype implementation of the DPMF. Afterwards, Sect. 7 validates the presented models on a realistic eHealth application. Sect. 8 then discusses specific aspects of the DPMF while detailing areas for future work. Finally, Sect. 9 concludes the paper.

## 2 Background and Motivation

This section provides the necessary background for this paper. First, Sect. 2.1 shortly summarizes the role and requirements for Data Protection Impact Assessments (DPIAs) as stipulated in EU data protection law. Then, Sect. 2.2 discusses the current state of the art in existing DPIA approaches and tools. Based on this, Sect. 2.3 discusses the findings and motivates this work.

### 2.1 EU data protection law

Data protection law requires *controllers* – and to a lesser extent *processors* – to comply with the numerous provisions of the GDPR when it comes to the *processing of personal data*. These crucial notions are explicitly defined by the Regulation and circumscribe its scope of application.

#### Scope of application

The first step of a traditional compliance exercise consists in identifying the involved actors and relevant processing activities to delineate where conformity with data protection rules is required. Art. 4(7) defines the *controller* as “*the natural or*

*legal person which, alone or jointly with others, determines the purposes and the means of the processing of personal data*”, while Art. 4(8), defines the *processor* as “*the natural or legal person which processed personal data on behalf of the controller*”. Most of the GDPR mainly impacts the former, while the latter faces limited obligations with regard to security, liability, breach notification, and its interactions with other actors.

Under Art. 4(2), *processing* is understood as “*any operation or set of operations which is performed on personal data or on sets of personal data*”. This is a broad definition in the sense that it encompasses everything that can be done with personal data, from their collection to their erasure. Personal data is defined by Art. 4(1) as “*any information relating to an identified or identifiable natural person*”.

#### GDPR principles

Every *controller* and *processor* involved in the processing of *personal data* must comply with the general principles of Art. 5, namely: (a) lawfulness, fairness and transparency, (b) purpose limitation, (c) data minimization, (d) accuracy, (e) storage limitation, (f) integrity and confidentiality, and (g) accountability. Most importantly, the collection of personal data must be paired with a specific, explicit, and legitimate purpose and be based on one of the lawful grounds listed in Art. 6(1). Every subsequent processing of the data must, in turn, *not be incompatible*<sup>2</sup> with the purposes for which they were initially collected.

#### Data Protection by Design (DPbD)

Art. 25(1) introduces the obligation for controllers to consider data protection issues right from the

<sup>2</sup> As emphasized by the Article 29 Working Party (2013a) opinion: “*rather than imposing a requirement of compatibility, the legislator chose a double negation: it prohibited incompatibility. By providing that any further processing is authorised as long as it is not incompatible, it would appear that the legislators intended to give some flexibility with regard to further use. Such further use may fit closely with the initial purpose or be different. The fact that the further processing is for a different purpose does not necessarily mean that it is automatically incompatible.*”

design, rather than retrofitting compliance measures into a final system. As such, it substantiates the shift to a risk-based approach towards data protection and ensures a degree of flexibility for controllers when implementing safeguards.

Art. 24 delineates the five components of DPbD, namely: (i) a risk-based approach that requires controllers to tailor their compliance efforts to the risks posed by the processing operations to data subjects' rights and freedoms, (ii) the obligation to ensure compliance with all the GDPR requirements, (iii) the implementation of both technical and organizational measures, (iv) the need to demonstrate that the processing is performed in accordance with the above-mentioned rules, and (v) the necessity to take those considerations into account both at the time of the determination of the means for processing and at the time of the processing itself.

#### **Data Protection Impact Assessment (DPIA)**

In order to achieve compliance with all the above, controllers usually rely on DPIAs. Not only is this exercise mandatory in case of processing activities with a high risk to data subjects' rights and freedoms (Art. 35), it also lays the groundwork for a sound risk-based approach. Even for cases that do not formally require a full-fledged DPIA, the threshold assessments required to determine the necessity of a DPIA are based on a comprehensive description of the different data processing activities. The GDPR indeed encourages the adoption of "appropriate" measures that are proportional to the likelihood and severity of the risks for data subjects' rights and freedoms. Quantifying that risk allows controllers to tailor the scope of their compliance duty (Art. 24(1)), implement data protection by design (Art. 25(1)), and address the security aspects related to their processing activities (Art. 32).

As highlighted above, a thorough DPIA starts with a description of the processing operations, including the involved actors, data subjects, and types of personal data. In turn, comprehensively capturing this information enables two types of reasonings. First, soundness criteria can be applied

to verify the consistency of the system description with the relevant rules (e. g., the need to pair every collection activity with a processing purpose and a lawful ground). Second, traditional legal assessments can be applied to a specific part of the processing activity (e. g., the compatibility and necessity assessment required by, respectively, the purpose limitation and lawfulness principles).

#### **2.2 State of the art in DPIA frameworks and approaches**

Conducting a DPIA is far from trivial, and numerous templates, guidelines, recommendations, and methodologies have been proposed to assist controllers in this task. This section outlines the existing state of the art and distinguishes between (i) recommendations and approaches suggested by National Supervisory Authorities (NSAs), (ii) dedicated legal literature, (iii) available commercial tools, and (iv) specific literature on privacy and security requirements engineering.

##### **Guidance from NSAs**

Most NSAs have released guidelines and templates for conducting DPIAs and these mainly rely on the guidelines from the European Data Protection Supervisor (2018) and the Article 29 Working Party (2017). Wright et al. (2014) have highlighted the necessity of studying NSAs methodologies separately in their comparative benchmark study. A non-exhaustive list of the most notable ones is discussed below.

In its PIA Manuals (CNIL 2018b,c,d), the French Commission National de l'Informatique et des Libertés (CNIL) has presented a straightforward, four-step PIA methodology based on ANSSI's EBIOS risk management framework (Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) 2010). They distinguish between the description of the context of the processing, the identification of existing or planned controls, the assessment of the privacy risks, and the decision to validate the entire PIA process. Furthermore, they provide templates, knowledge bases, and a user-friendly, template-based tool (CNIL 2018a) in support

of the performance of a thorough DPIA. The software facilitates the entire process by providing contextual knowledge based directly on the GDPR, the PIA guides, and their security guide.

The German Standard Data Protection Model (SDM) (ULD 2017) has outlined a series of data protection goals and paired these with generic reference measures that have been tried and tested in data protection investigations and audits, and that may be used to ensure compliance with these principles. The ULD describes a 5-step methodology in which its SDM could fit (analysis of the processing context, substantive legal assessment, characteristics of the protection goals, variance analysis or target plan/comparison relative to the respective authority, and feedback) (Bitkom 2017).

The Belgian NSA (GBA/APD) has issued a recommendation (APD 2018b) listing the essential elements of a DPIA according to Art. 35(7), namely: (i) a description of the processing operations and their purposes, and (ii) an assessment of the necessity and proportionality in relation to those purposes and an assessment of the risks to the data subjects' rights and freedoms and the measures to address these risks. They also offer a template to document the processing activities and provide relevant recommendations (APD 2018a).

### Legal literature

Alnemr et al. (2015) have developed a questionnaire-based approach specific to cloud applications, the main outcome of which is the establishment of a privacy score. Bieker et al. (2016) have outlined a full-fledged methodology to implement Art. 35, which distinguishes between preparation, evaluation, and reporting.

Oetzel and Spiekermann (2014) have proposed a seven-step PIA methodology based on the risk assessment method of the German Federal Office for Information Security (BSI). Their approach distinguishes between the characterization of the system, the definition of privacy targets, the evaluation of the degree of protection required for each target, the identification of threats, the identification and recommendation of new and existing controls suited to protect against these threats,

the assessment and documentation of the residual risks and the documentation of the PIA process. Recommendations are made in terms of the characterization of the system in terms of four complementary views (system, functional, data and physical environment view). However, explicit modeling support is lacking.

Palmirani et al. (2018) have presented PrOnto, a legal ontology summarizing the concepts and relations inherent to the GDPR. Processing activities are represented as workflows comprised of multiple steps, and explicit modeling support is foreseen in a dedicated language (LegalRuleML). This legal ontology is meant as a generic meta-structure for capturing essential GDPR notions and relations but is not explicitly tailored to the performance of a full-fledged DPIA.

### Privacy requirements engineering

In the domain of requirements engineering, numerous approaches to systematically list privacy and data protection requirements have been proposed, with varying degrees of consideration for the applicable regulatory framework.

Starting from the legal perspective, Breaux and Antón (2008) and Breaux et al. (2006) have looked into natural language processing to extract technical requirements from regulatory objectives. Islam et al. (2010) have proposed a framework to assist the elicitation and management of security and privacy requirements starting from the relevant legislative frameworks. Siena et al. (2009) have created a conceptual meta-model to elicit of law-compliant system requirements. Meis et al. (2015), on the other hand, have developed a taxonomy of requirements derived from the principle of transparency. Acknowledging the interpretable nature of the law, Muthuri et al. (2016) have built on the recent development in legal informatics to suggest an interpretative process to orient the acquisition and specifications of legal requirements. For meeting such requirements, Compagna et al. (2009) have proposed a framework to model security and privacy patterns. Finally, Colesky et al.

(2016) discuss privacy design strategies and corresponding tactics for those strategies in order to realize privacy by design.

Security- and privacy-driven software engineering starts in the early stages of the software development life cycle (Howard and Lipner 2006) with activities of, respectively, threat modeling (Shostack 2014), assessment (Brüggemann et al. 2016; Joyee De and Le Métayer 2016) and mitigation. The LINDDUN (Deng et al. 2011) privacy threat modeling framework supports the systematic elicitation and mitigation of privacy threats. Conducting such threat analyses starts from an architectural model of the system, more specifically a Data Flow Diagram (DFD) (DeMarco 1979), that models how data flows through the system and is commonly used in this context (Deng et al. 2011; Dhillon 2011; Howard and Lipner 2006; Shostack 2008, 2014). The DFD notation is based on five distinct element types and, due its relative simplicity, is broadly applicable. As the model lacks much of the necessary legal information, it is not suited to perform more advanced analyses.

DFD extensions have been proposed in the literature (Berger et al. 2016; Sion et al. 2018a,b; Tuma et al. 2017) to provide information on security or privacy solutions (Sion et al. 2018b) in order to take existing countermeasures into account and enable the up-front elimination of non-applicable or already mitigated threats. Further extensions include adding risk assessment information such as asset values, countermeasures strengths, and explicit attacker models (Sion et al. 2018a) to enable a full-fledged risk analysis of the resulting security and privacy threats.

Oliver (2014) has extended DFD modeling with an ontological approach that uses a number of legal concepts but lacks support for lawful grounds and does not have (publicly available) tool support. PA-DFD (Antignac et al. 2016) explicitly introduces data protection-specific concepts derived from the GDPR and ISO 29100. This includes several relevant notions such as purpose, personal data, and storage period, but it does not support notions such as lawful grounds, representatives, and assets.

Petri-net-based approaches (Rahman 2017) have been used to model the business flows and algorithms within processes, similar to an activity diagram. They can theoretically be used to model the system in its entirety. Given their limited building blocks (i. e. states and transitions of a process, connected by arcs), there is no formal possibility to make certain concepts mandatory in the system.

Muntés-Mulero et al. (2019) have built on threat modeling frameworks such as STRIDE and LINDDUN – that both rely on DFD notations for the description of the system – to develop a continuous risk modeling approach that specifically supports risk-aware, trustworthy IoT systems.

Blanco-Lainé et al. (2019) have provided an enterprise architecture approach based on ArchiMate, in which the GDPR principles are encoded in seven complementary architectural goals. These are then further refined in a goal-oriented fashion until specific privacy-enhancing technologies and services have been selected. The impact of these architectural decisions is then further modeled in complementary models (e. g., business processes).

### Modeling approaches

Not focused on DFDs, but also on enriching models with privacy-related information, Ahmadian et al. (2018a,b) have proposed a UML extension for privacy that includes stereotypes for «sensitiveData», granularity, objectives and ABAC (attribute based access control) and privacy preferences purpose, visibility, granularity, and retention. Tool support is available with the integration in CARiSMA (Ahmadian et al. 2018a).

Alshammari and Simpson (2018) have presented APDL, a UML profile dedicated to expressing a system in support of privacy compliance. Their meta-model allows the descriptions of processing activities, actors, and data types, and also supports the notion of purpose. Based on the Object Constraint Language (OCL), a number of legal verifications have been implemented and operationalized (e. g., purpose limitation).

Similarly to DFDs, Fotiou et al. (2014) have proposed the use of Information-Centric Networking (ICN) models as input for a privacy analysis. An ICN network consists of data owners (which are in control of the data, and hence map to controller), consumers (recipients), storage nodes (data sources), resolvers (processes), and two information containers: data flows and data pools.

The CAIRIS platform implements the IRIS (Integrating Requirements and Information Security) approach for identifying security-related requirements. Coles et al. (2018) have extended this framework with explicit support for DPIA activities. As such, they offer a comprehensive modeling approach based on a meta-model that allows to express key concepts related to DPIA activities. Based on the created models, they support the model-based verification of a number of relevant assessments (such as the lawfulness).

While not directly aimed at conducting DPIAs, Tom et al. (2018) have presented a conceptual model of the GDPR-related principles and concepts. Their modeling approach enables the description of data processing activities and the expression of data subject's rights, but is preliminary in nature and lacks, for example, support for the specification of compliance rules and their enforcement.

Ghanavati et al. (2014) have suggested an approach to deduct goal-oriented requirements from regulations which could be applied in a similar fashion as in the work of Blanco-Lainé et al. (2019). Since it precedes the GDPR, however, it cannot be adopted as-is, without refinements.

Agarwal et al. (2018) have developed a generic legislative compliance assessment framework that can support various legislative frameworks. Their modeling approach mainly focuses on text-based analysis and data modeling, augmented with boilerplates and templates. The end result is a set of policies encoded and enactable in Open Digital Right Language (ORDL) format. While the main contribution of this work is the development of a promising, end-to-end, and technically integrated approach, their coverage of GDPR-related concepts and notions is still lacking.

The experience report of Torre et al. (2019) provides valuable insights in terms of their attempts to automate a number of compliance assessments, using the UML and OCL constraints. While this is a promising approach, this is in essence a preliminary technology feasibility study and these results are not sufficiently elaborate to be adopted as-is in the implementation of a DPIA.

### Commercial solutions

The GDPR and the ensuing necessity to perform a DPIA to avoid fines has raised awareness and led to the development of numerous commercial offerings. Although details on these initiatives are generally not publicly available – and thus fair comparison with other approaches is impossible – some of those are shortly discussed below.

Nymity (2019) provides dedicated tooling for DPIAs. Their approach is based on questionnaires and the generation of suitable accountability documentation is explicitly supported. In-depth information on the level of support of, for instance, the key-abstractions and steps listed by the Article 29 Working Party is, however, not publicly available.

AvePoint (2019) Privacy Impact Assessment is also based on a form-based questionnaire system and supports the generation of PIA reports based on document templates.

RealDPG (2019) assists organizations in building and documenting a comprehensive processing register and to perform data protection risk assessment on all processing involving risks to data subjects. It also allows for compliance proofing and visualization of data flows.

OneTrust (2019) Privacy Management is a fully integrated privacy management platform in support of GDPR compliance. OneTrust offers, amongst others, a readiness and accountability questionnaire-based tool designed to assess the level of compliance of an undertaking with regard to a given legislative framework and that allows for easy reporting. The OneTrust platform also facilitates compliance with Art. 30 GDPR by providing customizable templates to record processing activities and automating the data mapping phase.

Niobase (Akarion AG 2019) implements a comprehensive modeling approach to describe the processing activities and augments this support with data flow visualization and risk analysis.

### 2.3 Findings and motivation

The previous section clearly demonstrates that many different approaches and methodologies have emerged in support of DPIA and, indirectly, of DPbD. A comprehensive overview of our analysis of the current state of existing approaches and tools is provided in Tab. 1.

The second column summarizes the nature of the support specifically aimed at describing of the data processing activities, which ranges from templates, questionnaires and checklists, to more structured modeling approaches. In the third column, an indication is given as to whether or not the proposed approach is embedded in a larger DPIA methodology. The fourth column represents the degree of coverage of the key concepts and abstractions<sup>3</sup> highlighted by the Article 29 Working Party (2017) as essential for the description of data processing activities of a comprehensive DPIA. Columns five and six respectively indicate whether there is any support for evaluating whether the description of the processing activities is complete, sound and consistent (soundness and completeness criteria), and to evaluate the described model in light of the GDPR principles (legal and compliance assessments). Then, the seventh column gives an indication of the state of the tool support underpinning these approaches. The eighth column highlights whether the generation of appropriate accountability documentation is supported, whereas the final column indicates the extent to which model management (e. g., versioning, model evolution) is supported.

#### Findings

As shown, the approaches promoted by NSAs unsurprisingly provide high coverage of these concepts and abstractions, but generally lack tool support and more notably support for the evaluation

<sup>3</sup> These concepts are processing, lawful grounds, purpose, personal data, data subjects, recipients, controllers, processors, representatives, third parties, storage period, and assets.

of model completeness and soundness, and lack support for identifying legal compliance issues.

Approaches originating from the area of requirements engineering provide some support for the core DPIA abstractions typically as an extension to existing modeling paradigms (such as DFDs), and are in some cases more mature in terms of the tooling. However, they generally lack built-in support for the evaluation of both soundness and legal assessments. The DFD-based approaches that explicitly model the processing activities as an extension to DFD models provide preliminary support for *model management*, in the sense that these models are artifacts in the development life cycle that can be stored in version control systems. However, explicit support for differencing between models, versioning at the level of the semantics of these models, evolution over time, and explicitly documenting the rationale and legal reasoning that supports changes to the models is lacking.

Many of the recent GDPR modeling approaches are promising. However, while most of them offer more automated compliance checking, they often lack exhaustive support for these assessments. The generation of DPIA documentation is a key requirement in light of accountability principle imposed by the GDPR. Here, the challenge is to support the generation of documentation tailored to different stakeholders. For example, a record of processing activities that will be made publicly available may deliberately omit information about the inner functioning of the system. As mentioned, several commercial tools are available and these generally support this requirement well but, given the lack of public documentation, this is difficult to verify and compare fairly.

#### Motivation

In light of the above, we highlight the need for a more comprehensive and structured model-driven approach that supports the exhaustive modeling of data processing activities and related information elements in a systematic and structured fashion. The key requirements include support for:

Table 1: Evaluation of existing modeling approaches w.r.t. legal and security/privacy architecture DPbD requirements.

Approach	Description of processing activities	DPIA methodology support	Coverage of Art. 29 WP concepts	Support for soundness criteria (R1)	Tool support	Document generation (R2)	Model generation (R3)	Model management (R4)
<b>National Supervisory Authorities (NSAs)</b>								
France: CNIL 2018a	Template	●●● 90%	●●● ●●○	●●● ●●●	●●● ●●●	●●● ●●●	●●○ ●●○	●●○ ●●○
Belgium: APD 2018a,b	Template	●○○ 90%	●○○ ●○○	○○○ ○○○	●○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
Germany: SDM (ULD 2017)	Unstructured	●●● n/a	○○○ ●○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
<b>Legal research</b>								
Alnemr et al. 2015	Questionnaire	●○○ 58%	●○○ ●●○	●●○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
Bieker et al. 2016	Unstructured	●●● n/a	●○○ ●●○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
Oetzel and Spiekermann 2014	Unstructured	●●● n/a	●○○ ●●○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
PrOnto (Palmirani et al. 2018)	Workflow/ontology	○○○ 67%	●○○ ●●●	●●○ ●●○	●○○ ●○○	●○○ ●○○	●○○ ●○○	●○○ ●○○
<b>Privacy and security requirements engineering</b>								
STRIDE/LINDDUN	DFD models	○○○ 25%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
DFD+dict.(DeMarco 1979)	DFD models	○○○ 42%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
PA-DFD (Antignac et al. 2016)	DFD models	○○○ 75%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
DFD+ontology (Oliver 2014)	DFD models	○○○ 83%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
Privacy-aware DFD (Rahman 2017)	DFD models	○○○ 58%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
Blanco-Lainé et al. 2019	Deliverables	●●○ 25%	○○○ ○○○	●●○ ○○○	●●○ ○○○	●●○ ○○○	●●○ ○○○	●●○ ○○○
<b>GDPR modeling approaches</b>								
CARiSMA (Ahmadian et al. 2018b)	Annotations (UML)	○○○ 50%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
APDL (Alshammari and Simpson 2018)	Annotations (UML)	●○○ 58%	●●● ●●○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○	○○○ ○○○
ICN arch. (Fotiou et al. 2014)	within ICN models	○○○ 25%	○○○ ○○○	●●○ ○○○	○○○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
CAIRIS-DPIA (Coles et al. 2018)	IRIS modeling	●○○ 42%	●●● ●●○	●●○ ●●○	●○○ ●○○	●○○ ●○○	●○○ ●○○	●○○ ●○○
Torre et al. 2019	Model (UML)	○○○ 33%	●○○ ○○○	●●○ ○○○	●○○ ●○○	●○○ ●○○	●○○ ●○○	●○○ ●○○
Tom et al. 2018	Model (UML)	○○○ 79%	●○○ ○○○	●●○ ○○○	●●○ ○○○	●○○ ○○○	●○○ ○○○	●○○ ○○○
Agarwal et al. 2018	Checklist (ORDL)	●●● n/a	○○○ ●●○	○○○ ○○○	●●● ●●●	●●● ●●●	●●● ●●●	●●● ●●●
<b>Commercial tools</b>								
Nymity GDPR (Nymity 2019)	Questionnaire	●○○ ---	--- ---	--- ---	●●● ●●●	●●● ●●●	--- ---	--- ---
APIA (AvePoint 2019)	Questionnaire	●○○ ---	--- ---	--- ---	●●● ●●●	●●● ●●●	--- ---	--- ---
RealDPG 2019	Modeling	●●○ ---	--- ---	--- ---	●●● ●●●	●●● ●●●	--- ---	--- ---
OneTrust 2019	Quest./template	●●○ ---	--- ---	--- ---	●●● ●●●	--- ---	--- ---	--- ---
NioBase (Akarion AG 2019)	Modeling	●●○ ---	--- ---	--- ---	●●● ●●●	●●● ●●●	--- ---	--- ---

Legend: - - -: unclear due to lack of public information, ○○○: unsupported, ●○○: limited, ad-hoc or partial support, ●●○: supported but extensive manual effort required, ●●●: automated support (tools, generators, macros) for criteria.

'Coverage of Art. 29 WP concepts' is based upon and extended from Dewitte et al. (2019) and focuses on built-in support for the concepts and abstractions put forward in the Art. 29 WP DPIA requirements (Article 29 Working Party 2017, Annex 2).

- R1** The automated and semi-automated (i. e. tool-assisted) verification and validation of a comprehensive set of model soundness and legal assessments (those considered common and necessary in the context of a DPIA).
- R2** The construction of legal argumentation and the explicit inclusion of such rationale in a model to ensure consistency and reuse.
- R3** The generation of appropriate accountability documentation that includes the above-mentioned arguments and verification results.
- R4** Model management over time (e. g., keeping track of the process rationale over evolving versions of the model).

In this paper, we present the DPMF, a novel framework to fulfil these key requirements. The next section provides a detailed overview of its support for creating data protection models.

### 3 The Data Protection Model (DPM)

This section presents the meta-model that supports the creation of Data Protection Models (DPMs) for documenting and analyzing the legal implications of data processing operations from the perspective of the GDPR. This section contains three main parts. First, the process of constructing the meta-model is discussed. Second, the key concepts of the meta-model are explained. Given the pervasiveness and complexity of algorithmic decision-making processes and the problem they pose to accountability (Bayamlioglu 2018; Casey et al. 2019), we illustrate those concepts by building the DPM-fragments of a News Recommender System, that we use as a running example throughout the following sections. Third, a number of constraints – some directly enforced by the meta-model – are presented to ensure the construction of sound DPMs that do not represent invalid or inconsistent data processing operations.

#### 3.1 Construction of the meta-model

Fig. 1 depicts the DPM meta-model, which is the result of intensive interdisciplinary collaboration between legal and model-driven engineering

experts to ensure a shared, accurate representation of the processing operations. The presented meta-model is the outcome of iterative refinement.

The terminology used for the modeling concepts follows entirely from the concepts and notions defined in the GDPR. Not only does this guarantee a certain degree of expressiveness of the modeling paradigm, but it also ensures the legibility of the models for external stakeholders in the context of administrative or judicial proceedings.

#### 3.2 Overview of the meta-model

The next subsections provide further detail on using the model in concrete examples to model the *who*, *how*, *what*, and *why* of the processing operations. Every section first explains the key modeling concepts and then provides an illustration contextualizing them in the running example.<sup>4</sup>

Tab. 6 in the appendix provides a legal glossary of all the concepts used in this section.

##### 3.2.1 Modeling the actors

The following modeling concepts are provided for modeling the actors involved in the processing:

###### Core modeling concepts

**Actor** 🗑️: An «Actor» represents an organizational entity involved in the processing of personal data. For each actor, details are provided about whether they are a public authority, an international organization, established in the EU, etc.

**Representative** 🗑️: A «Representative» is a natural or legal person that represents an «Actor» not established in the EU (Art. 27).

**LegalRole** 🗑️: The «LegalRole» specifies in which capacity an actor is involved in a processing activity. The GDPR distinguishes between controllers, processors, recipients, third parties, and, when controllers or processors are not established in the EU, representatives (Art. 4(7–10,17)). Such a qualification is essential for allocating responsibilities under the GDPR. Decoupling an «Actor» from its specific «LegalRole» in the context of a

<sup>4</sup> It is worth noting that the DPMF does not impose an order in using these concepts. The used order of presentation is purely illustrative.

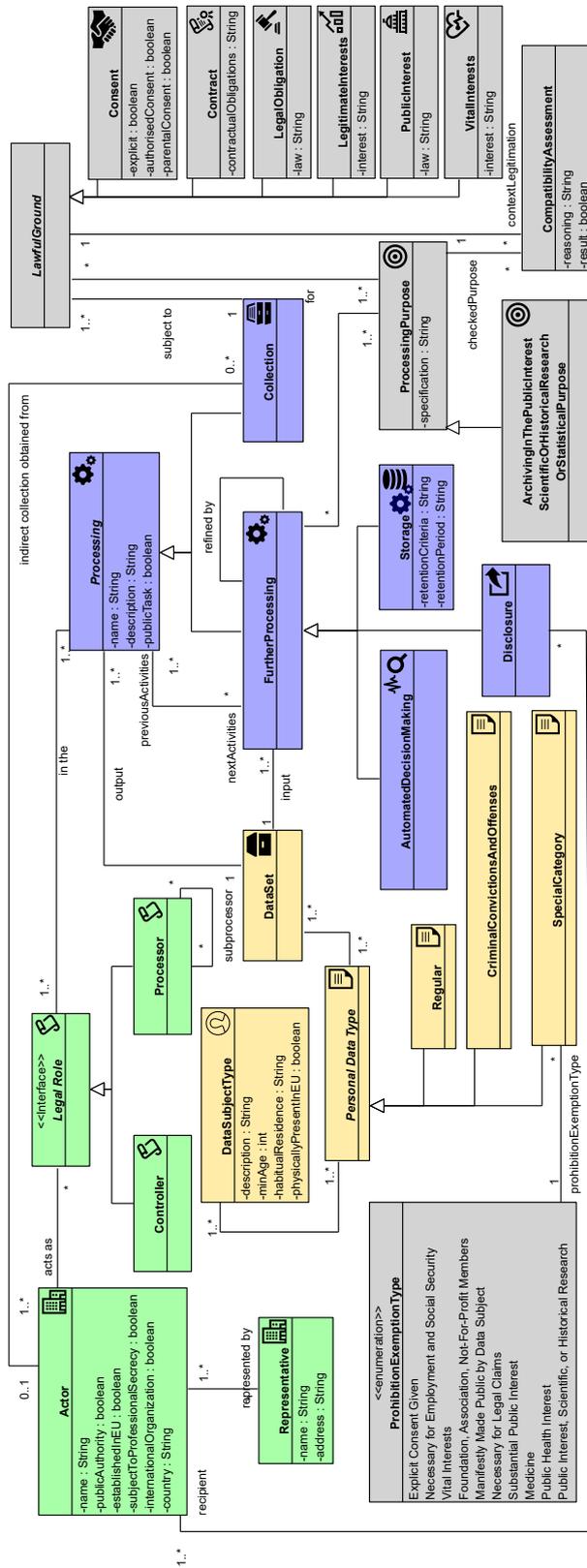


Figure 1: The meta-model for creating Data Protection Models (DPMs).

The corner of each class includes the icon used in the instance model to more easily make a distinction between the different types. Sect. 3.3 lists the additional checks that are performed to ensure the creation of a sound Data Protection Model (DPM). All the colored classes on the left-hand side and center indicate concepts used for modeling the processing operations, the involved data, and the entities performing the processing (the *who* and *what*), while the gray classes on the right-hand side indicate the legal reasoning of processing (the *why*).

processing activity allows an entity to have multiple roles in different processing activities (e. g., being a «Controller» for one processing activity, while acting as «Processor» for another one).

**Controller**  $\mathfrak{S}$ : The «Controller» is the entity that, alone or jointly with others, determines the purposes and means of the processing (Art. 4(7)). The separate «Controller» role supports complex situations such as, for example, joint controllership.

**Processor**  $\mathfrak{S}$ : The «Processor» is the entity that processes data on behalf of the «Controller» (Art. 4(8)).

**Recipient**: A recipient is any «Actor» to which the data are disclosed (Art. 4(9)). It can be a third party or not, depending on whether it qualifies as a «Controller» or «Processor» with regard to the «Processings» at stake. A recipient can also act as a «Controller» or «Processor» in its own right with regard to another set of «Processings». It can be derived from the *recipient* reference to an actor.

**Third Party**: A third party is an «Actor» that is neither a data subject, the «Controller», or the «Processor» with regard to the «Processings» at stake. It can also act as a «Controller» or «Processor» in its own right with regard to another set of «Processings» (Art. 4(10)).

### Illustration

Fig. 2 shows the «Actors» and «LegalRoles» in the context of a news aggregator website to illustrate how the multiplicities in the meta-model support the modeling of complex situations. In this example, *NewsRecommender* «Actor» acts as a «Controller» for a certain set of processing operations and relies on *CloudProvider* acting as a «Processor». Since it is not established in the EU, the *NewsRecommender* organization has appointed *RecommenderRepresentative* as its «Representative». *NewsRecommender* also combines its own data with additional data obtained via *DataBroker*. As part of its processing operations, *NewsRecommender* also discloses the

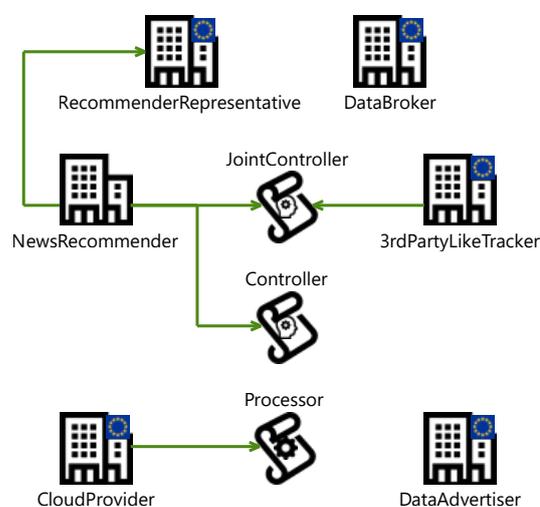


Figure 2: Modeling parties involved in the processing. This diagram illustrates the different «Actors» and their «LegalRoles». For readability purposes, the same graphical notation from Fig. 1 is adopted, instead of a more verbose UML object-instance notation.

data to *DataAdvertiser*. Finally, in joint controllership with *3rdPartyLikeTracker*, *NewsRecommender* also jointly determines the purposes and the means of another set of processing operations to keep track of popular news articles.

### 3.2.2 Modeling the processing operations

The following modeling concepts are introduced for modeling the processing operations:

#### Core modeling concepts

**Processing**  $\mathfrak{A}$ : A «Processing» activity is any operation performed on personal data by the actors listed above (Art. 4(2)). The following subtypes can be used for a more detailed specification.

**Collection**  $\mathfrak{B}$ : Every processing activity needs to start with an initial «Collection» of personal data – either directly from the data subject or from another actor (Article 29 Working Party 2013a). This is enforced by the meta-model by requiring every «FurtherProcessing» to link to a previous «Processing» to create a chain. Only a «Collection» at the start does not specify a previous one. This link of «Processings» implies the reuse of the personal data collected or processed by previous «Processings» in the chain.

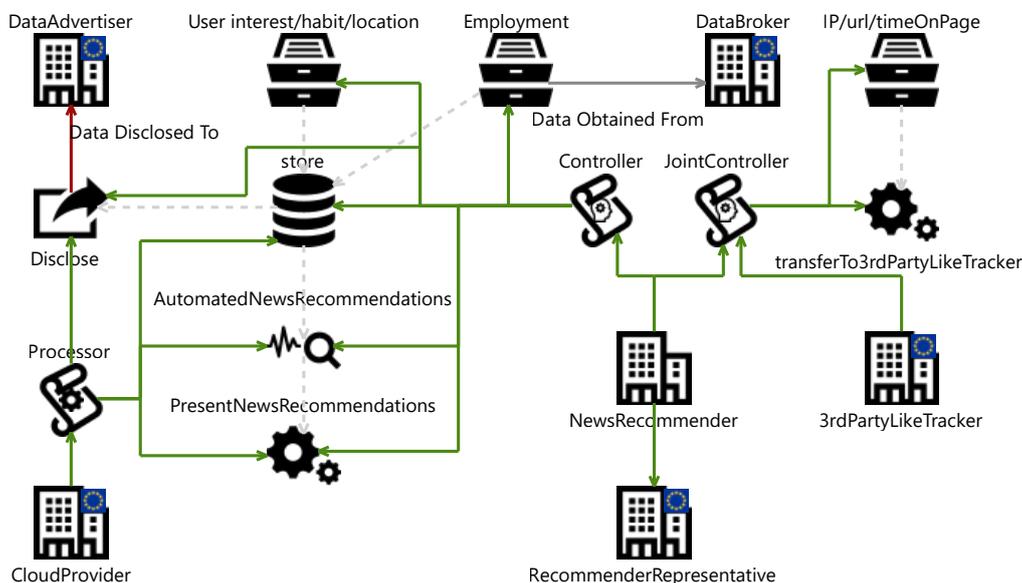


Figure 3: «Processings» and «Actors» in the news recommender example.

This diagram illustrates the «Processings» for a news recommender system. It starts with three initial «Collections», followed by several «FurtherProcessings» (including «Storage», «AutomatedDecisionMaking», and «Disclosure»). The dashed arrows on the diagram point to the *next* «Processing» activities.

**FurtherProcessing**  $\ast$ : Every processing after the initial «Collection» is a «FurtherProcessing». They can branch into multiple «FurtherProcessings» that are performed in isolation from each other. There are three subtypes with additional legal implications:

**Storage**  $\ast$ : A «Storage» activity requires specifying either a *retentionPeriod* for the data or the *retentionCriteria* for determining how long they will be kept (Art. 13(2)a; 14(2)a).

**AutomatedDecisionMaking**  $\ast$  $\alpha$ : This is defined quite narrowly as a “*decision based solely on automated processing, including profiling, which produces legal effects concerning the data subject or similarly significantly affects him/her*” (Art. 22(1)). It has implications on the transparency obligations (Art. 13(2)f; 14(2)g; 15(1)h), the type of lawful grounds on which the controller has to rely (Art. 22(2)), the measures that must be implemented (Art. 22(3); Rec. 71) and the type of data that may be processed (Art. 22(4)). Regardless of the (numerous) legal controversies

surrounding the exact scope of this notion (Goodman and Flaxman 2017; Malgieri and Comandé 2017; Selbst and Powles 2017; Wachter et al. 2017), the meta-model supports this concept as well as some of those rules.

**Disclosure**  $\ast$ : A «Disclosure» represents an explicit activity of disclosing personal data to a *recipient*. This requires specifying to which «Actors» the data are disclosed. Keeping track of that information is particularly useful to facilitate compliance with the transparency obligations (Art. 13(1)e; 14(1)e).

### Illustration

Fig. 3 builds on the example provided in Fig. 2 and expresses the sequence of «Processings» for the personalization of items on a news aggregator website. In this example, there are three chains of «Processings». The first is initiated by *NewsRecommender* acting as «Controller» and is comprised of a direct «Collection» and three «FurtherProcessings» in the form of a «Storage», an «AutomatedDecisionMaking» and a «Disclosure» to *DataAdvertiser*. The



### Illustration

Fig. 4 extends the DPM model specified in Fig. 3 with information on the «DataSets», «PersonalDataTypes», and the «DataSubjectType». The three chains of «Processings» each rely on different «DataSets». *NewsRecommender* processes the interests, browsing habits, location history, and political orientation of its users. It also enriches its database with employment details as obtained from *DataBroker*. Moreover, it discloses the interests, browsing habits, location history, political orientation, and employment details to *DataAdvertiser*. Finally, *NewsRecommender* and *3rdPartyLikeTracker* jointly collect the IP address, the visited URLs and the time spent on the web page. The model forces the specification of the collected data as a «DataSet» output of a «Collection» and as a «DataSet» input to other «FurtherProcessings». The «PersonalDataTypes» in turn link to the «DataSubjectType» they pertain to. This ensures that there is an unambiguous and complete specification of the input and output for every «FurtherProcessing».

### 3.2.4 Modeling lawful grounds and purposes

The following modeling concepts are provided for modeling the «LawfulGrounds» and «ProcessingPurposes» of the «Processings»:

#### Core modeling concepts

**LawfulGround:** According to the lawfulness principle, each specified «Collection» activity must be explicitly motivated with one or more specifications of «LawfulGrounds». It must be one of the types listed in Art. 6(1): «Consent» 🗑️, «Contract» 📄, «LegalObligation» ⚖️, «LegitimateInterests» 📊, «PublicInterest» 🏛️, or «VitalInterests» ❤️.

**ProcessingPurpose** 🎯: The purpose limitation principle – more specifically its ‘purpose specification’ component – imposes that each «Processing» is paired with one or more «ProcessingPurposes», be it for a «Collection» through the «LawfulGround» or a «FurtherProcessing» (Article 29 Working Party 2013a) (Art. 5(1)b).

**CompatibilityAssessment** 📄: As specified by Art. 5(1)b, the purpose limitation principle requires that data are not further processed in a manner that is incompatible with the «ProcessingPurpose» for which they were originally collected (Article 29 Working Party 2013a). The «CompatibilityAssessment» documents the outcome of that legal assessment and specifies whether the «ProcessingPurpose» of a «FurtherProcessing» is compatible with the «ProcessingPurpose» of the «Collection».

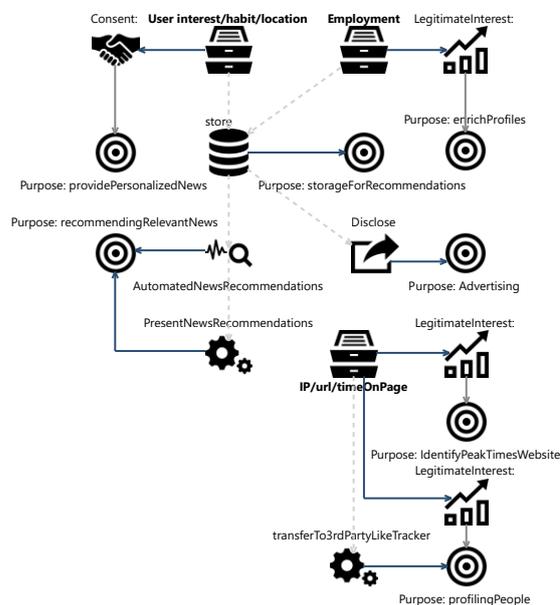


Figure 5: Assigning «LawfulGrounds» and «ProcessingPurposes» in the news Recommender system. This extends Fig. 4 with «LawfulGrounds» and «ProcessingPurposes» assigned to the different «Collections» and «FurtherProcessings».

### Illustration

Fig. 5 further extends Fig. 4 with «ProcessingPurposes» and appropriate «LawfulGrounds». In the example, *NewsRecommender* processes the user interests, etc. on the basis of the «Consent» of the data subjects to personalize their news feed. For that, it collects, stores, and engages in automated decision-making. Thanks to *DataBroker*, *NewsRecommender* also indirectly collects and processes the employment information on the basis of its «LegitimateInterests» to enrich the

profiles it already holds. Similarly, *NewsRecommender* discloses some data to *DataAdvertiser*, on the basis of its «LegitimateInterests» in order to generate extra revenue. Finally, *NewsRecommender* and *3rdPartyLikeTracker* collect and further process analytics data, each of them on the basis of their «LegitimateInterests» respectively to identify peak usage times on the website and profiling the users of the website.<sup>6</sup>

Fig. 5 shows the modeling of the «ProcessingPurposes» and «LawfulGrounds» for each «Collection», as well as the «ProcessingPurposes» of the «FurtherProcessings», which, in turn, link back to the previous «Processing» activities. This traceable link between «Processings» and the original «Collection» of the data is essential to perform the «CompatibilityAssessment» and verify that all the collected data are actually necessary for the processing.

### 3.3 Model constraints

One of the benefits of adopting a systematic modeling approach is that it enables explicit support for completeness and soundness criteria to verify that the model created is correct and sound. The meta-model depicted in Fig. 1 imposes a number of constraints on the model elements via the stated multiplicities between the concepts. These constraints implement a set of rules for obtaining a suitable system description that will serve as a solid basis for a comprehensive compliance exercise. Tab. 7 in the appendix provides an overview of all the constraints discussed below.

#### 3.3.1 Meta-model constraints

The section presents the constraints that are directly enforced by the meta-model with the relations and the multiplicities specified therein.

**Model Constraint 1** *Every chain of processing must start with a collection (Art. 4(2)).*

<sup>6</sup> The present example does not necessarily depict a legally valid situation. While the use of «LegitimateInterests» as a «LawfulGround» for the disclosure of personal data to an advertiser is debatable, the point here is to highlight the support for representing and analyzing this situation.

Every chain of «Processings» needs to start with a «Collection». Every «FurtherProcessing» requires a previous activity except the «Collection». This way, every «FurtherProcessing» activity can be evaluated in light of the original «LawfulGround» and «ProcessingPurpose» for which the data were collected.

**Model Constraint 2** *Every collection must specify a lawful ground and a processing purpose (Art. 5(1)a,b; 6(1))*

The multiplicity of the relation between «Collection» and «LawfulGround» (1..\*) enforces the specification of at least one «LawfulGround» per «Collection», which in turn requires specifying a «ProcessingPurpose» (1..\*).

**Model Constraint 3** *Every further processing must specify a processing purpose (Art. 5(1)b).*

The multiplicity of the relation between «FurtherProcessing» and «ProcessingPurpose» (1..\*) enforces at least one «ProcessingPurpose» for each subsequent activity, which enables the «CompatibilityAssessment».

**Model Constraint 4** *Every processing of special categories of personal data must specify an exemption type (Art. 9(1)(2)).*

The multiplicity between «SpecialCategory» and «ProhibitionExemptionType» (1) forces the specification of an exemption for every «SpecialCategory» data type.

**Model Constraint 5** *Every legal role must specify at least one actor.*

The multiplicity between «LegalRole» and «Actor» (1..\*) forces the specification of at least one «Actor» so that there are no empty «LegalRoles» (i. e. not filled in by a specific «Actor»).

**Model Constraint 6** *Every disclosure must specify the actor(s) that receive the data (Art. 4(9)(10); 13(1)e; 14(1)e).*

A «Disclosure» requires at least one «Actor» to be specified as the *recipient*.

### 3.3.2 Soundness constraints

The constraints listed below are not captured in the relations and their multiplicities but can be expressed as additional constraints on the model.

**Soundness Constraint 1** *Consistency of input and output datasets of personal data types.*

Every «PersonalDataType» used in an input «DataSet» to a «Processing» must be in the output «DataSet» of a «Processing» that is located earlier in the chain of «Processings».

**Soundness Constraint 2** *A non-EU actor must appoint a representative (Art. 27(1)).*

If the «Actor» with the role of «Controller» or «Processor» is *establishedInEU* = FALSE and *publicAuthority* = FALSE, then it must be represented by a «Representative» in the Union.

**Soundness Constraint 3** *A collection must be followed by a further processing (Art. 4(2)).*

Every «Collection» should be followed by one or more «FurtherProcessings». A simple query on the model can verify that there are no «Collections» without «FurtherProcessings».

**Soundness Constraint 4** *Consistency between indirect collection and disclosure.*

When modeling the indirect «Collection» and «FurtherProcessings» that follow from a «Disclosure» to a *recipient*, the «Actor» providing the data for the indirect «Collection» should be consistent with the «Actor» controlling the «Disclosure» providing the data.

**Soundness Constraint 5** *Every processing must specify a controller (Art. 4(7)).*

Every «FurtherProcessing» must have at least a «Controller» in its set of «LegalRoles».

**Soundness Constraint 6** *Every processing with a processor specified, must also specify a controller (Art. 4(8)).*

Every «Processing» which links to a «Processor» must also specify a «Controller».

**Soundness Constraint 7** *A public authority acting in the performance of its public tasks cannot rely on its legitimate interests (Art. 6(1) ind. 2).*

If the «Actor» acting as «Controller» is a public authority (*publicAuthority* attribute set to TRUE) and the «Processing» is a public task (*publicTask* attribute set to TRUE), then the «LawfulGround» of the «Collection» cannot be «LegitimateInterests».

**Soundness Constraint 8** *A controller processing special categories of personal data for preventive or occupational medicine must be subject to professional secrecy (Art. 9(3)).*

If the data type is «SpecialCategory» and the «ProhibitionExemptionType» is *Medicine*, then the «Actor» with the «LegalRole» of «Controller» must be *subjectToProfessionalSecrecy*.

**Soundness Constraint 9** *Every storage activity must specify a retention period or retention criteria (Art. 5(1)e; 13(2)a; 14(2)a).*

Each «Storage» activity must specify either a *retentionPeriod* or the *retentionCriteria* used to unambiguously determine the retention period.

**Soundness Constraint 10** *Every automated decision-making within the meaning of Art. 22(1) must specify contract, legal obligation or explicit consent as lawful ground (Art. 22(1)(2)).<sup>7</sup>*

Each «AutomatedDecisionMaking» must be based on either «Contract», «LegalObligation» or *explicit* «Consent» as a «LawfulGround».

With the appropriate tool support (Sect. 6), the above rules can be implemented and checked when the model is constructed and this in turn allows providing direct feedback to the modeler.

<sup>7</sup> While Art. 22(1) GDPR is phrased as a right granted to data subjects, the WP29 has interpreted it as a general prohibition of «AutomatedDecisionMaking» (Article 29 Working Party 2018a). Concluding otherwise would have resulted in data subjects being protected merely upon request.

## 4 Legal Assessments in Support of DPIA

The meta-model presented in the previous section allows for a detailed description of data processing activities and the key related aspects. On the basis of this model, a wide range of legal assessments can be performed to assist a legal expert in performing DPIAs.

### 4.1 Approach for determining the legal assessments

The legal assessments presented in this section are based on an in-depth analysis of the principles and rules stemming from the GDPR. In order to come up with such a list, the provisions of the Regulation were first analyzed by a legal expert and documented to pinpoint the rules that govern the actual processing of personal data. Each of the relevant provisions was then dissected and expressed as one or more *legal assessment(s)* to allow the modeling framework to automatically apply and enforce the underlying rule.

In collaboration with experts on model-driven engineering, each of those legal assessments can be translated to queries using the modeling concepts of the DPM meta-model to support the operationalization of the rules. These assessments have been described in accordance with the template structure shown below.

**Assessment 0** *Assessment name* <sup>Ⓝ</sup>

*Legal description referring to the GDPR.*

Description using the DPM concepts.

In this template structure, the <sup>Ⓝ</sup> variable expresses the level of support in the DPMF:

- ① A DPM provides little to no information to perform the legal assessment.
- ② A DPM provides relevant information (in terms of the colored concepts of the meta-model) but the legal assessment has to be performed separately and does not involve any change in the DPM.

- ③ A DPM allows extension with legal rationale (in terms of the grey concepts of the meta-model) so that the legal assessment can be performed on the basis of the DPM itself.

Depending on the reader's familiarity with the GDPR, some of the legal descriptions may be skipped or glanced over and referred back to when reading the scenario-based validation (see Sect. 7) that applies some of these legal assessments on a concrete DPM. A complete and comprehensive overview of all the legal assessments discussed in this section can be found in Tab. 8 to 10. Sect. 4.2 discusses assessments that can be supported at the basis of DPM models (assessments ranked either ② or ③) whereas Sect. 4.3 discusses the remaining assessments (ranked ①).

### 4.2 Legal assessments supported by DPM

This section provides, for the identified set of legal assessments, a reference to the corresponding provision(s) in the GDPR and a short description of the approach to implement these assessments using the core modeling concepts of the meta-model. Where necessary, some of the legal assessments are split into multiple *assessments* that provide more details on the overarching principle or rule. The DPMF does not impose any order in applying these assessments on concrete DPMs.

#### Legal Assessment 1 *Lawfulness* <sup>②</sup>

*"Personal data shall be processed lawfully" (Art. 5(1)a; 6(1)).*

*"As long as the further processing is compatible with the purposes for which the data were initially collected, no separate legal basis is necessary" (Rec. 50).*

As enforced by Model Constraint 2, each «Collection» needs to specify a «LawfulGround» and a «ProcessingPurpose». If the «ProcessingPurpose» of the «FurtherProcessing» is *not incompatible* with the «ProcessingPurpose» of the «Collection» (denoted by the «CompatibilityAssessment» objects resulting from Assessment 2.2), no additional «LawfulGround» needs to be specified (i. e. the «LawfulGround» of the «FurtherProcessing» is deemed identical to the one specified for the «Collection» activity at

the start). If it is incompatible, a new «LawfulGround» and «ProcessingPurpose» must be specified for the «Collection».

**Assessment 1.1** *Performance of a contract, legal obligation, vital interests, public task, interest, legitimate interests – Necessity* ②

*When the processing is based on the performance of a contract, compliance with a legal obligation, the protection of the vital interests of the data subject, the performance of a task carried out in the public interest or the legitimate interests of the controller, the processing must be objectively necessary to achieve that (Art. 6(1)b–f).*

If a chain of «Processing» activities is based on any of the aforementioned «LawfulGrounds», the relevant details must be specified and the «ProcessingPurpose» of all the «Processing» activity must be objectively necessary. For example, for a «Processing» based on «Contract», the contractual obligations must be documented and the «Processing» must be objectively necessary to achieve those obligations (European Data Protection Board 2019).

**Assessment 1.2** *Compliance with a legal obligation to which the controller is subject – Existence of Union or Member State law* ②

*When the processing activity is based on either a legal obligation or the performance of a task carried out in the public interest, the applicable Union or Member State law must be specified (Art. 6(3)).*

If a «Processing» is based on either «LegalObligation» or «PublicInterest» as «LawfulGround», the corresponding Union or Member State law must be specified in the *law* attribute of the said «LawfulGround».

**Assessment 1.3** *Legitimate interests of the controller – Balancing test* ②

*When the processing is necessary for the purposes of the legitimate interests pursued by the controller, it is necessary to verify that such interests are not overridden by the interests or fundamental rights and freedoms of the data subject, in particular when the data subject is a child (Art. 6(1)f).*

If a «Processing» is based on «LegitimateInterests» as «LawfulGround», the interests of the controller and the interests or fundamental rights and freedoms of the data subjects must be

documented. If Assessment 1.1 succeeds, it is also necessary to balance those interests against each other.<sup>8</sup> While the DPMF does not automate the balancing test, it can nonetheless streamline this assessment by providing a comprehensive overview of the «Processings», «LawfulGrounds», «ProcessingPurposes», «DataSubjectType» and associated «PersonalDataTypes» to consider.

**Assessment 1.4** *Consent – Attributes* ②

*A valid consent means “any (i) freely given, (ii) specific, (iii) informed and (iv) unambiguous indication of the data subject’s wishes by which he or she, by (v) a statement or by a clear affirmative action, signifies agreement to the processing of personal data” (Art. 4(11)) (European Data Protection Board 2020).*

Since Model Constraint 2 enforces that each «Collection» specifies a «LawfulGround» and a «ProcessingPurpose», the DPMF directly supports the requirement for the consent to be “specific” (ii). The DPMF also facilitates the assessment of the other attributes (i, iii, iv, and v) by providing the full picture of the «Processings», «ProcessingPurposes», «PersonalDataTypes», and the involved «Actors».

**Assessment 1.5** *Consent – Demonstrability* ②

*“Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his/her personal data” (Art. 7(1)) (European Data Protection Board 2020).*

Since Model Constraint 2 enforces that each «Processing» based on «Consent» as a «LawfulGround» is clearly specified, the DPMF enables the corresponding «Actor» to identify the «Processings» for which it is necessary to implement measures to demonstrate that the consent has been given by the data subject.

**Assessment 1.6** *Consent – Separation and intelligibility* ②

*“If the data subject’s consent is given in the context of a*

<sup>8</sup> As highlighted by the Article 29 Working Party (2014): “this is not a straightforward balancing test which would simply consist of weighing two easily quantifiable and easily comparable ‘weights’ against each other. Rather, carrying out the balancing test may require a complex assessment taking into account a number of factors”.

written declaration which also concerns other matters, the request for consent shall (i) be presented in a manner which is clearly distinguishable from the other matters, (ii) in an intelligible and easily accessible form, using clear and plain language” (Art. 7(2)) (European Data Protection Board 2020).

Since Model Constraint 2 enforces that each «Processing» activity based on «Consent» as a «LawfulGround» is clearly indicated, the DPMF allows the corresponding «Actor» to identify the «Processings» for which it is necessary to implement measures to demonstrate that the consent by the data subject for the «Processing» has been given separately from the other matters.

#### **Assessment 1.7 Consent – Withdrawal** ②

“The data subject shall have the right to withdraw his/her consent at any time. The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. Prior to giving consent, the data subject shall be informed thereof” (Art. 7(3)) (European Data Protection Board 2020).

Since Model Constraint 2 enforces that each «Processing» activity based on «Consent» as a «LawfulGround» is clearly indicated, the DPMF allows the corresponding «Actor» to identify the «Processings» for which it is necessary to implement measures to easily allow the data subject to withdraw his/her consent. As such, each «ProcessingPurpose» with «Consent» as «LawfulGround» must be revocable by the data subject, which must be communicated to the data subject. As these are clearly and unambiguously modeled in a DPM, they can be easily identified together with the «DataSubjectTypes» that must be systematically informed.

#### **Assessment 1.8 Consent – Age requirement** ③

If the data subject is a child, consent is only valid if it has been given or authorised by the holder of parental responsibility. While the age foreseen in the GDPR is 16, Member States may provide by law for a lower age provided that it is not below 13 (Art. 8(1)).

In case of a child, the «Processing» can only be based on «Consent» if it is either *authorised-Consent* or *parentalConsent*. If *minAge* of the «DataSubjectType» is > 16, the data subject is

not a child. If *minAge* is < 13, the data subject is a child. If  $13 < \text{minAge} < 16$ , the qualification of the «DataSubjectType» as a child follows from the combination of the *habitualResidence*, the *minAge*, and the corresponding national implementing act.

#### **Assessment 1.9 Consent – Verification of parental authorization or consent** ②

When consent is given by a child, “the controller shall make reasonable efforts to verify that consent is given or authorized by the holder of parental responsibility over the child, taking into consideration available technology” (Art. 8(2)).

If, following Assessment 1.8, a «Processing» is based on a «Consent», that consent must be either *authorisedConsent* or *parentalConsent*. The «Consent» specifies which of those and allows the corresponding «Actor» to identify the «Processings» activities for which it is necessary to implement measures to verify that the authorization or consent has been given by the holder of parental responsibility.

#### **Legal Assessment 2 Purpose limitation**

##### **Assessment 2.1 Purpose specification** ③

“Personal data shall be collected for specified, explicit and legitimate purposes” (Art. 5(1)b) (Article 29 Working Party 2013a).

Because of Model Constraint 2, each «Collection» of «PersonalDataTypes» specifies a «LawfulGround» and a «ProcessingPurpose». As a result, the DPMF fully supports this assessment by enforcing the specification of this information for every «Collection» in the DPM.

##### **Assessment 2.2 Compatibility assessment** ②

“Personal data shall not be further processed in a manner that is incompatible with the purposes for which they were initially collected” (Art. 5(1)b) (Article 29 Working Party 2013a).

The «ProcessingPurposes» of the «Further-Processings» must not be incompatible with the «ProcessingPurposes» of the «Collection». The meta-model enables such an assessment not only by modeling all the necessary concepts, but also by traversing the chain of previous «Further-Processings» to verify compatibility with the «LawfulGrounds» and «ProcessingPurposes» of the «Collection».

**Assessment 2.3** *Presumption of non-incompatibility for further processing for archiving purposes in the public interest, scientific or historical research, or statistical purposes* ③

“Further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall not be considered to be incompatible with the initial purposes” (Art. 5(1)b).

As an exception to Assessment 2.2, if the «ProcessingPurpose» of a «FurtherProcessing» is «ArchivingInThePublicInterestScientificOrHistoricalResearchOrStatisticalPurpose», then it must not, *per se*, be considered incompatible with the «ProcessingPurpose» of the «Collection», provided that the necessary safeguards are implemented. Information on the implemented organizational or technical measures is not expressed in the model.

**Legal Assessment 3** *Data minimization* ②

Personal data shall be limited to what is necessary for the purposes for which they are processed (Art. 5(1)c).

For each «Processing», the «PersonalDataTypes» must be strictly necessary to achieve the corresponding «ProcessingPurpose». The DPMF cannot automate this assessment, but it can assist the user in performing a manual assessment by presenting all the relevant information on each «Processing» and the «PersonalDataTypes» involved in that «Processing».

**Legal Assessment 4** *Storage limitation*

**Assessment 4.1** *Necessity* ②

“Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed” (Art. 5(1)e).

Soundness Constraint 9 ensures that every «Storage» specifies either a *retentionPeriod* or the *retentionCriteria* used to determine that period. For each «Storage» of a «DataSet», the *retentionPeriod* or the *retentionCriteria* must not be longer than what is necessary to achieve the «ProcessingPurposes» of the «Collection». The assessment of whether such storage is longer than necessary is a manual assessment that has to be performed by the user.

**Assessment 4.2** *Exemption for further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes* ③

“Personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes” (Art. 5(1)e).

As an exception to Assessment 4.1, if the «ProcessingPurpose» of a «Storage» is «ArchivingInThePublicInterestScientificOrHistoricalResearchOrStatisticalPurpose», then the *retentionPeriod* of that «DataSet» can be longer than necessary to achieve the «ProcessingPurposes» of the «Collection».

**Legal Assessment 5** *Processing of special categories of personal data – Exemption to the general prohibition – Necessity* ②

When the processing of special categories of personal data is authorised under one of the exemptions listed in Art. 9(2)b,c,f-j, the processing must be objectively necessary to achieve the essence of that specific exemption.

Model Constraint 4 requires each «SpecialCategory» of personal data to be paired with a «ProhibitionExemptionType». If the «ProhibitionExemptionType» is any of the aforementioned (Art. 9(2)b,c,f-j), the relevant details must be provided and the «ProcessingPurpose» of the «Processing» of those «SpecialCategory» of personal data must be objectively necessary. For example, if the «Processing» of «SpecialCategory» of personal data is based on *substantialPublicInterest* as a «ProhibitionExemptionType», the actual substantial public interest must be provided and the «Processing» must be objectively necessary to achieve that interest. This information documenting the necessity of processing the «SpecialCategory» is not captured directly in the model, but the model assists in identifying the instances for which it has to be recorded.

**Legal Assessment 6** *Processing of personal data relating to criminal convictions and offences* ②

“Processing of personal data relating to criminal convictions and offences or related security measures based on Article 6(1) shall be carried out only under the control of official authority or when the processing is authorised by Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects” (Art. 10).

If personal data related to «CriminalConvictionsAndOffences» are processed, that specific «Processing» can only happen under the control of official authority or if the applicable Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects is specified. The DPMF can assist by locating these «Processings» together with information on the involved «Actors» (and, for example, their country) to support performing this assessment. The DPM does not capture the result of this assessment.

#### **Legal Assessment 7** *Decision based solely on automated processing*

##### **Assessment 7.1** *Measures to be implemented in case of exemption to the general prohibition* ②

“In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject’s rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his/her point of view and to contest the decision” (Art. 22(3); Rec. 71).

If, as detailed in Soundness Constraint 10, the «AutomatedDecisionMaking» is based on either «Contract» or *explicit* «Consent» and passes – for the contract – the corresponding necessity test (Assessment 1.1), then the corresponding «Actor» must also implement suitable measures to allow the data subjects to, at least, obtain human intervention of the controller, express their point of view, and contest the automated decision.<sup>9</sup>

##### **Assessment 7.2** *Special categories of personal data* ③

“Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject’s rights and freedoms and legitimate interests are in place” (Art. 22(4)).

A «AutomatedDecisionMaking» can only have as an input a «DataSet» that contains «SpecialCategory» of personal data if the specified «ProhibitionExemptionType» is either *explicit-Consent* or *substantialPublicInterest*.

<sup>9</sup> The possibility for the data subject to obtain an ‘explanation’ of the decision reached after such an assessment is only mentioned in the non-binding Recital 71 GDPR.

#### **Legal Assessment 8** *Joint controllers* ②

*Joint controllership requires a transparent specification of the respective responsibilities for compliance with the obligations under the GDPR by means of an arrangement between the joint controllers (Art. 26(1)(2)).*

If a «Controller» role specifies more than one actor for a «Processing», those «Actors» are considered as joint controllers with regard to that «Processing». As such, they must determine their respective responsibilities by means of an arrangement between them, designate a single point of contact, and make the essence of that arrangement available to the data subject. The DPM supports identifying these instances of joint controllership. The respective responsibilities are not modeled in the DPM.

#### **Legal Assessment 9** *Representative of controllers or processors not established in the EU* ③

*A controller established outside the EU but falling within the territorial scope of the GDPR according to Article 3(2) must designate a representative in the EU (Art. 27(1)(2)).*

If an «Actor» with the role of «Controller» or «Processor» has *establishedInEU* = FALSE and *publicAuthority* = FALSE, then it must be represented by a «Representative» in the Union.

#### **Legal Assessment 10** *Prohibition to engage another processor without prior specific or general approval of the controller* ②

“The processor shall not engage another processor without prior specific or general written authorisation of the controller. In the case of general written authorisation, the processor shall inform the controller of any intended changes concerning the addition or replacement of other processors” (Art. 28(2)).

When, with regard to the same «Processing», a «Processor» engages another «Processor», it must have the prior specific or general approval of the «Controller» on whose behalf it acts.

#### **Legal Assessment 11** *Controller-processor agreement* ②

*The processing by a processor must be governed by a contract or other legal act that is binding on the processor with regard to the controller and that sets out the details of the processing (Art. 28(3)).*

When, for a specific «Processing», there is an «Actor» acting as a «Processor», there must be a contract with the «Controller» that contains all the elements listed in Art. 28(3).

#### **Legal Assessment 12** *Transfer to third countries or international organizations*

##### **Assessment 12.1** *Adequacy decision* ②

“A transfer of personal data to a third country or an international organisation may take place where the Commission has decided that the third country, a territory or one or more specified sectors within that third country, or the international organisation in question ensures an adequate level of protection” (Art. 45(1)).

In case of a «Disclosure» to an «Actor» not established in EU or an international organization, there must be an adequacy decision issued by the European Commission concerning the country of the recipient or the international organization.

##### **Assessment 12.2** *Appropriate safeguards* ②

In the absence of an adequacy decision, “a controller or processor may transfer personal data to a third country or an international organisation only if the controller or processor has provided appropriate safeguards, and on condition that enforceable data subject rights and effective legal remedies for data subjects are available” (Art. 46(1)).

In case of a «Disclosure» to an «Actor» not established in EU or an international organization and there is no decision as referred to in Assessment 12.1, then the «Actor» disclosing the personal data must provide appropriate safeguards as required by Art. 46(2).

### **4.3 Legal Assessments of Obligations Facilitated by the DPM**

The previous section outlined a number of legal assessments and concerns that can be supported directly through analysis – and in some cases, through extension – of DPM models. In this section, we discuss remaining and relevant legal assessments for which such support can not be found in a DPM model, but in the broader context of Data Protection by Design (DPbD) as implemented by the DPMF.

#### **Legal Assessment 13** *Record of processing activities* ①

“Each controller and, where applicable, the controller’s representative, shall maintain a record of processing activities under its responsibility” (Art. 30(1)).

The DPM does not capture whether a «Controller» keeps track of their «Processings» so it cannot automate this assessment. However, since the DPM does capture all the relevant information on the «Processings», the use of DPMs by a «Controller» does facilitate the compliance with this obligation to maintain a record of processing activities (see Sect. 8.3).

#### **Legal Assessment 14** *Transparency*

##### **Assessment 14.1** *Towards data subjects* ①

“Personal data shall be processed in a transparent manner” (Art. 5(1)a; 12–14).

While the DPMF does not assess whether the transparency obligations are met, the DPM does help in meeting these obligations because it captures all the relevant information on the the «Actors», «Processings», «LawfulGrounds», «ProcessingPurposes», and «PersonalDataTypes». This information enables the DPMF to automatically generate suitable documentation directly from the model (see Sect. 8.3).

##### **Assessment 14.2** *Timing – Personal data collected from the data subject* ①

“Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information” (Art. 13(1)).

In case of a direct «Collection» (i. e. when the «DataSets» originate directly from the data subject), the «Controller» must provide all the information listed in Art. 13 to the data subject at the time of the «Collection».

##### **Assessment 14.3** *Timing – Personal data not obtained from the data subject* ①

“The controller shall provide the information (a) within a reasonable period after obtaining the personal data, but at the latest within one month [. . .]; (b) if the personal data are to be used for communication with the data subject, at the latest at the time of the first communication to that data subject; or (c) if a disclosure to another recipient is envisaged, at the latest when the personal data are first disclosed” (Art. 14(3)).

In case of an indirect «Collection» (i. e. when the «DataSets» originate from another «Actor» after a «Disclosure»), the recipient acting as «Controller» must provide the information listed in Art. 14 to the data subject either within a reasonable period not exceeding one month, at the time of the first communication with the data subject, or at the time of the «Disclosure». The model can assist in determining the «Actor» responsible for communicating this to the data subject.

**Legal Assessment 15 Integrity and Confidentiality** ①

“Personal data shall be processed in a manner that ensures appropriate security” (Art. 5(1)f).

For every «Processing», the appropriate technical and organizational measures must be implemented to protect against, amongst others, unauthorized or unlawful processing and against accidental loss, destruction, or damage.

**Legal Assessment 16 Accuracy** ①

Personal data shall be accurate and kept up to date; reasonable steps must be taken to ensure that inaccurate data are erased or rectified without delay (Art. 5(1)d).

For each «Processing» the associated «DataSet» must be accurate with regard to the «ProcessingPurpose». The corresponding «Actor» must implement measures to ensure that inaccurate «DataSets» are erased or rectified.

**Legal Assessment 17 Security of processing** ①

Appropriate technical and organization measures must be implemented to secure the processing and protect the rights and freedoms of natural persons (Art. 32).

The DPMF provides guidance in raising the issue of providing appropriate security measures for the processing and can assist in identifying the «Processing» operations with the highest risk to the data subjects because of, for instance, the sensitivity of the information being processed (e. g., medical information). For the actual assessment of the technical measures, complementary security and privacy analysis (Deng et al. 2011; Shostack 2014) and risk assessment approaches (Lund et al. 2010; Sion et al. 2019b) can be applied to ensure appropriate measures are in place.

## 5 Methodology

The meta-model presented in Sect. 3 provides the means to document data processing activities in much detail. This section describes the proposed methodology to populate such a DPM in a structured manner. Rather than being a sequential process, it embodies an approach that iterates between (i) describing the processing operations and (ii) specifying the legal rationale. While iterating between these two activities, the level of detail in each of them increases.

Similar to the Twin Peaks model to software engineering (Nuseibeh 2001), where requirements and architectural specifications are concurrently developed with increasing detail, populating a DPM will alternate between describing the processing operations and specifying the legal rationale while gradually increasing the level of detail in each of them.

Fig. 6 visualizes these aspects in two parts. The top part illustrates the iterative refinement of the processing descriptions and legal rationale as two peaks which are expanded as more details are included. The bottom part shows a cross-section, illustrating the relevant DPM concepts (from Sect. 3) used in these peaks and their relationships with each other.

### 5.1 Threshold assessment

When processing personal data, a *threshold assessment* is needed to determine whether an in-depth DPIA is required in the context of the data processing activities. In order to facilitate such an assessment, the Article 29 Working Party (2017) has issued a list of criteria to consider when determining whether a processing operation is likely to result in high risk to data subjects’ rights and freedoms. Regardless of the result of that assessment (as emphasized in Sect. 2.1), conducting a DPIA is also the cornerstone of a risk-based approach.

### 5.2 Description of processing operations

As shown in Fig. 6, the description of processing activities revolves around three main concepts, namely: actors, processings, and data (which includes the data subjects).

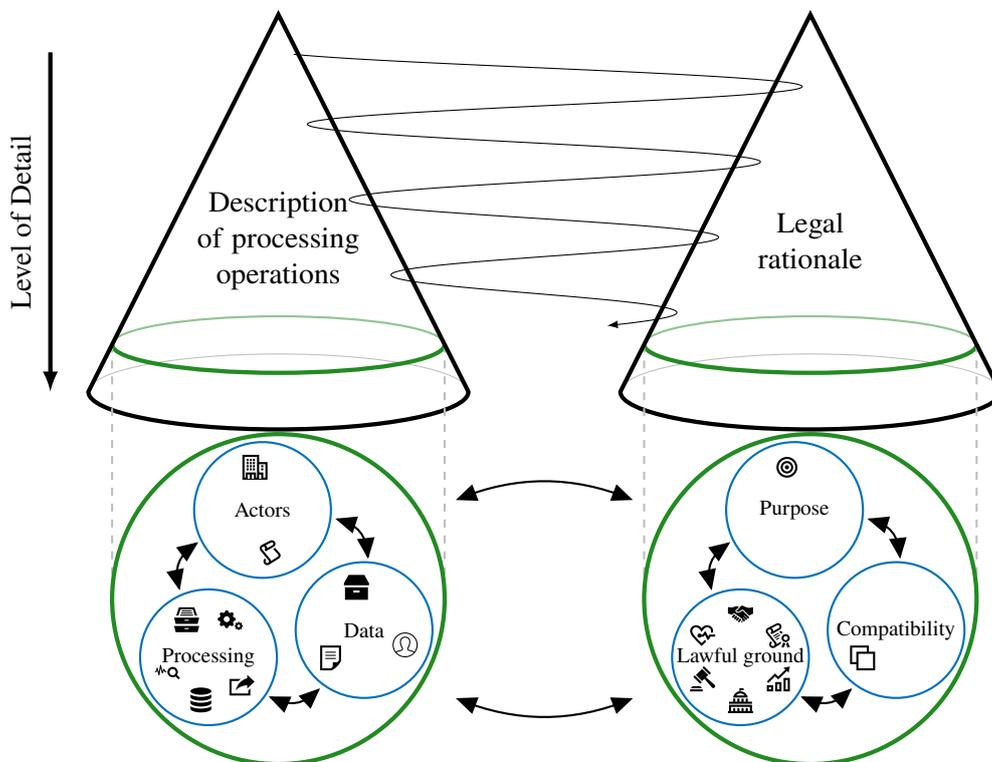


Figure 6: Overview of the iterative approach to build Data Protection Models.

**Actors** This involves specifying the entities («Actors») involved in the processing operations and their legal roles («Controller», «Processor», or recipient).

**Processings** This involves describing the specific types of processing activities. These concepts link to the involved entities through the «Legal-Roles» or as recipient from a «Disclosure» or providing data for an indirect «Collection».

**Data** This involves modeling the different «Data-Sets», which link to the processing activities, the «PersonalDataTypes» contained in these «DataSets», and the «DataSubjectTypes», the «PersonalDataTypes» belong to.

There is no strict sequential order in the description of the processing operations since associations between different elements can trigger the refinements of other elements. For example, splitting up the processing activities in more granular ones impacts the data sets, and processing additional data types influences both the processing activities (by,

for example, adding a new «Collection») and the corresponding «Actors» for these activities.

### 5.3 Description of legal rationale

The right-hand side of Fig. 6 shows how the legal rationale involves the specification of purposes, lawful grounds, and compatibility assessments.

**Purposes** This involves specifying the appropriate «ProcessingPurposes» of each of the «Processings» in the DPM.

**Lawful grounds** This involves specifying the «LawfulGrounds» of the «Collections». They also require a «ProcessingPurpose».

**Compatibility assessments** In the DPM, the «CompatibilityAssessments» document the results of the assessments by linking «ProcessingPurposes» to «LawfulGrounds» together with the rationale and the assessment result.

The specification of these elements of legal rationale can again influence the processing activities, involved data, and actors.

## 5.4 Refining the model

The meta-model (Sect. 3.3.1) and soundness (Sect. 3.3.2) constraints will indicate the inconsistencies and incomplete parts in the model that must be addressed (e. g. requiring the specification of a «LawfulGround» for the «Collections»), while the legal assessments listed and described in Sect. 4.2 provide guidance in ensuring GDPR compliance by raising issues such as the purpose compatibility of all «FurtherProcessings», or reducing the «PersonalDataTypes» for a specific «Processing» to minimize the personal data being processed. Extensions and refinements are necessary until all modeling constraints and legal assessments are successful. Evidently, changes to the system itself should be reflected in the model, and vice-versa.

## 5.5 Documentation extraction

Based on the DPM model, the main outcome of this methodology, suitable documentation can be extracted that describes the processing activities and justifies the choices made during the design phase. Sect. 8.3 provides a more elaborate discussion on the different types of documentation that can be generated from the DPM.

## 6 Implementation

We have implemented the DPMF in a research prototype that allows the creation DPMs and performing soundness and legal assessments.<sup>10</sup> Sect. 6.1 first discusses the implementation of the prototype. Subsequently, Sect. 6.2 presents the implementation of a selection of the legal assessments introduced in Sect. 4.2.

### 6.1 Prototype

To validate the presented framework for constructing DPMs, the meta-model is implemented in an Eclipse-based prototype. Fig. 7 provides a high-level overview of the technological components that constitute the DPMF.

<sup>10</sup> More information on the prototype is available at: <https://distrinet.cs.kuleuven.be/software/dpmf/>

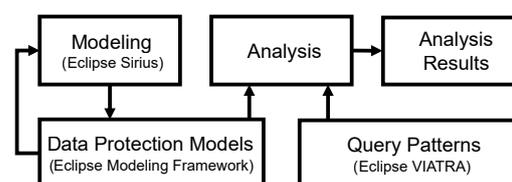


Figure 7: Overview of the DPMF components.

This diagram shows the high-level components of the prototype implementation of the DPMF.

The support for creating and manipulating DPMs is provided by implementing the meta-model from Fig. 1 in the Eclipse Modeling Framework (Steinberg et al. 2008). The implementation ensures the soundness of the DPMs by enforcing the constraints developed in Sect. 3.3.

To create concrete DPMs, graphical modeling support is implemented using an Eclipse Sirius Viewpoint Specification. This specification provides two diagram views for creating and editing DPMs. The first view is a class-based editor, which provides access to a detailed specification of the DPM elements. The second diagram view is a more user-friendly graphical visualization which shows the different elements and their roles, and defers the lower-level details to properties panes (see Fig. 8). Every DPM fragment used for illustration purposes in Sect. 3 and 7 has been created in and exported from the DPMF prototype.

To evaluate the legal requirements from Sect. 4 on concrete DPMs, the DPMF implements these criteria as patterns specified in VIATRA, a graph-based pattern language.<sup>11</sup> Subsequently, concrete DPMs can be queried for these patterns using the VIATRA query engine. The next section discusses the implementation of a number of legal assessments as VIATRA query patterns. The matched patterns are then used in the subsequent analysis in which the legal stakeholder is provided the results of the legal assessments. Subsequently, the appropriate modifications can be made to the DPM to resolve the identified issues. As a benefit of implementing the assessments as model query

<sup>11</sup> <https://www.eclipse.org/viatra>

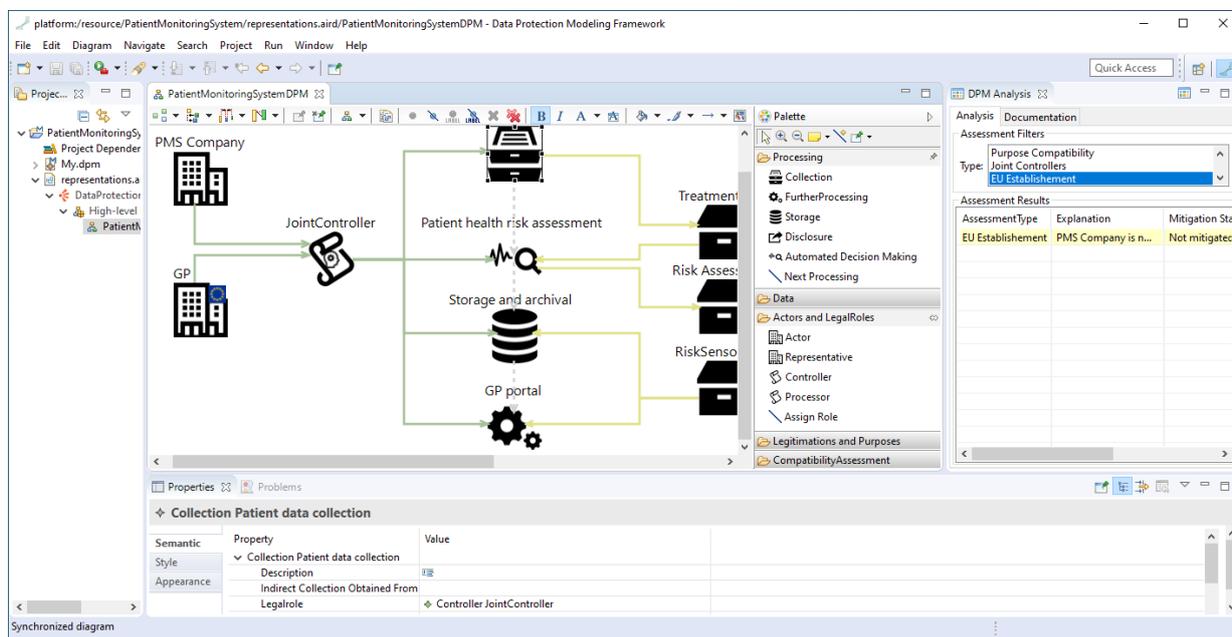


Figure 8: Data Protection Modeling Framework screenshot.

It shows the graphical editor in the center, element properties at the bottom, and filtered assessment results in the right pane.

patterns, the modified model can be efficiently re-assessed to verify that the changes resolve the earlier uncovered issues. The implementation of all the legal assessments is currently in progress.

### 6.2 Implementation of the assessments

This section revisits the legal assessments elicited in Sect. 4.2 and presents the concrete implementation of these assessments as VIATRA model query patterns. The implementation of a legal assessment consists of two parts:

**Identifying DPM locations**, which involves the identification of the relevant elements in a DPM that indicate an area of concern worth a more detailed assessment by a legal stakeholder.

**Guiding the assessment**, which requires a legal stakeholder to determine whether the potential issues call for a modification of the DPM or the implementation of organizational measures.

VIATRA uses the concepts from the meta-model presented in Sect. 3 as keywords, while properties are separated with a ‘.’ which gives, for example, *Actor.representedBy(actorInstance, representativeInstance)*. Combined with *Representative.name(representativeInstance, “TEST”)*, this

allows us to find all «Actors» that have a «Representative» with the specified name “TEST”.

Given that many of the smaller checks are quite trivial, this section focuses on revisiting some of the more complex assessments to illustrate the potential of DPM model queries in guiding legal stakeholders in conducting their compliance exercise. The discussion presents different assessments with increasing complexity from Sect. 4.2.

#### 6.2.1 Legal Assessment 9

*Representative of controllers or processors not established in the EU.* Performing this assessment requires the identification of all «Actors» that have a «LegalRole» as «Controller» or «Processor» and that are not established in the EU. The pattern below matches with «Actors» which: (i) have a «LegalRole» (can be either controller or processor), (ii) are *not* established in the EU, (iii) are *not* a public authority, and (iv) do *not* have a «Representative» in the EU.

```

pattern representativeMissing(a:Actor) {
    Actor(a);
    Actor.actsAs(a,r);
}
    
```

```

LegalRole(r);
Actor.establishedInEU(a, false);
Actor.publicAuthority(a, false);
neg Actor.representedBy(a, _);
}

```

After identifying the problematic «Actors», the necessary «Representatives» can be specified in the DPM, after which the above query will no longer identify these «Actors» as problematic.

### 6.2.2 Legal Assessment 8

*Joint controllers.* Another requirement stemming from the GDPR is that joint controllers need to determine their respective responsibilities for compliance with their obligations by means of an arrangement between them. This situation can be easily identified in a DPM by querying the model for all «Controllers» to check whether the number of linked «Actors» is greater than 1.

```

pattern jointControllers(c:Controller,
    ctr : java Integer) {
    Controller(c);
    ctr == count Actor.actsAs(_,c);
    true == eval(ctr > 1);
}

```

After identifying all the instances of joint «Controllers» in the model, the legal stakeholder will have to ensure that the appropriate division of responsibilities has been made for each of them. The explicit specification of such role divisions is currently not supported in DPMs.

### 6.2.3 Assessment 7.2

*Automated decision-making – Special categories of personal data.* The performance of automated decision-making on personal data has several legal implications, especially in the case of the «SpecialCategory» of personal data. In this case, the «ProhibitionExemptionType» needs to be specified as either *explicitConsent* or *substantialPublicInterest*. The pattern below checks whether there is automated decision-making on «SpecialCategories» without a valid exemption.

```

pattern AdmSpecialCategories(adm
    :AutomatedDecisionMaking) {

```

```

AutomatedDecisionMaking.input(adm, ds);
DataSet.datatype(ds, dt);
SpecialCategory(dt);
neg find AdmValidExemption(dt);
}

```

```

private pattern AdmValidExemption(dt
    :SpecialCategory){
    SpecialCategory.exemption(dt,
        ::EXPLICIT_CONSENT);
} or {
    SpecialCategory.exemption(dt,
        ::SUBSTANTIAL_PUBLIC_INTEREST);
}

```

Any match to this pattern is a direct violation of the restrictions imposed by the Art. 22(4) and will need to be fixed. There is no further manual assessment required by the legal stakeholder. However, a similar pattern could be used to find all cases of valid automated decision-making based on «SpecialCategory» for which additional measures might also need to be implemented (see Assessment 7.1 and Art. 22(3)).

### 6.2.4 Assessment 2.2

*Purpose limitation – Compatibility assessment.* Finally, another important assessment that has to be performed involves ensuring that the «ProcessingPurpose» of every «FurtherProcessing» is not incompatible with the «ProcessingPurpose» for which the data were originally collected. The pattern below queries a DPM for every pair of a «Collection» and «FurtherProcessing» following from that initial «Collection» to check whether there is a «CompatibilityAssessment» specifying that the «FurtherProcessing»'s specific «ProcessingPurpose» is not incompatible with the «ProcessingPurpose» specified for the initial «Collection».

```

pattern CompatibilityAssessment(c:
    Collection, fp:FurtherProcessing,
    pp:ProcessingPurpose) {
    find NextProcessingActivity+(c, fp);
    ProcessingPurpose.furtherprocessing(pp,
        fp);
    neg find Compatibility(pp, c);
}
private pattern Compatibility(pp:

```

```

    ProcessingPurpose, c:Collection) {
Collection.subjectTo(c, lg);
CompatibilityAssessment(ca);
CompatibilityAssessment
    .processingpurpose(ca, pp);
CompatibilityAssessment
    .lawfulground(ca, lg);
} or {
Collection.subjectTo(c, lg);
LawfulGround.purpose(lg, pp);
}

```

This pattern provides all instances in a DPM that provide no information on the non-incompatibility of the «ProcessingPurpose» of a «FurtherProcessing» with the «LawfulGround» and «ProcessingPurpose» of the original «Collection». To mitigate this, the legal stakeholder can either specify a «CompatibilityAssessment» indicating that these are not incompatible, or change the «ProcessingPurpose» or «FurtherProcessing» so that they are no longer incompatible. This is an example of a legal assessment that can guide the inclusion of new knowledge on the compatibility of the «ProcessingPurposes» of the «FurtherProcessings» with those of the «Collection» to keep track of the progress in the compatibility assessment and enable the reuse of «CompatibilityAssessments» in future DPMs.

## 7 Scenario-Based Validation

This section provides a scenario-driven validation of the DPMF by applying it on a concrete eHealth application for monitoring patients with cardiovascular diseases, the Patient Monitoring System (PMS). The scenario-driven validation discusses three development evolutions of the PMS and showcases the role of the DPMF as an effective implementation of Data Protection by Design (DPbD) in such a development context: (i) the first development of the PMS which is accompanied with the initial description of the involved data processing operations (Sect. 7.1), (ii) an extension of the PMS with support for automated emergency service notification for patients in critical condition (Sect. 7.2), and (iii) a second extension which

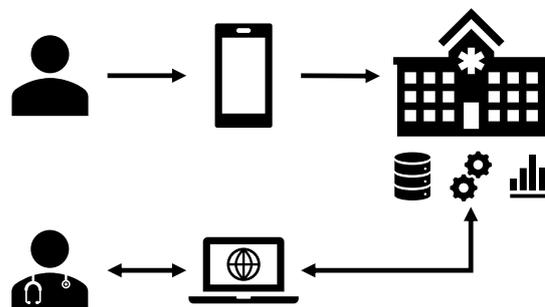


Figure 9: High-level overview of the system.

This diagram provides a very high-level overview of the Patient Monitoring System (PMS). At the top it shows the collection of sensor data from patients via a smartphone to the patient monitoring system. Those data are stored and analyzed, after which the results are communicated to the patient's general practitioner who can add treatment notes.

involves using the collected data for scientific research purposes (Sect. 7.3).

Before delving into the details of the validation, it is worth mentioning that this section does not intend to provide legal advice on the compliance of such a system, but rather aims at demonstrating the possibilities and the flexibility of the DPMF.<sup>12</sup>

### 7.1 Initial PMS version

Before any legal assessments can be performed using the DPMF, an initial description of the system needs to be provided. The Patient Monitoring System (PMS) is an eHealth system for the treatment and monitoring of patients with cardiovascular diseases. Fig. 9 presents a high-level overview of the system. The primary goal of the PMS is to support extra-mural, continuous and remote monitoring, timely decision-making, and prediction of malignant events. This is done by fitting patients with wearables to measure health parameters such as body temperature and electrocardiograms (ECGs).

<sup>12</sup> While all the assessments performed below draw on the relevant soft law instruments and consider the recent legal literature, the choices outlined in Sect. 7.1 to 7.3 are hypothetical and serve to illustrate the application of the DPMF. Examining all legal implications would distract from the validation and is therefore out of scope.

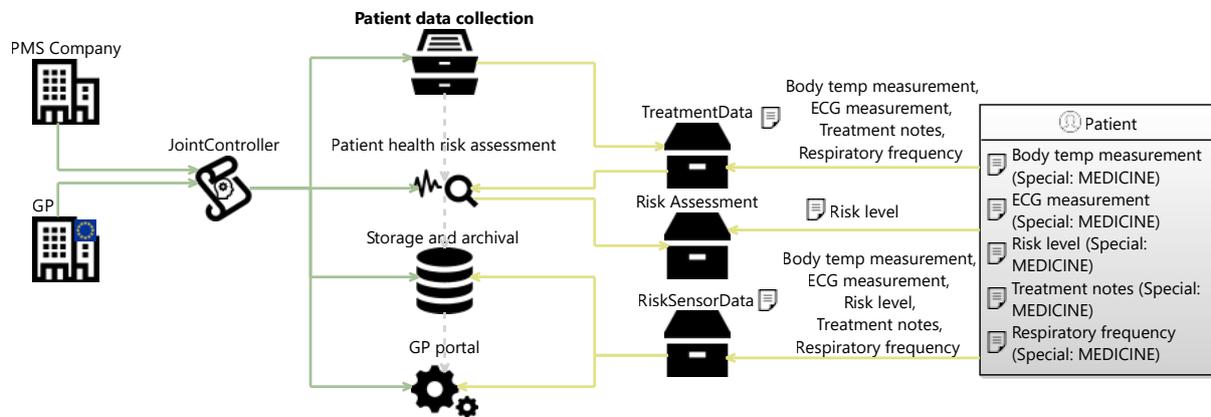


Figure 10: Data Protection Model for the initial PMS.

This view is constructed with the concepts from the meta-model (Fig. 1). It shows: (i) the «Actors» involved in the «Processing» activities as joint controllers (*left*), (ii) the sequence of «Processing» activities starting with a collection (*center*), and (iii) the processed «PersonalDataTypes» and «DataSubjectType» (*right*).

The data are sent from the smartphone app to the PMS platform, where the sensor data is used for risk assessment. Both the received sensor data and the outcome of the risk assessment are stored to make them available for general practitioners (GPs), who can use a web-based portal to access the information on their patients. The GPs can use the same system to record additional information on their patients such as treatment notes.

After technical stakeholders provide the description of the system (as outlined above), the legal stakeholder can create a DPM to describe the processing operations of the initial PMS (Fig. 10). It involves two «Actors», namely the *PMS Company* and the *GP*, in a situation of joint controllership with regard to all the «Processing» activities: an initial «Collection», followed by an «AutomatedDecisionMaking» for the cardiovascular risk assessment, a «Storage» and a portal accessible to the *GP*. Finally, the DPM also describes the involved «DataSets» and «PersonalDataTypes» of the *Patient* «DataSubjectType».

### 7.1.1 Model and soundness checks

With the creation of the DPM (depicted in Fig. 10), model and soundness constraints can be verified automatically. Tab. 2 presents the results of these

Table 2: Outcome of the model and soundness constraints for the initial PMS DPM (Fig. 10).

Model Constraint	Result
1: Start with collection	✓
2: Collection has lawful ground/purpose	✗
3: Processing has purpose	✗
4: Special category requires exemption	✓
5: Legal role has actor	✓
6: Disclosure has recipient	✓
Soundness Constraint	Result
1: Input/output consistency	✓
2: EU representative	✗
3: Collection has further processing	✓
4: Indirect collection consistency	✓
5: Every processing has controller	✓
6: Every processor has controller	✓
7: Lawfulness of public task	✓
8: Medical data—professional secrecy	✓
9: Storage has retention period/criteria	✗
10: Automated decisions—special categories	✗

checks. Each of the failed constraints is explained in more detail below.

**Model Constraint 2:** *Every collection must specify a lawful ground and a processing purpose.* For the *Patient data collection* activity, no «LawfulGround» or «ProcessingPurpose» is specified in the DPM.

**Model Constraint 3:** *Every further processing must specify a processing purpose.* For the three «FurtherProcessings», no «ProcessingPurposes» are specified.

**Model Constraint 4:** *Every processing of special categories of personal data must specify an exemption type.* While this constraint is met in the model, it is worth highlighting that lawfulness (Art. 5(1)a and 6(1)) and the exemption to the general prohibition of the processing of special categories of personal data (Art. 9(2)) are two distinct requirements. When processing special categories of personal data, it is therefore necessary to specify both.<sup>13</sup> In the current model, *Medicine* has been selected as the «ProhibitionExemptionType» for all «SpecialCategories».

**Soundness Constraint 2:** *A non-EU actor must appoint a representative.* There is no «Representative» specified for the *PMS company*.

**Soundness Constraint 8:** *A controller processing special categories of personal data for preventive or occupational medicine must be subject to professional secrecy.* Given the *Medicine* exemption, both «Controllers» will need to make sure that they are subject to professional secrecy as defined under the applicable law.

**Soundness Constraint 9:** *Every storage must specify a retention period or retention criteria.* The «Storage» activity named *Storage and archival* does not specify a retention period nor any retention criteria.

**Soundness Constraint 10:** *Every automated decision-making must specify contract, legal obligation or explicit consent as lawful ground.* This

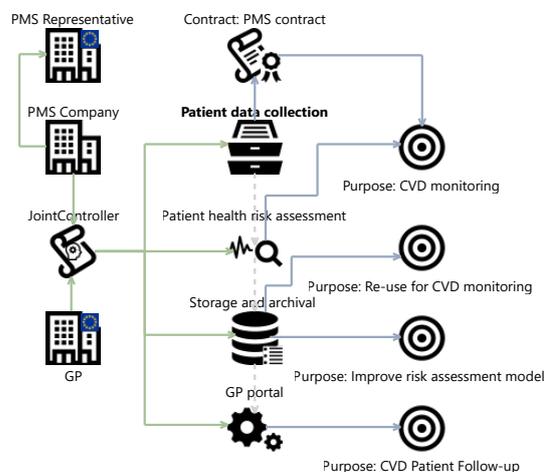


Figure 11: Updated DPM for the PMS to resolve the issues raised by the model and soundness constraints. This updates Fig. 10 to add a PMS «Representative» and assign «LawfulGrounds» and «ProcessingPurposes». It also adds the *retentionPeriod* for the «Storage» activity.

refines Model Constraint 2 because of the «AutomatedDecisionMaking». Since the latter is not met by the DPM, this one isn't either.

Fig. 11 shows the resulting DPM of the PMS after resolving the different issues identified by the model and soundness constraints. The required details in the properties of the elements have been filled in. The model is extended with the «LawfulGrounds», «ProcessingPurposes», and an EU representative for the *PMS company*.

### 7.1.2 Legal assessments

With the included DPM extensions depicted in Fig. 11, a number of legal assessments can be performed on this model. Assessments 1.1 and 7.2 aim at finding the right lawful grounds for the PMS, while Assessments 2.2, 3, 4.1, and 7.1 illustrate compliance with other GDPR principles.

**Assessment 1.1: Lawfulness – Performance of a contract, legal obligation, vital interests, public task, legitimate interests – Necessity.** The goal is to assess the necessity of the «Processings» with regard to the corresponding «LawfulGround». As Fig. 11 shows, «Contract» is stated as the «LawfulGround» for the «Collection» and

<sup>13</sup> The choice of «LawfulGround» does not dictate the choice of «ProhibitionExemptionType», and vice versa (Information Commissioner's Office 2019).

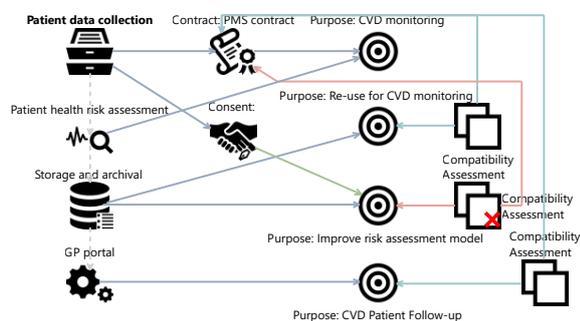


Figure 12: Compatibility assessments.

This DPM fragment shows the result of the compatibility assessments. The incompatibility for archival is resolved by relying on the data subject's consent for this storage activity.

«FurtherProcessing» of patient data. As explained by the European Data Protection Board (2019), such «Processings» must therefore be objectively necessary to perform the contractual service in question. In other words, if there are realistic, less intrusive alternatives, then those processing operations are not necessary. The DPM captures the contract details in the *contractualObligations* property and the «ProcessingPurposes» of the «Collection» and «FurtherProcessings». On that basis, the legal stakeholder can then perform the above-mentioned necessity test.

In this case, it can be argued that the collection and processing is objectively necessary to achieve the *contractualObligations* since those involve the continuous monitoring of the patient's cardiovascular disease and the real-time prediction of malignant events, which cannot be achieved through more traditional, less privacy-invasive ways (e. g., periodic appointments with a cardiologist).

**Assessment 7.2: Automated decision-making – Special categories of personal data.** This assessment only concerns «AutomatedDecisionMaking» activities that rely on «SpecialCategories» of personal data. Since it is the case for the *Patient health risk assessment*, such processing can only happen (i) if the data subject has given his/her explicit consent or (ii) when the processing is necessary for reasons of substantial public interest. This requirement is not met since *Medicine* has

been selected as the «ProhibitionExemptionType» following Model Constraint 4. Shifting to *explicitConsent* as the «ProhibitionExemptionType» resolves this issue and also avoids the need for the controllers to comply with Soundness Constraint 8.<sup>14</sup>

**Assessment 2.2: Purpose limitation – Compatibility assessment.** This assessment is one of the key principles underpinning the GDPR, according to which the «ProcessingPurposes» of every «FurtherProcessings» must not be incompatible with the «ProcessingPurposes» of the original «Collection». The DPMF allows to systematically iterate over the «FurtherProcessings» and associated «ProcessingPurposes» to assist legal experts when performing the compatibility test (outcome documented in Fig. 12).

When a «FurtherProcessing» refers to the exact same «ProcessingPurpose» of the «Collection», the compatibility assessment automatically succeeds. Otherwise, the DPMF will prompt the user for manual assessment and allow them to document the results in a «CompatibilityAssessment». For the PMS, this leads to the following assessments.

**Compatibility of the patient health risk assessment.** This assessment automatically succeeds because it refers to the exact same «ProcessingPurpose» as the «Collection» (i.e. *CVD monitoring*).

**Compatibility of the storage and archival.** While the *Re-use for CVD monitoring* is certainly compatible with the «ProcessingPurpose» of the «Collection» (i.e. *CVD monitoring*), the compatibility of *Improve risk assessment model*

<sup>14</sup> While it seems that this requirement is not met since the entire processing chain remains based on «Contract» and not «Consent» as the «LawfulGround», it is worth re-emphasizing that the specification of a «LawfulGround» and a «ProhibitionExemptionType» are two cumulative requirements that do not (necessarily) impact each other. Hence, it perfectly possible to consider *explicitConsent* as the «ProhibitionExemptionType» even if «Consent» is not used as the «LawfulGround» (Information Commissioner's Office 2019). As a result, there is no need to change the «LawfulGround».

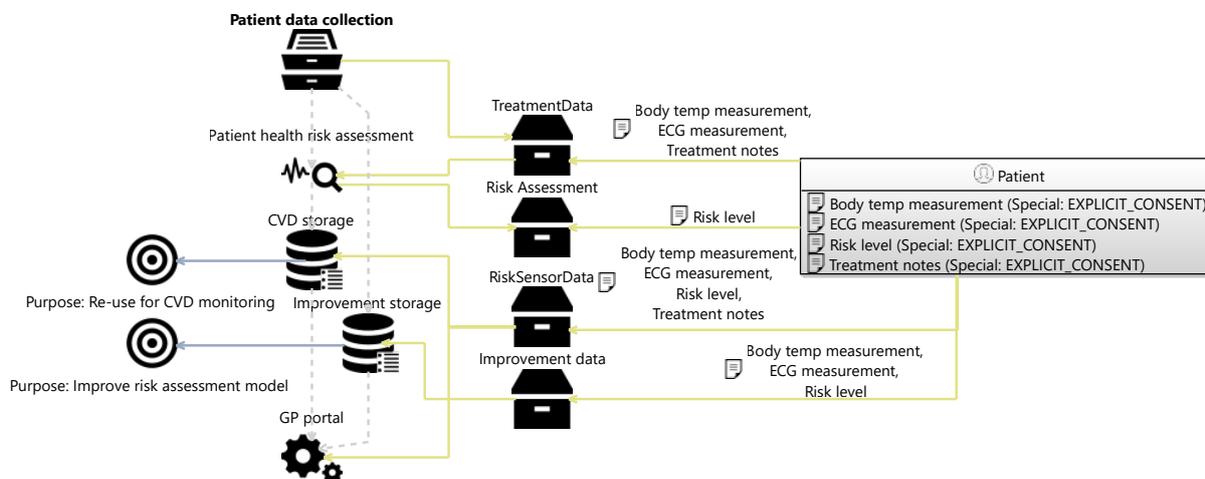


Figure 13: Removed respiratory frequency and split up storage and archival.

This DPM fragment shows the removal of the respiratory frequency and the splitting of the storage to exclude the risk levels from the long-term medical record and its impact on the linked purposes.

is debatable. In order to resolve this, it is possible to either: (i) consider it as a new chain of «Processing» activities based on the data subject's «Consent»<sup>15</sup> or (ii) refrain from storing the data for longer than strictly necessary to perform CVD monitoring. Fig. 12 shows a fragment of the resulting model in which the former solution has been implemented.

**Compatibility of the GP portal.** The *CVD patient follow-up* is considered to be compatible with the *CVD monitoring* «ProcessingPurpose» specified for the «Collection».

**Legal Assessment 3: Data minimization.** The goal of this assessment is to ensure that all the processed «PersonalDataTypes» are strictly necessary for the stated «ProcessingPurposes». The

<sup>15</sup> As emphasized by the Article 29 Working Party (2013a), the consequence of an incompatibility is a prohibition to further process those data which cannot be fixed by considering it as a new processing with a different lawful ground. At first, it seems that the solution in Fig. 12 does exactly that. However, this is only the case if the fix is implemented *a posteriori*, once the data have already been collected. Rather, the DPMF aims at preventing this situation by identifying the issue at design stage, to ensure compliance with the lawfulness and purpose limitation principles from the start. This will result in the controller asking the data subject's consent for improving the risk assessment model *before* the processing takes place, rather than extending the contract afterwards.

DPMF can assist by systematically considering every «Processing» with the relevant «ProcessingPurpose» and «PersonalDataTypes».

As illustrated by Fig. 10, the PMS collects the following patient data: (i) the body temperature, (ii) the ECG measurement, (iii) the respiratory frequency, and (iv) the treatment notes provided by the GP. Following a discussion between the GP and the PMS company, it appears that keeping track of the respiratory frequency is not strictly necessary to accomplish CVD monitoring. Furthermore, the *Storage and archival* includes all the «PersonalDataTypes». While this is certainly justified for the *CVD monitoring*, storing the treatment notes is not strictly necessary for improving the risk assessment model. Fig. 13 shows the updated DPM fragment in which the processing of the respiratory frequency was removed, and in which the initial «Storage» has been split into two separate ones with their own «ProcessingPurpose» and the relevant «PersonalDataTypes».

**Assessment 4.1: Storage limitation – Necessity** This assessment triggers the necessity test for the *CVD storage* and the *Improvement storage* «Storage» activities. The *retentionPeriod* of both «Storages» is set to 2 years in the DPM. While this seems reasonable to allow the PMS Company

and the GP to fulfil the second «ProcessingPurpose» (i. e. *Improve the risk assessment model*), such a long retention period is not strictly necessary for the *Re-use for CVD monitoring* «ProcessingPurpose». To resolve this, it is necessary to limit the *retentionPeriod* of the *CVD storage* by replacing the 2 years *retentionPeriod* with *retentionCriteria* that could consider, for instance, the following: (i) the evolution of the cardiovascular disease, (ii) the number of critical incidents that have happened during the monitoring period, and (iii) the decision of the GP to end the treatment.

**Assessment 7.1:** *Automated decision-making – Measures to be implemented in case of exemption to the general prohibition.* This assessment checks the existence of «AutomatedDecisionMakings» based on «Contract» or *explicit* «Consent» to highlight the need to implement suitable measures as detailed in Assessment 7.1. Since the *Patient health risk assessment* is such a form of automated decision making, the DPMF will raise the need for suitable measures to provide patients with meaningful information about the logic behind the risk calculation and allow them to influence its outcome by, for example, obtaining human intervention by the GP or providing more input on specific aspects of their lifestyle.

## 7.2 PMS extension: automated notification of emergency services

In the second scenario, the PMS Company decides to extend the system with automated functionality for notifying the emergency services to intervene in case of a critical incident. Fig. 14 shows the updated DPM in which the «Collection» is extended to include recording the *Real-time location* of the patients. Two «FurtherProcessings» have also been added to support the new functionality, namely: (i) an «AutomatedDecisionMaking» which uses the result from the *Patient health risk assessment* as input and (ii) a «Disclosure» which involves sharing the *InterventionData* to the new *Medical emergency services* «Actor», a new recipient that represents the emergency services that will be notified.

Table 3: Model and soundness constraints for the emergency notification extension DPM (Fig. 14).

Model Constraint	Result
1: Start with collection	✓
2: Collection has lawful ground/purpose	✓
3: Processing has purpose	✗
4: Special category requires exemption	✓
5: Legal role has actor	✓
6: Disclosure has recipient	✓
Soundness Constraint	Result
1: Input/output consistency	✓
2: EU representative	✓
3: Collection has further processing	✓
4: Indirect collection consistency	✓
5: Every processing has controller	✗
6: Every processor has controller	✓
7: Lawfulness of public task	✓
8: Medical data—professional secrecy	✓
9: Storage has retention period/criteria	✓
10: Automated decisions—special categories	✓

### 7.2.1 Model and soundness checks

The model and soundness constraints can be re-evaluated on the extended DPM to verify its consistency and soundness (Tab. 3). The failed constraints are discussed in more detail below.

**Model Constraint 3:** *Every further processing must specify a processing purpose.* This constraint is not met since the «FurtherProcessings» *Emergency notification decision* and *Disclosure to medical emergency services* do not specify a «ProcessingPurpose». Because these processings happen to automate the notification and to inform the emergency services, the legal stakeholder specifies these as the «Processing-Purposes» for the new «FurtherProcessings». Fig. 15 shows the result of this.

**Soundness Constraint 5:** *Every processing must specify a controller.* This constraint is not met since the *CVD storage* and *Improvement storage* (replacing the *Storage and archival* in the original DPM), and the *Emergency notification decision* do not specify a controller. Because the same two «Actors» (*PMS Company* and the *GP*) determine the means and purposes of the processing they are

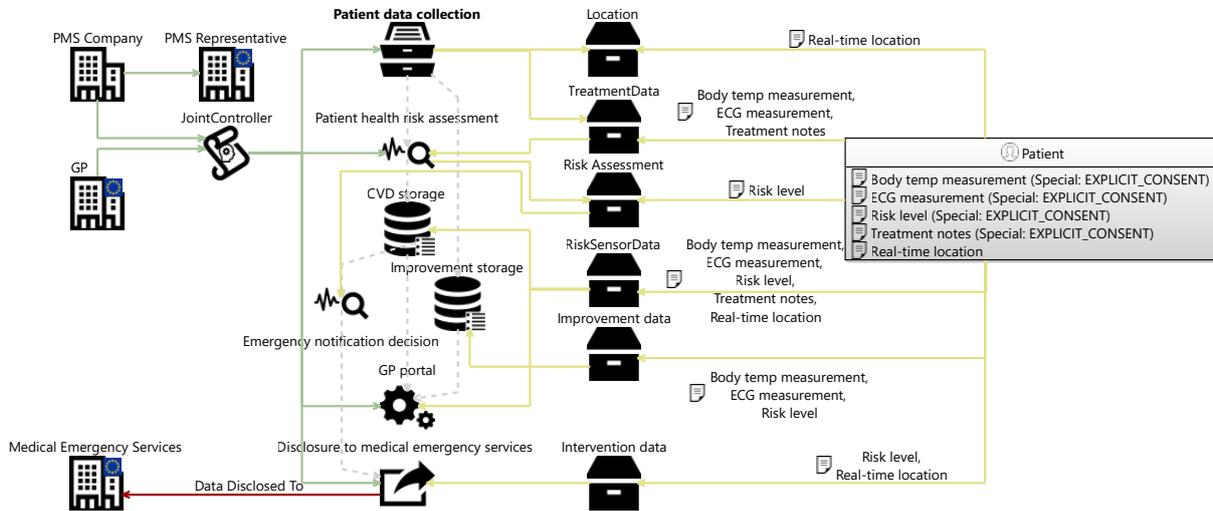


Figure 14: Updated DPM of the PMS extended with the new automatic notification functionality.

This diagram extends the PMS DPM resulting from Sect. 7.1 and Fig. 13 with new functionality for automatically notifying the medical emergency services in case of a critical incident.

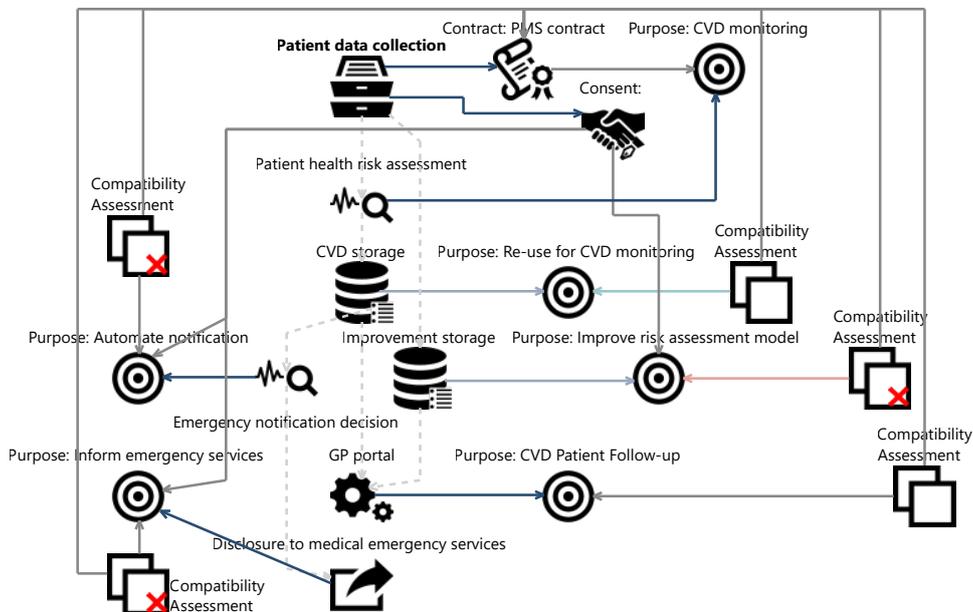


Figure 15: Updated DPM with the purposes and compatibility assessments for the notification functionality.

This diagram extends the PMS DPM from Fig. 14 with the necessary purposes and outcomes of the compatibility assessments resulting from Model Constraint 3 and Assessment 2.2.

again specified as joint controllers for these newly introduced «FurtherProcessings».

### 7.2.2 Legal Assessments

After resolving the issues, a number of legal assessments (2.2, 3, and 4.1) can be (re-)conducted on the extended DPM. For the sake of conciseness, only the most relevant ones are detailed below.

**Assessment 2.2: Purpose limitation – Compatibility assessment.** The compatibility of the two newly introduced «FurtherProcessings» with the original «Collection» needs to be assessed. In light of the key factors outlined by the Article 29 Working Party (2013a), it is unlikely that the «ProcessingPurpose» of the «AutomatedDecisionMaking» *Emergency notification decision* is compatible with the *CVD monitoring* specified for the «Collection», as the «Contract» between the PMS Company, the GP, and the patients only covers *CVD monitoring* and not the intervention of emergency services. This also holds true for the *Disclosure to medical emergency services*. Fig. 15 shows the updated DPM with *explicit* «Consent» as the «LawfulGround».

**Legal Assessment 3: Data minimization.** As illustrated in Fig. 14, the new functionality relies on the collection and disclosure of the patient's *Real-time location*. However, continuously collecting the patient's location is not strictly necessary for automatic notification of emergency services. A viable design alternative that is less privacy-invasive is to only collect and disclose the location of the patient when an incident occurs. In consequence, the DPM is updated to only include the *Incident location* instead of the *Real-time location*.

**Assessment 4.1: Storage limitation – Necessity.** Since the *Incident location* (previous *Real-time location*) is also stored in the *CVD storage*, the *retentionCriteria* are too broad, in the sense that the retention period will exceed the retention period that is strictly necessary to allow the emergency services to intervene in case of a critical incident. To resolve this, it is possible to modify the DPM to either: (i) exclude the *Incident location* from the *RiskSensorData* «DataSet» so as to not store

Table 4: Model and soundness constraints for the research extension DPM (Fig. 16).

Model Constraint	Result
1: Start with collection	✓
2: Collection has lawful ground/purpose	✓
3: Processing has purpose	✗
4: Special category requires exemption	✗
5: Legal role has actor	✓
6: Disclosure has recipient	✓
Soundness Constraint	Result
1: Input/output consistency	✓
2: EU representative	✓
3: Collection has further processing	✓
4: Indirect collection consistency	✓
5: Every processing has controller	✗
6: Every processor has controller	✓
7: Lawfulness of public task	✓
8: Medical data—professional secrecy	✓
9: Storage has retention period/criteria	✓
10: Automated decisions—special categories	✓

that information at any point, or (ii) create a new «Storage» for which the *retentionCriteria* will not exceed the time necessary for the emergency services to arrive at the incident location.

### 7.3 Extended version of the PMS to include research activities

The third and final scenario involves the PMS company deciding to use the collected data to study the evolution of cardiovascular diseases over time. Fig. 16 shows the updated DPM in which the PMS Company further processes some of the «PersonalDataTypes» to that aim. As before, Tab. 4 documents the results of running the model and soundness constraints, which led to: (i) the addition of the «ProcessingPurpose» *Scientific research* for the «FurtherProcessing» *PMS research*, (ii) the choice of *publicInterestScientificHistoricalResearch* as the new «ProhibitionExemptionType» for the processing of «SpecialCategory» and (iii) the specification of the PMS Company as the sole «Controller» for the new «FurtherProcessing». Some of the most relevant assessments are detailed below.

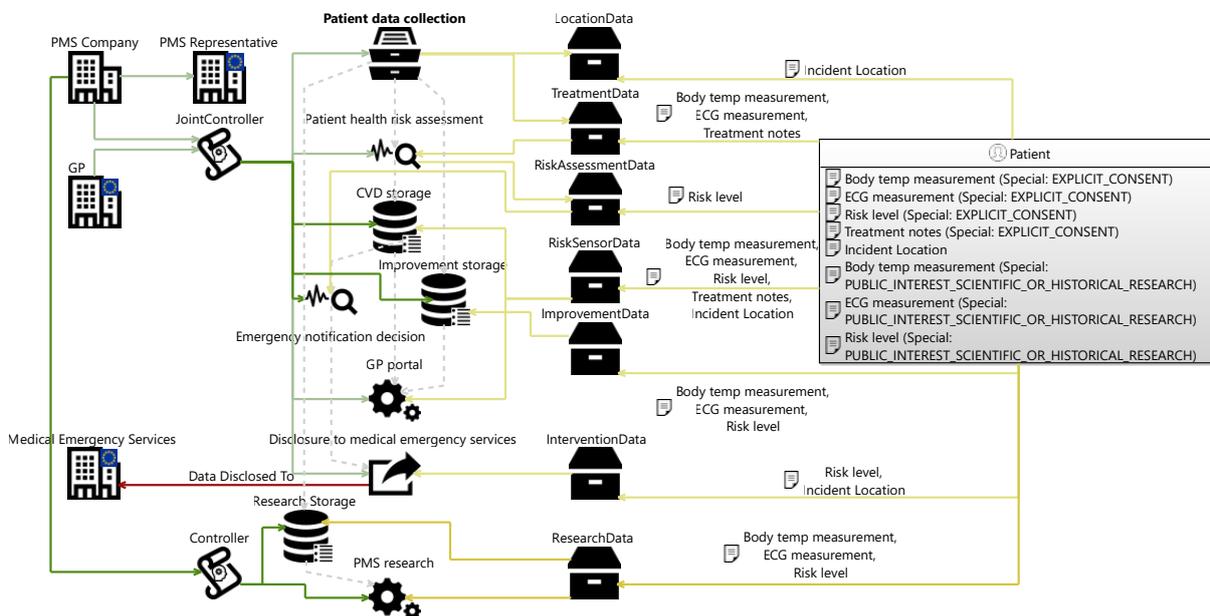


Figure 16: Updated DPM of the PMS extended to include the further processing for scientific research purposes.

This diagram extends the PMS DPM from Fig. 14 to include research activities undertaken by the PMS Company on the basis of the patient body temperature, ECG measurements and risk levels.

**Assessment 2.3: Purpose limitation – Presumption of non-incompatibility for further processing for archiving purposes in the public interest, scientific or historical research, or statistical purposes.** Adding a new «FurtherProcessing» requires examining its compatibility with the «ProcessingPurposes» specified for the «Collection». In the present case, since the «ProcessingPurpose» of the «FurtherProcessing» is *Scientific research*, the GDPR states that it shall not be considered as incompatible with the initial purposes (Art. 5(1)b). However, as clarified by the European Data Protection Supervisor (2020), this presumption of non-incompatibility does not amount to a blanket authorization to further process personal data for research purposes. In that sense, the criteria listed in the GDPR (Art. 6(4)) and outlined by the Article 29 Working Party (2013a) must still be taken into account.

The DPMF supports this assessment by raising the presumption of compatibility. On that basis, legal stakeholders can then perform and document the assessment. Here, as long as the activities

undertaken by the PMS Company fall under the broadly defined scope of “research” (Rec. 159) and the necessary safeguards (Art. 89(1)) are implemented, the outcome of the compatibility assessment is likely to be positive.

**Legal Assessment 3: Data minimization.** The «FurtherProcessing» PMS research relies on the *ResearchData* «DataSet», which includes the patient’s: (i) temperature, (ii) ECG and (iii) risk level computed by the algorithm. The DPMF allows to systematically present each pair of «Processing» and «ProcessingPurpose» with the associated «PersonalDataTypes», which in turn streamlines the necessity test developed in Legal Assessment 3. Here, following a thorough discussion between legal and medical stakeholders, it appears that the above-mentioned «PersonalDataTypes» are strictly necessary to efficiently perform the type of research envisioned by the PMS Company; there is therefore no need to further limit the scope of the *ResearchData* «DataSet».

To also ensure compliance with data minimization, the controller has to properly anonymize the

«DataSet» as long as the «ProcessingPurposes» of the research can be fulfilled in that manner (Art. 89(1)). Pseudonymization can be used as a secondary solution in case proper anonymization would reduce the overall utility of the «DataSet» for research. In that sense, data minimization is not only about adjusting the quantity of data to the «ProcessingPurposes», but also its form.

**Assessment 4.2:** *Storage limitation – Exemption for further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.* The *retentionPeriod* of the *Improvement storage* was set to two years, and appropriate *retentionCriteria* were elicited for the *CVD storage*. However, this might be too short of a period to efficiently achieve the *PMS research* «ProcessingPurpose». To resolve this, an additional «Storage» has been created for the *ResearchData* «DataSet» with a *retentionPeriod* of 5 years. Such a long period of time can be justified as long as the sole «ProcessingPurpose» is scientific research.

## 8 Discussion

This section elaborates on how and to what extent the DPMF: (i) substantiates the DPbD paradigm, (ii) lays the groundwork for a comprehensive, in-depth DPIA, (iii) supports the generation of documentation, (iv) sustains changes in the model, and (v) offers extension possibilities with elements from past models and dedicated knowledge bases.

### 8.1 Enabling data protection by design

Sect. 2.1 elicited the five components of DPbD. This section highlights how and to what extent the DPMF supports each of them.

**A risk-based approach.** DPbD requires controllers to take various criteria into account when complying with data protection rules (e. g., nature, scope, context, and purposes of the processing, risks for data subject's rights and freedoms). This is referred to as the risk-based approach. The DPMF explicitly supports those criteria through dedicated modeling concepts in the meta-model<sup>16</sup>

– the choice of which is explained is Sect. 3.2. This, in turn, allows legal and technical stakeholders to orient and refine their decisions based on a comprehensive, dynamic overview of the system. While the current version of the DPMF distinguishes between serious and minor issues, proper risks quantification will be explored in future work. Such risk analyses will also need to consider the effect of technical countermeasures on the privacy risk (Freund and Jones 2014; Lund et al. 2010; Sion et al. 2019b) by analyzing complementary software engineering views (Sion et al. 2019a).

**The obligation to ensure compliance with the requirements stemming from the GDPR.** As documented above, these requirements have been translated into: (i) directly enforceable model and soundness constraints (Sect. 3.3.1 and 3.3.2), and (ii) legal assessments (Sect. 4.2), which can be systematically performed on the created models. By operationalizing the requirements within a modeling framework, the DPMF can provide automated assistance to ensure compliance with the GDPR requirements, taking the form of either: (i) the identification of issues when modeling one or more specific elements (e. g., need to rely on *explicitConsent* or *substantialPublicInterest* as the «ProhibitionExemptionType» when modeling «AutomatedDecisionMaking» based on «SpecialCategory» of personal data) or (ii) the possibility to easily retrieve information from the model in order to streamline a given legal assessment (e. g., systematically considering the «ProcessingPurposes» and «PersonalDataTypes» associated with the «Storage» when performing the storage limitation test). Future work will include the development of more fine-grained guidance to

<sup>16</sup> While this is apparent for the first set of criteria, how the DPMF takes the data subject's rights and freedoms into account is not so straightforward. By interpreting, as suggested by Gellert (2018), the notion of "risk" under the GDPR as comprised of both an "event" (i. e. the lack of compliance) and a "consequence" (i. e. the actual risks for data subject's rights and freedoms), one can argue that the DPMF offers in-depth support for those risks by operationalizing the GDPR principles through specific constraints and assessments.

Table 5: Evaluation of our approach w.r.t. legal and security/privacy architecture DPbD requirements.

Approach	Description of processing activities	DPIA methodology support	Coverage of Art. 29 WP concepts	Support for soundness criteria (R1)	Support for legal criteria (R2)	Tool support	Document generation (R3)	Model management (R4)
GDPR modeling approaches								
Data Protection Modeling Framework Modeling	●●● 95%	●●●	●●●	●●●	●●●	●●○	●●○	●●○

Legend: ○○○: unsupported, ●○○: limited, ad-hoc or partial support, ●●○: supported but extensive manual effort required, ●●●: automated support (tools, generators, macros) for criteria.

orient the modeler when dealing with those issues and assessments (see Sect. 8.5).

**The implementation of technical and organizational measures.** The information in the DPMs and the legal assessments performed on them can be leveraged to raise issues on the need for appropriate measures. For example: (i) a joint controllership triggers the discussion on the allocation of responsibilities, (ii) relying on consent, requires a valid consent form, or (iii) the need for additional technical measures when processing «SpecialCategories» of data. Furthermore, the model also supports these activities by providing the relevant information for them, such as, in the case of the consent form, by providing information on the «Actors», «Processings», «ProcessingPurposes», and «PersonalDataTypes». Analogously to the risk analysis above, these analyses can benefit from the alignment and integration with complementary software engineering views (Sion et al. 2019a) to ensure that appropriate technical countermeasures are put in place. These integrated analyses will be addressed in future work.

**The need to demonstrate that the processing is performed in accordance with the Regulation.** By explicitly modeling DPMs and updating these

models as different issues raised by the legal assessments are resolved, the relevant documentation to demonstrate compliance is already automatically constructed as part of this process. Sect. 8.3 provides further detail on how the DPMF supports generating various types of documentation from a DPM adjusted to different types of audiences.

**The necessity to take all the above those considerations into account at the design stage, and throughout the entire data processing life cycle.** Since DPIAs are meant to be living instruments, DPbD should be a continuous exercise to addressing data protection issues throughout the different iterations of a data processing operations. As extensively demonstrated in Sect. 7, the Twin Peaks-inspired methodology outlined in Sect. 5 provides support during both the design and the processing phases. In that sense, the description of the processing activities and the identification and mitigation of the legal issues are concomitant and iterative processes. As a result, refining the former will inevitably lead to refining the latter, and vice-versa. Ultimately, every change made to the system and reflected in the DPM will trigger additional constraints and assessments that will need to be addressed by legal or technical stakeholders. The DPMF therefore assists in the

continuous consideration of data protection issues throughout the data processing life cycle.

## 8.2 Laying the groundwork for DPIAs

Sect. 2 provided an overview of the most notable approaches for conducting a DPIA and evaluated them using a set of key criteria (Tab. 1). This section highlights to what extent the proposed DPMF fulfils these criteria (shown in Tab. 5).

As mentioned in Sect. 2, DPIAs involve the following key elements: (i) the description of the processing, (ii) the identification of the risks to data subject's rights and freedoms, (iii) the implementation of appropriate countermeasures and (iv) the generation of accountability documentation. Since the meta-model presented in Fig. 1 was built using concepts stemming from the GDPR and the relevant soft-law instruments, the DPMF provides unmatched coverage of the WP29 concepts. The DPMF also offers automated support for model and soundness (Sect. 3.3.1 and 3.3.2) constraints as well as an extensive range of legal assessments (Sect. 4.2), and prototype tool support (Sect. 6). In addition, Sect. 8.3 discusses the extent to which preliminary documentation can be generated from the DPMs, while Sect. 8.4 elaborates on model management.

## 8.3 Generating documentation

The GDPR requires controllers to document their processing activities (Art. 5(2), 24(1), 25(1), 30). Since a DPM includes most of the information required to perform a DPIA, it already contains all the necessary details to comply with the above provisions. Furthermore, the DPMF offers a number of key benefits, as described below.

### **Generating records of processing activities.**

The GDPR requires controllers and processors to maintain a record of their processing activities (Art. 30(1,2)), and to be able to provide it upon request to supervisory authorities (Art. 30(4)). The DPMF can leverage the information in the DPM to automatically generate these records. Snippet 1 illustrates this with the record of the last version of the PMS detailed in Sect. 7.3. Such an export

provides a solid basis for the involved stakeholders to further extend as desired.

**Generating DPIA reports.** The DPM serves as the basis for a number of legal assessments, each involving structured queries over the model to answer legal questions to verify compliance with the GDPR. Therefore, this model provides – together with the answers to those questions – an excellent source for documenting the DPIA.

### **Generating documentation for different types of audiences.**

Because of the central representation of the processing operations in the DPM, a wide range of outputs for different audiences can be generated from the same model. Besides the record of processing activities, it can also help generate suitable documentation for other purposes, such as: (i) an overview of all the information that must be communicated to the data subjects (Art. 13, 14), which also lays the groundwork for the drafting of a comprehensive privacy policy (Article 29 Working Party 2018b); (ii) a list of all the information that must be provided to data subjects when exercising their right of access (Art. 15(1)); (iii) for processors, fragments of their own DPMs in order to provide the controller with the necessary information to demonstrate compliance (Art. 28(3)h); or (iv) any information requested by NSAs (Art. 31). These examples illustrate how a DPM can serve as the source for a wide range of documentation outputs.

## 8.4 Model and process management

Closely aligned with the documentation is keeping track of changes in the model. The changes to a DPM are indicative of either: (i) the evolution of the data processing operations over time in response to business changes, (ii) the changes applied in response to previously unanticipated risks, and (iii) the evolution of the legal assessments with regard to regulatory developments.

Therefore, it is important to not only document the current state, but also the evolution over time. Keeping track of those changes is a major challenge for which the model-driven engineering community has studied different approaches

---

Records of Processing Activities - PMS Company Inc.

Controller:

PMS Company Inc. (Address: 1 PMS Company Way, CA, US) (not established in EU)  
Representative: PMS Company Europe Ltd. (Address: Street 1, 1000 Brussels, Belgium)

Joint controller(s):

General Practitioner (Address: (local general practitioner), EU country) (established in EU)

Purposes of the processing:

- \*Improve risk assessment model
- \*CVD monitoring
- \*Re-use for CVD monitoring
- \*Inform emergency services
- \*Automate notification
- \*Scientific research
- \*CVD Patient Follow-up

Categories of data subjects and personal data:

- \* DataSubject: Patient (Residence: EU) - MinAge: 18
  - \* Body temp measurement
  - \* Treatment notes
  - \* Risk level
  - \* Incident Location
  - \* ECG measurement

Recipients:

- \* Medical Emergency Services (Address: null, null) (established in EU)

Time limits:

- \* Improvement storage: 5 years  
Data Types: Body temp measurement, ECG measurement, Risk level
- \* CVD storage: (i) the evolution of the cardiovascular disease,  
(ii) the number of critical incidents that have happened during the monitoring period, and  
(iii) the decision of the GP to put an end to the treatment  
Data Types: Risk level, Body temp measurement, ECG measurement, Treatment notes, Incident Location
- \* Research Storage: 5 years  
Data Types: Body temp measurement, ECG measurement, Risk level

Processings:

- \* Patient data collection, Collection
- \* Patient health risk assessment, AutomatedDecisionMaking
  - \* CVD storage, Storage
    - \* GP portal, FurtherProcessing
  - \* Emergency notification decision, AutomatedDecisionMaking
    - \* Disclosure to medical emergency services, Disclosure
- \* Improvement storage, Storage
  - \* GP portal, FurtherProcessing
- \* Research Storage, Storage
  - \* PMS research, FurtherProcessing

---

*Snippet 1: DPMF export of the record of processing activities for the last version of the PMS.*

This fragment shows an example export by the DPMF of the record of processing activities. This information should be readily available for controllers to maintain their record and to hand over to national supervisory authorities upon request.

and techniques to model versioning (Brosch et al. 2012). Model version control systems (Altmanninger et al. 2009a,b) mainly focus on the technical challenges of migrating, differencing, and merging diverging models in a semantically correct manner given the meta-models that underpin these MDE approaches, but rarely support the proper documentation of the rationale behind these changes. Additionally, core challenges related to the independent evolution of the meta-model and appropriate tooling remain high on the agenda (Paige et al. 2016). In related research in software architecture, similar challenges have been identified in architecture knowledge management (Babar et al. 2009; Kruchten et al. 2006; Weinreich and Groher 2016) to properly document not only the different versions of a software design, but also the design process and rationale, which is often implicit, and known to ‘evaporate’ over time (Feilkas et al. 2009).

The DPMF deliberately incorporates elements of legal argumentation in the models. For example, the «LawfulGround» and «CompatibilityAssessment» are not just descriptive but actually represent the outcome of a legal reasoning about the lawfulness of the processing activities and the second component of the purpose limitation principle, respectively. In our vision, the modeler has to explicitly encode these arguments to ensure they are documented and versioned.

The documentation of these elements can also prove useful from a legal perspective. For example, consider a new judicial precedent according to which certain processing purposes are incompatible with a specific lawful ground (e. g., collecting metrics data to improve the performance of a service is not objectively necessary for the performance of a specific contract (see Assessment 1.1)). Should that happen, the model could be easily queried to retrieve all purposes that were previously considered ‘*not incompatible*’ («CompatibilityAssessment» objects) with that lawful ground and thus enable a more efficient assessment of the impact of the new precedent on the modeled data processing operations.

## 8.5 Knowledge bases

The representation of the legal rationale such as the «CompatibilityAssessments» triggers the possibility of storing and reusing this knowledge across multiple models. When modeling similar systems, many previously defined purposes could be reused. By relying on existing «CompatibilityAssessments», the purpose limitation analysis can be optimized by not requiring re-assessment of already evaluated combinations.

In addition to the construction of knowledge bases from past experience in existing DPMs, there are a number of knowledge bases that could be constructed up-front and used across multiple DPMs to support other legal assessments. For example, the age at which a data subject qualifies as a ‘child’ is left up to Member States (Art. 8(1)). Without delving into the intricacies of private international law, that age mostly depends on (i) the habitual residence of the data subject or the country where the controller is established and (ii) the corresponding national transposing act. Since the former is already modeled in the DPM, a knowledge base can provide the relevant age for each Member State to streamline the qualification of the data subject as a child to facilitate Assessments 1.8 and 1.9.

Another example is the specification of the applicable Union or Member State law in cases where the processing activities are based on either «LegalObligation» or «PublicInterest» as a lawful ground (Art. 6(3)). The legislation may contain specific provisions governing, for instance, the lawfulness of processing, the types of data which are subject to the processing, etc. Compliance with these provisions could be facilitated by providing a knowledge base of standard scenarios based on some of the most relevant Union or Member State laws paired with a corresponding reference model of the processing operations.

## 9 Conclusion

In this paper, we have presented the Data Protection Modeling Framework (DPMF), a model-based approach for data protection that is rooted

upon the concepts and requirements imposed by the GDPR (European Union 2016) and related recommendations (Article 29 Working Party 2013a,b, 2014, 2017; European Data Protection Board 2020). The meta-model has been constructed following an in-depth, interdisciplinary analysis and investigation of these sources to determine the key concepts and ensure a common and unambiguous language to represent and reason about the processing operations.

The presented modeling approach supports and enforces soundness and completeness criteria to ensure that DPM models are sufficiently complete and accurate and thus can be used as comprehensive descriptions of data processing operations in the context of a DPIA exercise. Additionally, the approach streamlines and automates a number of complementary legal assessments, however, under the clear vision that these steps rely heavily upon the argumentations and manual assessment inputs provided by the legal stakeholders conducting these assessments. By explicitly documenting these argumentations as part of the modeled DPMs, the legal rationale is recorded alongside the model, which in turn helps in meeting the accountability requirements imposed by the GDPR.

These contributions were subsequently validated in the context of a real-world eHealth application to demonstrate the value of tool-supported model construction and analysis. Adopting a model-based approach instead of a document- or template-based approach – yet ensuring that these can in fact be generated from the constructed DPMs – is a strong prerequisite towards longer term support and evolution of these models. Indeed, the model-based approach promotes the central role of these models throughout the development life cycle and throughout the lifetime of the described data processing activities.

The DPMF as presented in this paper is the first step of a long-term research effort, and future extensions will include: (i) explicit support for modeling and assessing privacy countermeasures from a legal, organizational, and system engineering perspectives; (ii) risk assessment and

quantification from different, complementary perspectives: risks of non-compliance, risks for data subject's rights and freedoms, and monetary risk in terms of fines or business loss; (iii) support for dynamic data protection impact re-assessment triggered, for example, by run-time changes in a system or service; and (iv) an analysis of the necessary symbiotic relationship between data protection impact assessments, and the methods and techniques of privacy engineering (e. g., threat analysis, systematic mitigation with the use of Privacy-Enhancing technologies (PETs), etc.).

## References

- Agarwal S., Steyskal S., Antunovic F., Kirrane S. (2018) Legislative Compliance Assessment: Framework, Model and GDPR Instantiation. In: Privacy Technologies and Policy (APF 2018). Lecture Notes in Computer Science Vol. 11079. Springer, pp. 131–149
- Agence Nationale de la Sécurité des Systèmes d'Information (ANSSI) (2010) EBIOS — Expression des Besoins et Identification des Objectifs de Sécurité <https://www.ssi.gouv.fr/guide/ebios-2010-expression-des-besoins-et-identification-des-objectifs-de-securite/>
- Ahmadian A. S., Jürjens J., Strüber D. (2018a) Extending Model-Based Privacy Analysis for the Industrial Data Space by Exploiting Privacy Level Agreements. In: Proceedings of ACM SAC 2018: PDP. ACM, pp. 1142–1149
- Ahmadian A. S., Strüber D., Riediger V., Jürjens J. (2018b) Supporting privacy impact assessment by model-based privacy analysis. In: Proceedings of ACM SAC 2018: Software Engineering. ACM, pp. 1467–1474
- Akarion AG (2019) NioBase: Master the GDPR. <https://niobase.com/en/>

Alnemr R., Cayirci E., Dalla Corte L., Garaga A., Leenes R., Mhungu R., Pearson S., Reed C., de Oliveira A. S., Stefanatou D., et al. (2015) A data protection impact assessment methodology for cloud. In: Annual Privacy Forum. Lecture Notes in Computer Science Vol. 9484. Springer, pp. 60–92

Alshammari M., Simpson A. (2018) A model-based approach to support privacy compliance. In: Information & Computer Security 26(4), pp. 437–453

Altmanninger K., Brosch P., Kappel G., Langer P., Seidl M., Wieland K., Wimmer M. (2009a) Why model versioning research is needed!? an experience report. In: Proceedings of the Joint MoDSE-MCCM 2009 Workshop Vol. 9. Springer, pp. 1–12

Altmanninger K., Seidl M., Wimmer M. (2009b) A survey on model versioning approaches. In: International Journal of Web Information Systems 5(3), pp. 271–304

Antignac T., Scandariato R., Schneider G. (2016) A privacy-aware conceptual model for handling personal data. In: Lecture Notes in Computer Science 9952, pp. 942–957

APD (2018a) Modèle de registre des activités de traitement Autorité de Protection des Données <https://www.autoriteprotectiondonnees.be/canevas-de-registre-des-activites-de-traitement>

APD (2018b) Recommandation n° 01/2018 du 28 février 2018 concernant l'analyse d'impact relative à la protection des données et la consultation préalable Autorité de Protection des Données [https://www.autoriteprotectiondonnees.be/sites/privacycommission/files/documents/recommandation\\_01\\_2018.pdf](https://www.autoriteprotectiondonnees.be/sites/privacycommission/files/documents/recommandation_01_2018.pdf)

Article 29 Working Party (2013a) Opinion 03/2013 on purpose limitation (WP203). Article 29 Working Party. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf)

Article 29 Working Party (2013b) Opinion 07/2013 on the Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems (WP209). Article 29 Working Party. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp209\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp209_en.pdf)

Article 29 Working Party (2014) Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC (WP217). Article 29 Working Party. [https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217\\_en.pdf](https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf)

Article 29 Working Party (2017) Guidelines on Data Protection Impact Assessment (DPIA) (WP248 rev.01). Article 29 Working Party. [http://ec.europa.eu/newsroom/document.cfm?doc\\_id=47711](http://ec.europa.eu/newsroom/document.cfm?doc_id=47711)

Article 29 Working Party (2018a) Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (WP251 rev.01). Article 29 Working Party. [https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc\\_id=49826](https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826)

Article 29 Working Party (2018b) Guidelines on transparency under Regulation 2016/679. Article 29 Working Party. [https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc\\_id=51025](https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025)

AvePoint (2019) AvePoint Privacy Impact Assessment System: Comply with GDPR and other key data protection regulations <https://www.avepoint.com/privacy-impact-assessment/>

Babar M. A., Dingsoyr T., Lago P., van Vliet H. (2009) Software Architecture Knowledge Management: Theory and Practice. Springer

Bayamlioglu E. (2018) Contesting Automated Decisions: A View of Transparency Implications. In: European Data Protection Law Review 4(4), pp. 433–446

Berger B., Sohr K., Koschke R. (2016) Automatically extracting threats from extended data flow diagrams. In: *Lecture Notes in Computer Science* 9639, pp. 56–71

Bieker F., Friedewald M., Hansen M., Obersteller H., Rost M. (2016) A process for Data Protection Impact Assessment under the European General Data Protection Regulation. In: *APF 2016: Privacy Technologies and Policy. Lecture Notes in Computer Science* Vol. 9857. Springer, pp. 21–37

Bitkom (2017) Risk Assessment und Datenschutz-Folgenabschaetzung <https://www.bitkom.org/sites/default/files/pdf/NP-Themen/NP-Vertrauen-Sicherheit/Datenschutz/FirstSpirit-1496129138918170529-LF-Risk-Assessment-online.pdf>

Blanco-Lainé G., Sottet J.-S., Dupuy-Chessa S. (2019) Using an Enterprise Architecture Model for GDPR Compliance Principles. In: *The Practice of Enterprise Modeling (PoEM 2019). Lecture Notes in Business Information Processing* Vol. 369. Springer, pp. 199–214

Breaux T., Antón A. (2008) Analyzing Regulatory Rules for Privacy and Security Requirements. In: *IEEE Transactions on Software Engineering* 34(1), pp. 5–20

Breaux T. D., Vail M. W., Anton A. I. (2006) Towards Regulatory Compliance: Extracting Rights and Obligations to Align Requirements with Regulations. In: *14th IEEE International Requirements Engineering Conference (RE'06)*. IEEE, pp. 49–58

Brosch P., Kappel G., Langer P., Seidl M., Wieland K., Wimmer M. (2012) An introduction to model versioning. In: *Formal Methods for Model-Driven Engineering (SFM 2012). Lecture Notes in Computer Science* Vol. 7320. Springer, pp. 336–398

Brüggemann T., Hansen J., Dehling T., Sunyaev A. (2016) An Information Privacy Risk Index for mHealth Apps. In: *Privacy Technologies and Policy. Lecture Notes in Computer Science* Vol. 9857. Springer, pp. 190–201

Casey B., Farhangi A., Vogl R. (2019) Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise en. In: *Berkeley Technology Law Journal* 34(1), p. 46

CNIL (2018a) PIA: Analyse d'impact sur la protection des données (Privacy Impact Assessment) Commission Nationale de l'Informatique et des Libertés <https://www.cnil.fr/en/open-source-pia-software-helps-carry-out-data-protection-impact-assesment>

CNIL (2018b) Privacy Impact Assessment (PIA) 1 : Methodology. Commission Nationale de l'Informatique et des Libertés. <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-en-methodology.pdf>

CNIL (2018c) Privacy Impact Assessment (PIA) 2 : Template. Commission Nationale de l'Informatique et des Libertés. <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-en-templates.pdf>

CNIL (2018d) Privacy Impact Assessment (PIA) 3 : Knowledge Bases. Commission Nationale de l'Informatique et des Libertés. <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-3-en-knowledgebases.pdf>

Coles J., Faily S., Ki-Aries D. (2018) Tool-supporting Data Protection Impact Assessments with CAIRIS. In: *2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE)*. IEEE, pp. 21–27

Colesky M., Hoepman J. H., Hillen C. (2016) A Critical Analysis of Privacy Design Strategies. In: *2016 IEEE Security and Privacy Workshops (SPW)*. IEEE, pp. 33–40

Compagna L., El Khoury P., Krausová A., Mascacci F., Zannone N. (2009) How to integrate legal requirements into a requirements engineering methodology for the development of security and privacy patterns. In: *Artificial Intelligence and Law* 17(1), pp. 1–30

DeMarco T. (1979) *Structured Analysis and System Specification*. Yourdon

Deng M., Wuyts K., Scandariato R., Preneel B., Joosen W. (2011) A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements. In: *Requirements Eng* 16(1), pp. 3–32

Dewitte P., Wuyts K., Sion L., Van Landuyt D., Emanuilov I., Valcke P., Joosen W. (2019) A Comparison of System Description Models for Data Protection by Design. In: *Proceedings of ACM SAC: PDP*. ACM, pp. 1512–1515

Dhillon D. (2011) Developer-Driven Lessons Learned in the Trenches. In: *IEEE Security Privacy* 9(4), pp. 41–47

European Data Protection Board (2019) *Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects*. European Data Protection Board. [https://edpb.europa.eu/sites/edpb/files/files/file1/edpb\\_guidelines-art\\_6-1-b-adopted\\_after\\_public\\_consultation\\_en.pdf](https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf)

European Data Protection Board (2020) *Guidelines 05/2020 on consent under Regulation 2016/679*. European Data Protection Board. [https://edpb.europa.eu/sites/edpb/files/files/file1/edpb\\_guidelines\\_202005\\_consent\\_en.pdf](https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf)

European Data Protection Supervisor (2018) *Accountability on the ground Part II: Data Protection Impact Assessments & Prior Consultation*. European Data Protection Supervisor. [https://edps.europa.eu/sites/edp/files/publication/18-02-06\\_accountability\\_on\\_the\\_ground\\_part\\_2\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/18-02-06_accountability_on_the_ground_part_2_en.pdf)

European Data Protection Supervisor (2020) *A Preliminary Opinion on data protection and scientific research*. European Data Protection Supervisor. [https://edps.europa.eu/sites/edp/files/publication/20-01-06\\_opinion\\_research\\_en.pdf](https://edps.europa.eu/sites/edp/files/publication/20-01-06_opinion_research_en.pdf)

European Union (2016) *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016*. In: *Official Journal of the EU*, pp. 1–88

Feilkas M., Ratiu D., Jurgens E. (2009) The loss of architectural knowledge during system evolution: An industrial case study. In: *2009 IEEE 17th International Conference on Program Comprehension*. IEEE, pp. 188–197

Ferra F., Wagner I., Boiten E., Hadlington L., Psychoula I., Snape R. (2020) Challenges in assessing privacy impact: Tales from the front lines. In: *Security and Privacy* 3(2), pp. 1–19

Fotiou N., Arianfar S., Särelä M., Polyzos G. C. (2014) A Framework for Privacy Analysis of ICN Architectures. In: *Privacy Technologies and Policy. Lecture Notes in Computer Science Vol. 8450*. Springer, pp. 117–132

Freund J., Jones J. (2014) *Measuring and Managing Information Risk: A FAIR Approach*. Butterworth-Heinemann

Gellert R. (2018) Understanding the notion of risk in the General Data Protection Regulation. In: *Computer Law & Security Review* 34(2), pp. 279–288

Ghanavati S., Amyot D., Rifaut A. (2014) Legal goal-oriented requirement language (legal GRL) for modeling regulations. In: *Proceedings of the 6th International Workshop on Modeling in Software Engineering (MiSE 2014)*. ACM, pp. 1–6

Goodman B., Flaxman S. (2017) European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”. In: *AI Magazine* 38(3), pp. 50–57

Howard M., Lipner S. (2006) *The Security Development Lifecycle*. Microsoft Press

Information Commissioner’s Office (2019) *What are the rules on special category data*. <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-are-the-rules-on-special-category-data/>

- Islam S., Mouratidis H., Wagner S. (2010) Towards a Framework to Elicit and Manage Security and Privacy Requirements from Laws and Regulations. In: Requirements Engineering: Foundation for Software Quality. Lecture Notes in Computer Science Vol. 6182. Springer, pp. 255–261
- Joyee De S., Le Métayer D. (2016) PRIAM: A Privacy Risk Analysis Methodology. In: Data Privacy Management and Security Assurance (DPM 2016, QASA 2016). Lecture Notes in Computer Science Vol. 9963. Springer, pp. 221–229
- Kruchten P., Lago P., van Vliet H. (2006) Building Up and Reasoning About Architectural Knowledge In: Quality of Software Architectures (QoSA 2006) Vol. 4214 Lecture Notes in Computer Science Springer, pp. 43–58
- Lund M. S., Solhaug B., Stølen K. (2010) Model-driven risk analysis: the CORAS approach. Springer
- Malgieri G., Comandé G. (2017) Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation en. In: International Data Privacy Law 7(4), pp. 243–265
- Meis R., Wirtz R., Heisel M. (2015) A Taxonomy of Requirements for the Privacy Goal Transparency. In: Trust, Privacy and Security in Digital Business. Lecture Notes in Computer Science Vol. 9264. Springer, pp. 195–209
- Muntés-Mulero V., Dominiak J., González E., Sanchez-Charles D. (2019) Model-driven Evidence-based Privacy Risk Control in Trustworthy Smart IoT Systems. In: Model-Driven Engineering for the Internet of Things (MDE4IoT) & Interplay of Model-Driven and Component-Based Software Engineering (ModComp). CEUR-WS.org
- Muthuri R., Boella G., Hulstijn J., Humphreys L. (2016) Argumentation-Based Legal Requirements Engineering: The Role of Legal Interpretation in Requirements Acquisition. In: 2016 IEEE 24th International Requirements Engineering Conference Workshops (REW). IEEE, pp. 249–258
- Nuseibeh B. (2001) Weaving together requirements and architectures. In: Computer 34(3), pp. 115–119
- Nymity (2019) Automate the entire PIA process and Provide Better Response Time to the Business with Improved Business Engagement. <https://www.nymity.com/wp-content/uploads/Nymity-PIA-DPIA-Datasheet.pdf>
- Oetzel M. C., Spiekermann S. (2014) A systematic methodology for privacy impact assessments: A design science approach. In: European Journal of Information Systems 23(2), pp. 126–150
- Oliver I. (2014) Privacy engineering: A dataflow and ontological approach. CreateSpace Independent Publishing Platform
- OneTrust (2019) OneTrust Privacy Management Software. <https://www.onetrust.com/solutions/privacy-compliance/>
- Paige R. F., Matragkas N., Rose L. M. (2016) Evolving models in model-driven engineering: State-of-the-art and future challenges. In: Journal of Systems and Software 111, pp. 272–280
- Palmirani M., Martoni M., Rossi A., Bartolini C., Robaldo L. (2018) Legal Ontology for Modelling GDPR Concepts and Norms. In: Legal Knowledge and Information Systems. Frontiers in Artificial Intelligence and Applications Vol. 313. IOS Press, pp. 91–100
- Rahman M. (2017) A Petri Nets Semantics for Privacy-Aware Data Flow Diagrams. Department of computer Science and Engineering, Gothenburg University. [https://gupea.ub.gu.se/bitstream/2077/53077/1/gupea\\_2077\\_53077\\_1.pdf](https://gupea.ub.gu.se/bitstream/2077/53077/1/gupea_2077_53077_1.pdf)
- RealDPG (2019) RealDPG Features and Benefits. <https://www.realdpg.com/en/features-benefits>
- Selbst A. D., Powles J. (2017) Meaningful information and the right to explanation. In: International Data Privacy Law 7(4), pp. 233–242
- Shostack A. (2008) Experiences threat modeling at Microsoft. In: Proceedings of the Workshop on Modeling Security (MODSEC'08). CEUR Workshop Proceedings Vol. 413. CEUR-WS.org

Shostack A. (2014) *Threat Modeling: Designing for Security*. Wiley

Siena A., Perini A., Susi A., Mylopoulos J. (2009) A Meta-Model for Modelling Law-Compliant Requirements. In: 2009 Second International Workshop on Requirements Engineering and Law. IEEE, pp. 45–51

Sion L., Dewitte P., Van Landuyt D., Wuyts K., Emanuilov I., Valcke P., Joosen W. (2019a) An Architectural View for Data Protection by Design. In: 2019 IEEE International Conference on Software Architecture (ICSA). IEEE, pp. 11–20

Sion L., Van Landuyt D., Wuyts K., Joosen W. (2019b) Privacy Risk Assessment for Data Subject-aware Threat Modeling. In: 2019 IEEE Security and Privacy Workshops (SPW). IEEE, pp. 64–71

Sion L., Yskout K., Van Landuyt D., Joosen W. (2018a) Risk-based Design Security Analysis. In: Proceedings of the 1st International Workshop on Security Awareness from Design to Deployment. IEEE, pp. 11–18

Sion L., Yskout K., Van Landuyt D., Joosen W. (2018b) Solution-aware Data Flow Diagrams for Security Threat Modeling. In: Proceedings of ACM SAC: Software Architecture: Theory, Technology, and Applications. ACM, pp. 1425–1432

Steinberg D., Budinsky F., Merks E., Paternostro M. (2008) *EMF: eclipse modeling framework*. Pearson Education

Tom J., Sing E., Matulevičius R. (2018) Conceptual Representation of the GDPR: Model and Application Directions. In: Perspectives in Business Informatics Research (BIR 2018). Lecture Notes in Business Information Processing Vol. 330. Springer, pp. 18–28

Torre D., Soltana G., Sabetzadeh M., Briand L. C., Auffinger Y., Goes P. (2019) Using Models to Enable Compliance Checking Against the GDPR: An Experience Report. In: 2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems (MODELS), pp. 1–11

Tuma K., Scandariato R., Widman M., Sandberg C. (2017) Towards security threats that matter. In: 3rd Workshop On The Security Of Industrial Control Systems & Of Cyber-Physical Systems (CyberICPS 2017). Lecture Notes in Computer Science Vol. 10683. Springer, pp. 47–62

ULD (2017) The Standard Data Protection Model: A concept for inspection and consultation on the basis of unified protection goals. Unabhängiges Landeszentrum für Datenschutz. [https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology\\_V1.0.pdf](https://www.datenschutzzentrum.de/uploads/sdm/SDM-Methodology_V1.0.pdf)

Wachter S., Mittelstadt B., Floridi L. (2017) Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. In: International Data Privacy Law 7(2), pp. 76–99

Weinreich R., Groher I. (2016) Software architecture knowledge management approaches and their support for knowledge management activities: A systematic literature review. In: Information and Software Technology 80, pp. 265–286

Wright D., Friedewald M., Gellert R. (2014) Developing and testing a surveillance impact assessment methodology. In: International Data Privacy Law 5(1), pp. 40–53

Table 6: Glossary of the Data Protection Modeling Concepts

Concept	Description of the concept
<b>Actors and LegalRoles</b>	
👤 «Actor»	Entity involved in a «Processing» of personal data.
👤 «Representative»	Entity representing an «Actor» not established in the EU (Art. 27).
👤 «LegalRole»	Role of an «Actor» involved in a «Processing» of personal data («Controller», «Processor», «Recipient», «ThirdParty» or «Representative» (Art. 4(7–10,17)).
👤 «Controller»	«LegalRole» of the «Actor» that, alone or jointly with others, determines the purposes and the means of the processing of personal data (Art. 4(7)).
👤 «Processor»	«LegalRole» of the «Actor» that processes personal data on behalf of a «Controller» (Art. 4(8)).
👤 «Recipient»	«LegalRole» of the «Actor» to which the personal data are disclosed (Art. 4(9)).
👤 «ThirdParty»	«LegalRole» of the «Actor» that is neither a data subject, the «Controller», or the «Processor» with regard to the «Processings» at stake (Art. 4(10)).
<b>Processing Operations</b>	
⚙️ «Processing»	Any operation performed on personal data by an «Actor» (Art. 4(2)).
📂 «Collection»	Initial «Processing»—whether the personal data are collected directly from the data subject or obtained from another «Actor».
⚙️ «FurtherProcessing»	All «Processings» performed after the initial «Collection».
📂 «Storage»	«FurtherProcessing» that consists of the persistent storage of the personal data and that requires specifying either a <i>retentionPeriod</i> for the data or the <i>retentionCriteria</i> for determining how long they will be kept (Art. 13(2)a; 14(2)a).
⚙️ «AutomatedDecisionMaking»	«FurtherProcessing» that consists of a decision based solely on automated processing, including profiling, which produces legal effects concerning the data subject or similarly significantly affects him/her (Art. 22(1)).
📄 «Disclosure»	«FurtherProcessing» that consists of the disclosure of personal data to a <i>recipient</i> .
<b>Data and Data Subjects</b>	
📂 «DataSet»	List of «PersonalDataTypes» used as input or output of a «Processing».
📂 «PersonalDataTypes»	Types of personal data being processed.
📂 «Regular»	All «PersonalDataTypes» that do not fall into one of the next two subtypes.
📂 «SpecialCategory»	«PersonalDataTypes» revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, and genetic data, biometric data, data concerning health or data concerning a natural person's sex life or sexual orientation. Those are subject to a general prohibition of processing <i>except</i> when one of the «ProhibitionExemptionTypes» (Art. 9(2)) is applicable.
📂 «CriminalConvictionsAndOffences»	«PersonalDataTypes» related to criminal convictions and offences. Those require specific supervision or authorization (Art. 10).
👤 «DataSubjectType»	Types of natural persons whose personal data are processed (Art. 4(1)).
<b>Lawful grounds and processing purposes</b>	
«LawfulGround»	Lawful ground for the «Collection» (in support of the <i>lawfulness</i> principle). Specific types: 👤: «Consent»      📄: «Contract»      📄: «LegalObligation» 👤: «LegitimateInterests»      📄: «PublicInterest»      📄: «VitalInterests»
👤 «ProcessingPurpose»	Purpose for which the personal data are processed (in support of the <i>purpose limitation</i> principle (Art. 5(1)b)).
📄 «CompatibilityAssessment»	Outcome of the «CompatibilityAssessment» that specifies whether the «ProcessingPurpose» of a «FurtherProcessing» is compatible with the «ProcessingPurpose» of the «Collection» and includes legal argumentation.

Table 7: Overview of the meta-model constraints and soundness constraints

Constraint	Description of the constraint
<b>Meta-Model Constraints</b>	
1 Start with collection	Every chain of «Processings» needs to start with an initial «Collection».
2 Collection has lawful ground/purpose	Every «Collection» needs to specify a «LawfulGround» and at least one «ProcessingPurpose».
3 Further processing has purpose	Every «FurtherProcessing» needs to specify a «ProcessingPurpose».
4 Special category requires exemption	Every «SpecialCategory» of processed personal data needs to specify a «ProhibitionExemptionType».
5 Legal role has actor	Every «LegalRole» needs to specify an «Actor».
6 Disclosure has recipient	Every «Disclosure» needs to specify the <i>recipient</i> to whom the data is disclosed.
<b>Soundness Constraints</b>	
1 Input/output consistency	Every input «PersonalDataType» in a «FurtherProcessing» must be the output of an earlier «Processing».
2 EU representative	Every «Actor» that is a «Controller» or «Processor» and that is not established in the EU needs to specify a representative in the EU unless it is a public authority.
3 Collection has further processing	Every «Collection» should be followed by one or more «FurtherProcessings».
4 Indirect collection consistency	The «Actor» controlling the disclosure to a <i>recipient</i> should be consistent with «Actor» providing the data to the indirect «Collection» corresponding with that «Disclosure».
5 Every processing has controller	Every «Processing» must specify a «Controller» in its set of «LegalRoles».
6 Every processor has controller	Every «Processing» with a «Processor», must also specify a «Controller».
7 Lawfulness of public task	A public authority that performs a public task cannot rely on «LegitimateInterests» as «LawfulGround».
8 Medical data—professional secrecy	A «Controller» processing «SpecialCategory» for medical purposes must be subject to professional secrecy.
9 Storage has retention period/criteria	Every «Storage» must specify either a <i>retentionPeriod</i> or the <i>retentionCriteria</i> to determine that period.
10 Automated decisions—special categories	Every «AutomatedDecisionMaking» must be based on «Contract», «LegalObligation», or <i>explicit</i> «Consent» as «LawfulGround».

Table 8: Legal Assessments in Support of DPIA

Nbr	A	Assessment	Description
1	②	Lawfulness	If the «ProcessingPurpose» of the «FurtherProcessing» is <i>not incompatible</i> with the «ProcessingPurpose» of the «Collection», no additional «LawfulGround» needs to be specified. If it is incompatible, a new «LawfulGround» and «ProcessingPurpose» must be specified for the «Collection».
1.1	②	Performance of a contract, legal obligation, vital interests, public task, interest, legitimate interests—Necessity	«PublicInterest» or «LegitimateInterests», the relevant details must be specified and the «ProcessingPurpose» of all the «Processings» must be objectively necessary.
1.2	②	Compliance with a legal obligation to which the controller is subject—Existence of Union or Member State law	If a «Processing» is based on either «LegalObligation» or «PublicInterest» as «LawfulGround», the corresponding Union or Member State law that is applicable must be specified in the <i>law</i> attribute of the corresponding «LawfulGround».
1.3	②	Legitimate interests of the controller—Balancing test	If a «Processing» is based on «LegitimateInterests», the interests of the controller and the interests or fundamental rights and freedoms of the data subjects must be documented and balanced against each other.
1.4	②	Consent—Attributes	Model Constraint 2 allows the consent to be “specific”.
1.5	②	Consent—Demonstrability	Model Constraint 2 allows the «Actor» to identify the «Processings» for which it is necessary to implement measures to demonstrate that the consent has been given by the data subject.
1.6	②	Consent—Separation and intelligibility	Model Constraint 2 allows the «Actor» to identify the «Processings» for which it is necessary to implement measures to demonstrate that the consent by the data subject for the «Processing» has been given separately from the other matters.
1.7	②	Consent—Withdrawal	Model Constraint 2 allows the «Actor» to identify the «Processings» for which it is necessary to implement measures to easily allow the data subject to withdraw his/her consent.
1.8	③	Consent—Age requirement	In case of a child, the «Processing» can only be based on «Consent» as a «LawfulGround» if it is either <i>authorisedConsent</i> or <i>parentalConsent</i> .
1.9	②	Consent—Verification of parental authorization or consent	Assessment 1.8 allows the «Actor» to identify the «Processings» for which it is necessary to implement measures to verify that the authorization or consent has been given by the holder of parental responsibility.
2.1	③	Purpose specification	Model Constraint 2 supports the requirement for each «Collection» of «PersonalDataTypes» to specify a «LawfulGround» and a «ProcessingPurpose».
2.2	②	Compatibility assessment	The «ProcessingPurposes» of the «FurtherProcessings» must not be incompatible with the «ProcessingPurposes» of the «Collection».
2.3	③	Presumption of non-incompatibility for further processing for archiving purposes in the public interest, scientific or historical research, or statistical purposes	If the «ProcessingPurpose» of a «FurtherProcessing» is «ArchivingInThePublicInterestScientificOrHistoricalResearchOrStatisticalPurpose», then it must not, <i>per se</i> , be considered incompatible with the «ProcessingPurpose» of the «Collection», provided that the necessary safeguards are implemented.

Legend: A: Level of automation provided by the DPMF. ① A DPM provides little to no information to perform the legal assessment. ② A DPM provides relevant information (in terms of the colored concepts of the meta-model) but the legal assessment has to be performed separately and does not involve any change in the DPM. ③ A DPM allows extension with legal rationale (in terms of the grey concepts of the meta-model) so that the legal assessment can be performed on the basis of the DPM itself.

Table 9: Legal Assessments in Support of DPIA (continued)

Nbr	A	Assessment	Description
3	②	Data minimization	For each «Processing», the «PersonalDataTypes» must be strictly necessary to achieve the «ProcessingPurpose» that have been specified following Assessment 2.1.
4.1	②	Storage limitation—Necessity	For each «Storage» of a «DataSet», the <i>retentionPeriod</i> or the <i>retentionCriteria</i> must not be longer than what is necessary to achieve the «ProcessingPurposes» of the «Collection».
4.2	③	Exemption for further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes	If the «ProcessingPurpose» of a «Storage» is «ArchivingInThePublicInterestScientificOrHistoricalResearchOrStatisticalPurpose», then the <i>retentionPeriod</i> of that «DataSet» can be longer than necessary to achieve the «ProcessingPurposes» of the «Collection».
5	②	Processing of special categories of personal data—Exemption to the general prohibition—Necessity	If the «ProhibitionExemptionType» is anything but <i>explicitConsent</i> , the relevant details of the «ProhibitionExemptionType» must be specified and the «ProcessingPurpose» of the «Processing» of those «SpecialCategory» of personal data must be objectively necessary.
6	②	Processing of personal data relating to criminal convictions and offences	If personal data related to «CriminalConvictionsAndOffences» are processed, that specific «Processing» can only happen under the control of official authority or if the applicable Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects is specified.
7.1	②	Measures to be implemented in case of exemption to the general prohibition	If the «AutomatedDecisionMaking» is based on either «Contract» or <i>explicitConsent</i> , the corresponding «Actor» must also implement suitable measures to allow the data subjects to, at least, obtain human intervention of the controller, express their point of view, and contest the automated decision.
7.2	③	Special categories of personal data	An «AutomatedDecisionMaking» can only have as an input a «DataSet» that contains «SpecialCategory» of personal data if the specified «ProhibitionExemptionType» is either <i>explicitConsent</i> or <i>substantialPublicInterest</i> .
8	②	Joint controllers	If there is more than one «Controller» for a specific «Processing», those «Actors» are considered as joint controllers with regard to that «Processing».
9	③	Representative of controllers or processors not established in the EU	If an «Actor» with the role of «Controller» or «Processor» has <i>establishedInEU</i> = FALSE and <i>publicAuthority</i> = FALSE, then it must be represented by a «Representative» in the Union.
10	②	Prohibition to engage another processor without prior specific or general approval of the controller	When a «Processor» engages another «Processor» acting as a sub-processor, it must have the prior specific or general written authorisation of the «Controller» on whose behalf it acts. In case of general authorisation, the «Processor» shall also inform the «Controller» of any change.
11	②	Controller-processor agreement	When there is a «Processor», there must be a contract (i. e. a controller-processor agreement) with the «Controller» that contains all the elements listed in Art. 28(3).

Legend: A: Level of automation provided by the DPMF. ① A DPM provides little to no information to perform the legal assessment. ② A DPM provides relevant information (in terms of the colored concepts of the meta-model) but the legal assessment has to be performed separately and does not involve any change in the DPM. ③ A DPM allows extension with legal rationale (in terms of the grey concepts of the meta-model) so that the legal assessment can be performed on the basis of the DPM itself.

Table 10: Legal Assessments in Support of DPIA (continued)

Nbr	A	Assessment	Description
12.1	②	Adequacy decision	In case of a «Disclosure» to an «Actor» not establishedInEU or an internationalOrganization, there must be an adequacy decision issued by the European Commission concerning the country of the recipient or the international organization.
12.2	②	Appropriate safeguards	In case of a «Disclosure» to an «Actor» not establishedInEU or an internationalOrganization and there is no adequacy decision, then the «Actor» disclosing the personal data must provide appropriate safeguards as required by Art. 46(2).
13	①	Record of processing activities	Since the DPMF allows the «Controller» to keep track of all the «Processings», it facilitates compliance with the obligation to maintain a record of processing activities.
14.1	①	Towards data subjects	The mapping of all the «Actors», «Processings», «LawfulGrounds», «ProcessingPurposes» and «PersonalDataTypes» facilitates the compliance with the transparency obligations detailed in Art. 12–14.
14.2	①	Timing—Personal data collected from the data subject	In case of a direct «Collection», the «Controller» must provide all the information listed in Art. 13 to the data subject at the time of the «Collection».
14.3	①	Timing—Personal data not obtained from the data subject	In case of an indirect «Collection», the recipient acting as «Controller» must provide the information listed in Art. 14 to the data subject either within a reasonable period not exceeding one month, at the time of the first communication with the data subject, or at the time of the «Disclosure».
15	①	Integrity & Confidentiality	For every «Processing», the appropriate technical and organizational measures must be implemented to protect against unauthorized or unlawful processing and accidental loss, destruction, or damage.
16	①	Accuracy	For each «Processing», the associated «DataSet» must be accurate with regard to the «ProcessingPurpose». The «Actor» must implement measures to erase or rectify inaccurate «DataSets».
17	①	Security of processing	The DPMF provides guidance in raising the issue of providing appropriate security measures for the processing and can assist in identifying the «Processing» operations with the highest risk to the data subjects because of, for instance, the sensitivity of the information being processed.

Legend: A: Level of automation provided by the DPMF. ① A DPM provides little to no information to perform the legal assessment. ② A DPM provides relevant information (in terms of the colored concepts of the meta-model) but the legal assessment has to be performed separately and does not involve any change in the DPM. ③ A DPM allows extension with legal rationale (in terms of the grey concepts of the meta-model) so that the legal assessment can be performed on the basis of the DPM itself.