An Interdisciplinary Methodology to Validate Formal Representations of Legal Text Applied to the GDPR

Cesare Bartolini[1], Gabriele Lenzini[1], and Cristiana Santos[2]

[1] University of Luxembourg,
Interdisciplinary Centre for Security, Reliability and Trust (SnT) *
{cesare.bartolini, gabriele.lenzini}@uni.lu
[2] Research Centre for Justice and Governance (JusGov)
School of Law, University of Minho, Portugal
cristiana.teixeirasantos@gmail.com

Abstract. The modelling of a legal text into a machine-processable form, such as a list of logic formulæ, enables a semi-automatic reasoning about legal compliance but might entail some anticipation of legal interpretation in the modelling. The formulæ need therefore to be validated by legal experts, but it is unlikely that they are familiar with the formalism used. This calls for an interdisciplinary validation methodology to ensure that the model is legally coherent with the text it aims to represent but that could also close the communication gap between formal modellers and legal evaluators. This paper discusses such a methodology, providing an human-readable representation that preserves the formulæ's meaning but that presents them in a way that is usable by non-experts. We exemplify the methodology on a use case where Articles of the GDPR are translated in the Reified I/O logic encoded in LegalRuleML.

Keywords: General Data Protection Regulation; GDPR; data protection; legal validation; usability.

1 Introduction

Representing legal text in a machine-readable form is a well-consolidated field of research and industry. Supported by plenty of adequate models and tools [5], it has facilitated knowledge base building, enabled information retrieval processes, and supported decision-making.

Modelling legal text in a machine-readable format is also preparatory to a semi-automatic reasoning about legal compliance. Machines do not deliberate, but the logical steps they follow to assess compliance can increase (or decrease) one's confidence in a legal argumentation. This seems to be an interesting application in relation to the newly-established General Data Protection Regulation (GDPR), where computer scientists and lawyers are debating what measures and practices for data protection are compelled by law.

* Bartolini and Lenzini are supported by the FNR CORE project C16/IS/11333956 "DAPRECO: DAta Protection REgulation COmpliance".

As in other domains, modelling requires an interdisciplinary approach between domain experts and knowledge builders. But in the legal doctrine, modelling text faces additional challenges due to a widely-recognized fact: law is not merely the meaning of the text of its articles but the meaning that emerges from a legal interpretation of the text. Thus, any representation of the meaning convened by a text should be aligned with current interpretative consensus on that meaning as it defined by authoritative groups [7].

This calls for a *validation of any legal knowledge formalization*, and the validation should not be exhausted only within the discipline that does the modelling (computer science in our case). In fact:

(i) although the modeller may not be fully aware, in representing the legal text into a machine-processable format, he or she applies a certain degree of legally-critical interpretation. The modeller has to decide on whether a word or sentence represents a concept that has legal relevance; choose representative names for entities in the model; recognize if the text expresses an obligation, a permission, or a prohibition; and so on;

(ii) even though the interpretation required to encode a legal text into a machine-processable model may have a more restricted scope than a general interpretation of the law, it is unjustified to bury modelling decisions under a formalism that only expert modellers can master;

(iii) although the modeller may be confident of the interpretation he or she gives, concluding whether it is legally correct and complete should be the result of a rigorous process, and not of an individual self-assessment. In their turn, legal experts may want to apply their own qualitative criteria on the anticipated legal interpretations introduced in the model;

Besides, contemporary legal practitioners (such as policy-makers, judges, lawyers, administrators, and legal professionals) can be interested in verifying the results of a formal representation of the law and its applications, e.g., to find evidence in the legally-binding text or look for authoritativeness in knowledge representation [10]. Consequently, any legal knowledge base should be grounded in the reality of a juridical conceptualization of the law [4], promoting a reasonable threshold of reliability and authoritativeness to facilitate relevant applications that would transfer academic research to legal industry.

A legal validation is therefore a significant activity, but performing it over a formal model of a legal text, such as a logic, raises a significant issue: experts in law may not be prepared to fully understand the expressions encoded in the model, a situation that is aggravated if the formalism used is meant to be machine-readable rather than human-understandable. Validating the model requires therefore that the formalization is commonly understandable, and the methodology should be driven by usability considerations.

Promoting accessibility and understandability constitutes a sound practice in general; it is also aligned with the modern introduced codes of ethical conduct in data science[3] and with the GDPR's principles of algorithmic transparency, fairness, and accountability[4].

Contribution This work discusses a methodology for a legal validation of an existing machine-readable formalization of the GDPR. The input model is a set of XML files —referred to as the Data Protection Regulation Compliance (DAPRECO) knowledge base— containing the GDPR's provisions, formalized in Reified Input/Output (RIO) logic formulæ and embedded in LegalRuleML.

A key element in the methodology is an intermediate representation, supposedly human-readable, of the RIO logic formalization of the GDPR's Articles. This human-readable model is what the experts in law should be able to understand in order to give feedbacks on the consistency and completeness of the translation of the articles in the logic formalization.

The work is still preparatory: after introducing the methodology, it discusses two intermediate representations, together with an analysis of their understandability performed by testers, experts in law, who were involved in a usability experiment. One of the intermediate representations has been unanimously assessed as understandable. It can therefore be a potentially good candidate for a human-readable model to use in the execution of the methodology to collect reliable legal feedbacks on the validation of the formalized GDPR articles.

A full validation of the DAPRECO knowledge base is however left as future work. Many details have to be sorted out first, including improving the scalability of the editing of the human-readable representation, an activity that is currently in part performed automatically, but finalized by hand.

2 Related work

The validation phase of legal modeling by domain legal experts is, to the best of our knowledge, mentioned in the methodologies referring to ontological expert knowledge evaluation.

For example, the Methodology for Modeling Legal Ontologies (MeLOn) [9] was created to build legal ontologies in order to help legal experts model legal concepts, using the principles of data modification. Evaluation parameters consist in: i) completeness of the definition of the legal concepts; ii) correctness of the explicit relationships between legal concepts; iii) coherence of the legal concepts modelisation; iv) applicability to concrete use cases; v) effectiveness for the goals; vi) intuitiveness for the non-legal experts; vii) computational soundness of the logic and reasoning; viii) reusability of the ontology and mapping with other similar ontologies.

An *ad hoc* experimental validation by legal experts of a legal ontology, the Ontology of Professional Judicial Knowledge (OPJK), is described in [6] whereby

[3] `http://www.fatml.org/`.

[4] See GDPR, Art. 5.2 for the principle of accountability, and Artt. 13.2(f) and 14.2(g) for algorithmic transparency.

the expert validation conforms to phases. It includes i) a 48-question questionnaire whereby experts were asked to express their opinion regarding their level of agreement towards the ontology conceptualization and provide suggestions for improvement; and ii) an experimental validation based on a 10-item usability questionnaire, the System Usability Scale (SUS), tailored to evaluate the understanding and acceptance of the contents of the ontology. The adopted ontology evaluation questionnaire could offer rapid feedback and support towards the establishment of relevant agreement, shareability or quality of content measurements in expert-based ontology evaluation.

An evaluation methodology based on Competency Questions (CQs) has been proposed [13] to identify and evaluate the transformation of legal knowledge from a semi-formal form (Semantics Of Business Vocabulary And Rules - Standard English (SBVR-SE)) [8] to a more structured formal representation (OWL 2). It is based on a set of questions with regard to the subject-matter knowledge and predefined answers the ontology is supposed to give. The answers to such CQs could be viewed as requirements of a thus-constructed ontology. Such methodology can be a great tool to enable the cooperation between legal expert and knowledge modeller. Ontology quality criteria are accounted for.

However, there is a fundamental difference between the methodologies proposed to validate legal ontologies and this work. Ontologies are about concepts, data, entities, and a limited set of relations among them; validation of an ontology is therefore inevitably about assessing the qualities of those objects. Formal models for legal compliance, such as the DAPRECO knowledge base, represent a further step, as they model the logical and deontic structure of a legal text, modalities such as time, and constructs like those enabling a defeasible reasoning. The validation should take these elements into account.

Finally, usability is a requirement, since the formulæ must be validated by people who are not expert in logic. Existing validation processes cannot therefore be reused in the present work.

3 Background

The target of the validation proposed in this work is the DAPRECO Knowledge Base, which contains a formalization of the data protection legislation, particularly with respect to the GDPR. The knowledge base is made up of three interconnected main components: 1.) the legal text; 2.) the conceptual model; 3.) the deontic rules. It is meant to be used to provide a semi-automated assistance to the legal expert, all three components need to be machine-readable. For this reason, consolidated standards and reference formats have been used to model each of the three components.

The *legal text* is modelled in Akoma Ntoso[5], which makes it easy to navigate the document, referencing specific portions of text, using ordinary XML parsers.

[5] Currently stored at `https://github.com/guerret/lu.uni.dapreco.parser/blob/master/resources/akn-act-gdpr-full.xml`.

The *conceptual model* is contained in a legal ontology. specifically designed using the Web Ontology Language (OWL) language in an XML serialization.The ontology itself has been developed following the MeLOn methodology, which is based on a glossary and a set of CQs. The output of this work is an ontology of privacy and data protection, called Privacy Ontology (PrOnto) [12,11][6].

Finally, the *deontic rules* of the GDPR are expressed in Reified Input/Output (RIO) logic, which is an extension of Input/Output logic [14] using reification, a technique added to the logic to avoid nested obligations. A full description of the process followed to build the formulæ is described in [2]. This set of RIO formulæ% , their consistency and completeness regarding the legal are the real target of the validation task[7].

To express RIO formulæ in a machine-readable format, LegalRuleML[8] was used in modelling the formulæ. LegalRuleML is an XML markup language and a developing OASIS standard for representing the fine-grained semantic contents of legal texts, which has elements to represent legal content [1]. The formulæ act as a sort of *trait d'union* between the other two components, as they contain references both to ontological elements of the conceptual model and to the textual portions of the legal document expressed in Akoma Ntoso format.

3.1 How the machine-readable text looks like

The objective of the validation are the formulæ regardless of their expressive form (logic or LegalRuleML serialization). As RIO logic is an extension of Input/Output logic, all formulæ are if-then rules in the form (x, y), such that when x is given in input, y is returned in output. When applied to the legal domain, there are three sets to which rules can belong to: C is the set of constitutive norms, which defines when something counts as something else in the domain. Every pair $(x, y) \in C$ reads as "$x \rightarrow y$",as standard first-order logic implications; O and P are respectively the set of obligations and the set of permissions of the normative system. A pair $(x, y) \in O$ reads as "given x, y is obligatory", while a pair $(x, y) \in P$ reads as "given x, y is permitted".

Both the "if" and the "then" part of each formula are composed by a conjunction of predicates. Each predicate is in the form of the predicate name followed by a list of attributes. The name can be a concept belonging to an ontology (e.g., the PrOnto ontology) or it can be a logical operator. For example, $(\mathsf{PrOnto} : \mathsf{PersonalDataProcessing}\, x\, z)$ refers to a concept in the PrOnto ontology and takes two arguments. The predicate alone is incomplete, because it also needs to describe the two predicates used as arguments. If x is a controller and z some personal data of a data subject, an example may be formula 1.

[6] Currently stored at `https://github.com/guerret/lu.uni.dapreco.parser/blob/master/resources/pronto-v8.graphml`.
[7] The formulæ are available at `https://github.com/dapreco/daprecokb`.
[8] `http://ruleml.org/index.html`.

$$((\mathsf{prOnto:Controller}\,\mathrm{x}) \wedge (\mathsf{prOnto:DataSubject}\,\mathrm{w}) \wedge$$
$$(\mathsf{prOnto:PersonalData}\,\mathrm{z}\,\mathrm{w}) \wedge (\mathsf{prOnto:PersonalDataProcessing}'\,\mathrm{x}\,\mathrm{z})), \quad (1)$$

Furthermore, in RIO logic a predicate can be *reified* to be used as arguments for other predicates. Thus $(\mathsf{prOnto:PersonalDataProcessing}'\,e_p\,x\,z)$ is a new predicate, different from $(\mathsf{prOnto:PersonalDataProcessing}, x\,z)$; it represents the *possibility* that there is a processing of personal data. This allows e_p to be used as argument to another predicate.

In essence, each formula is expressed as a LegalRuleML rule containing two parts: premise (if) and the consequence (then). The predicates (and their arguments) composing both parts are serialized as RuleML atoms (and variables). The example above, with reification added, is serialized as follows:

Listing 1. LegalRuleML representation of formula 1.

```
<ruleml:Exists>
  <ruleml:Var key=":z">z</ruleml:Var>
  <ruleml:Var key=":x">x</ruleml:Var>
  <ruleml:And>
    <ruleml:Atom>
      <ruleml:Rel iri="prOnto:DataSubject"/>
      <ruleml:Var key=":w">w</ruleml:Var>
    </ruleml:Atom>
    <ruleml:Atom>
      <ruleml:Rel iri="prOnto:PersonalData"/>
      <ruleml:Var keyref=":z"/>
      <ruleml:Var keyref=":w"/>
    </ruleml:Atom>
    <ruleml:Atom>
      <ruleml:Rel iri="prOnto:Controller"/>
      <ruleml:Var key=":x">x</ruleml:Var>
      <ruleml:Var keyref=":z"/>
    </ruleml:Atom>
    <ruleml:Atom keyref=":A3">
      <ruleml:Rel iri="prOnto:PersonalDataProcessing"/>
      <ruleml:Var key=":ep">ep</ruleml:Var>
      <ruleml:Var keyref=":x"/>
      <ruleml:Var keyref=":z"/>
    </ruleml:Atom>
  </ruleml:And>
</ruleml:Exists>
```

4 Validation Methodology

From the experience in the DAPRECO project, it turned out that even IT experts required several and repeated explanations by the modeller to understand the formulæ. Hence the need for a *human-readable* representation of the formulæ, something which preserves the meaning of the machine-readable model while expressing that meaning clearly to non-expert in logic, ontologies, or XML. A few questions emerged: (i) What goal, beside usability, this human-readable representation should achieve? (ii) How to assess that the human-readable representation preserves the meaning of the formulaæ? (iii) How to assess that that

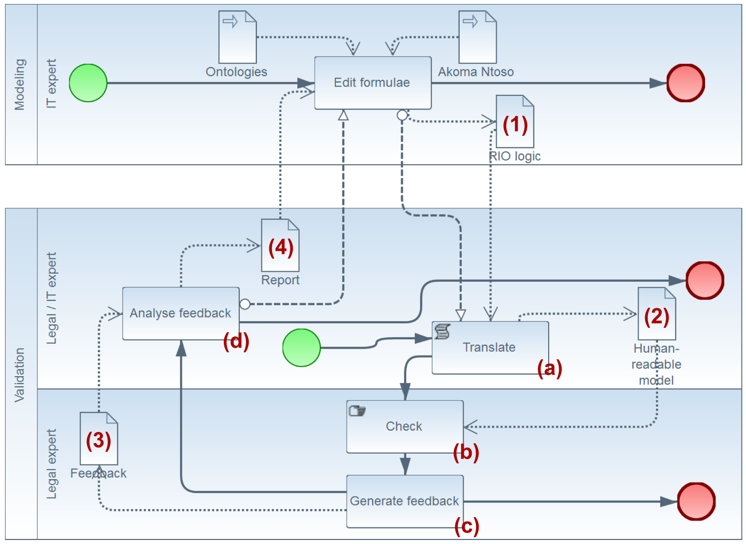

Fig. 1. Workflow of the Modelling and of the Validation methodology.

representation unambiguously and intelligibly conveys that meaning to those not familiar with the modelling?

It also became clear that, to answer these questions, the goal of the validation should be clarified; the human-readable form should be clearly understandable at least with respect to the feedbacks needed from legal evaluators, to verify that the model is done "in the legal right way".

Assuming to have already sorted out what features such a human-readable model must have, the methodology workflow for this work is shown in Figure 1.

The part about the methodology lays on the lower portion of the diagram ("Validation"). The machine-readable version of the modelling of the legal text —in our case, the DAPRECO Knowledge Base— is the output of the modelling effort by the IT expert. That file needs to be processed and rewritten ("Translated", (a) in Figure) into a human-readable representation. The "Human-readable model" (2) is then validated ("Check", (b)) against specific measures telling whether the modelling was correct from a legal point of view. The checking produces a list of "Feedbacks" (3) expressing the assessment of the model's legal qualities, likely in the form of quality measures or answers to a questionnaire. The feedbacks are then analyzed ("Analyze feedback", (d)), e.g., the statistical significance of certain answers will be measured, to compile a "Report" (4) for the IT experts and for the knowledge base builders. The report contains suggestions to review and improve their modelling. This workflow can be iterated until both parties are satisfied.

Due to space constraints, the present work will not delve into the details of each individual step, but only reports on the most critical step in the methodology: “Translate”. This step generates a human-readable version of the formal representation produced by the “Modelling” phase.

4.1 Translating into a human-readable model

The “Translate” step generates a representation of the formulæ that legal evaluators can use to assess their *legal quality*. Several metrics can be used to measure that quality, such as *completeness* (is all the required domain knowledge explicitly stated, or can it at least be inferred from the vocabulary?); *conciseness* (is there any amount of redundancy in the representation, or is it concise?); *accuracy* (do the deontic modalities represented in the formulæ match the corresponding legal provisions?); *consistency* (is the representation consistent with the law, or does it contain any contradictory knowledge?); and *clarity* (is the representation easy to understand?).

The human-readable representation fed to the legal evaluators to assess legal quality must be easily understandable: the evaluator’s mental strain to read the human-readable model to provide legal quality feedbacks should not overcome the effort required to provide feedbacks. A mental effort can be measured [3], but here we propose to measure the understandability of the human-readable representation indirectly as the interrater reliability[9] among a few raters judging whether and how easily they are able to answer a few questions for which they have to understand the meaning of the model.

4.2 Use Case: Translating LegalRuleML of RIO logic formulæ

Our input is the DAPRECO knowledge base, a LegalRuleML file of RIO formulæ supposedly expressing the legal meaning of articles of the GDPR. Reading the knowledge base presents some difficulties, although slightly facilitated by accompanying comments. For instance, in the LegalRuleML serialization, reading the enumerated prohibitions, obligations, reparations, exceptions is not straightforward. According to [15] “the list of [LegalRuleML] elements and their definitions are not sufficient for the consistent and accurate application of the annotations to text, nor is there clarification about how to analyse source text into LegalRuleML. Thus, an annotation methodology is required to connect text to LegalRuleML”.

Those issues have to be resolved within the specific validation (“Check”) that we foresee in our use case, consisting in the collection of feedback from the evaluators about the questions in Table 1.

To elicit a set of usability requirements for the human-readable model, we performed an internal unstructured inquiry where legal experts were asked to spell out what was making the reading hard and mentally burdensome while trying to answer the previous questions. The inquiry highlighted the following obstacles to a clear understanding of the LegalRuleML of a RIO formula: 1) a formula has little structure, and there are many variables and cross-references

[9] `https://en.wikipedia.org/wiki/Inter-rater_reliability`.

Table 1. The questions that we expect to ask for the "Check" step.

q1	Does the deontic modality of the formula (obligation, permission, constitutive rule) match that of the article?
q2	Does the formula capture all the relevant legal concepts that are expressed in the article (explicit concepts)?
q3	Does the formula capture all other relevant legal concepts not expressed in the article (implicit concepts)?
q4	Does the meaning of the formula completely represent the meaning of the article? What is missing?
q5	Does the meaning of the formula consistently represent the meaning of the article? What is wrong?

between them, forcing the reader to move up and down the code; 2) external references may refer to concepts expressed in the PrOnto ontology, or to logical operators from the RIO logic; 3) the choice of the names of predicates and arguments is not driven by a clear strategy, so that the formula can appear as confusing; 4) whether a formula is an obligation, a permission or an entailment is not immediately readable from its syntax, as it depends on the context, which is defined elsewhere according to LegalRuleML practices; 5) negations are hard to read, as they are structured with two predicates, the first introducing the negation of the second predicate that is expressed positively; 6) RIO logic avoids nesting of obligations and permissions, separating the content of the deontic rule from its bearer in two distinct formulæ. This decision, motivated by the purposes of the logic, can create some confusion, as ultimately there will generally be two separate, and almost identical, formulæ, with the same premises and almost the same consequence.

We address all these problems in a two-step "Translation": the first step is a software that parses the XML, expands and reorders the predicates of the formula; this addresses obstacles 1, 4, 5 and 6. The second is hand-made, to derive an almost natural language break-up version of the formula which, we believe, removes obstacles 2 and 3.

Step One: Automatic Parsing. The output of the automatic translator[10] overcomes the problems enumerated above in the following way: (**i**) variables are substituted with the predicate (taken from PrOnto) that restricts their type; (**ii**) predicates from PrOnto are clearly highlighted in **bold**, whereas predicates from RIO logic and terms that have been introduced for readability's sake are not; (**iii**) the translation of a predicate introduces some terms to try to put everything into context. This technique works quite well due to a good structure of the ontology; (**iv**) the context of a formula (obligation, permission, constitutive) is carried over to the translation; (**v**) negations are treated by translating the predicates in an inline negative sentence. Additionally, when a negation is the object of an obligation, the latter is renamed to a prohibition, and its content

[10] Available at `https://github.com/guerret/lu.uni.dapreco.parser.git`.

expressed positively; (**vi**) if the parser can find another formula with the exact same IF conditions, then they are most likely the content and bearer of an obligation or permission, so the two formulæ are merged into a single translation, which includes both content and bearer.

Article 7.1 of the GDPR can serve as an example: "*Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data*"[11].

The (simplified) RIO formula that IT experts wrote (and later encoded in LegalRuleML) to model the provision is shown in formula 2.

$$\begin{aligned}&([\,(\mathsf{RexistAtTime}\ a_1\ t_1) \wedge (\mathsf{and}\ a_1\ e_p\ e_{hc}\ e_{au}\ e_{dp}) \wedge (\mathsf{DataSubject}\ w) \wedge \\ &\quad (\mathsf{PersonalData}\ z\ w) \wedge (\mathsf{Controller}\ y\ z) \wedge (\mathsf{Processor}\ x) \wedge (\mathsf{nominates'}\ e_{dp}\ y\ x) \wedge \\ &\quad (\mathsf{PersonalDataProcessing'}\ e_p\ x\ z) \wedge (\mathsf{Purpose}\ e_{pu}) \wedge (\mathsf{isBasedOn}\ e_p\ e_{pu}) \wedge \\ &\quad (\mathsf{Consent}\ c) \wedge (\mathsf{GiveConsent'}\ e_{hc}\ w\ c) \wedge (\mathsf{AuthorizedBy'}\ e_{au}\ e_{pu}\ c)\,] \rightarrow \\ &\quad [\,(\mathsf{RexistAtTime}\ e_a\ t_1) \wedge (\mathsf{AbleTo'}\ e_a\ y\ e_d) \wedge (\mathsf{Demonstrate'}\ e_d\ y\ e_{hc})\,]\,) \in O\end{aligned} \tag{2}$$

The parser translates the formula as follows:

IF, in at least a situation,

- At time $:t_1$, the following situation exists:
  - (All of the following ($:a_1$))
    1. **Processor** ($:x$) does **PersonalDataProcessing** ($:e_p$) of **PersonalData** ($:z$)
    2. **DataSubject** ($:w$) performs a **GiveConsent** ($:e_{hc}$) action on **Consent** ($:c$)
    3. **Purpose** ($:e_{pu}$) is **AuthorizedBy** ($:e_{au}$) **Consent** ($:c$)
    4. **Controller** ($:y$) **nominates** ($:e_{dp}$) **Processor** ($:x$)
- **PersonalData** ($:z$) is relating to **DataSubject** ($:w$)
- The **Controller** ($:y$) is controlling **PersonalData** ($:z$)
- **PersonalDataProcessing** ($:e_p$) **isBasedOn Purpose** ($:e_{pu}$)

THEN it must happen that, in at least a situation,

- At time $:t_1$, **Controller** ($:y$) is **Obliged** to **AbleTo** ($:e_a$)
- **Controller** ($:y$) **Demonstrate** ($:e_d$) **GiveConsent** ($:e_{hc}$)

Although the translation still requires some mental effort to be processed, it is at least understandable without expertise in logic. The automatic processing also allowed the modeller to verify that the intended meaning has not been changed and is preserved in the translation.

Step Two: Hand Made Break-up. The automatic translation has been further hand-processed. The output is a natural language break-up that highlights the following elements: Premises and the Conclusion of the formula; the Deontic Modality, the Ontological Concepts that can be recognized in the article, Other Ontological Concepts present in the formula but not mentioned in the article; the Contextual meaning, which is what the formula expresses but is not in the article, and the Overall Meaning of the formula. The break-up of Article 7.1 is shown in Table 2.

[11] The full translations for Articles 5.1 and 7.1 can be found in the repository from note 10, in the "jurisin" folder.

Table 2. Structure of the formula's meaning.

Premise	Where processing is based on consent,
Conclusion	the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data.
Modality	Obligation
Ont. Concepts	Where [Processing] is based on [Consent], the [Controller] shall be [Able to] [Demonstrate] that the [Data subject] [Has consented = GiveConsent] to [Processing] of his or her [Personal data]
Other Ont. Concepts	[Purpose]; [Processor]; [IsAuthorizedBy]; [Nominates]; [IsBasedOn]; [BeAbleTo]
Context	There is a processing, which has a purpose authorized by a consent given by a data subject, and that is what a processor, whom a controller controlling the personal data nominates, does on personal data of the data of the data subject.
Overall Meaning	**Whenever** there is a processing, which has a purpose authorized by a consent given by a data subject, and that is what a processor, whom a controller controlling the personal data nominates, does on personal data of the data of the data subject **then** the controller is **obliged** to able to demonstrate that "data subject gave consent".

4.3 Measuring the usability of the human-readable model

Before collecting the legal experts' feedbacks on the quality of the model, the human-readable model must be read consistently and correctly by the evaluators. The experiment consisted in letting the four legal evaluators (two with knowledge of deontic logic, two without it) answer a few yes/no questions about their understanding of the models of two GDPR provisions, Articles 5.1(a) and 7.1. The input is the human-readable model, but we also fed the original XML formalization and the pre-processed output as control cases, measuring the (pure, not Fleiss Kappa) average interrater agreement between the answers of the evaluators for each model. The questions, built in the wake of the ones used for the validation check in Table 1, were the following. 1. *Can you identify the formula's premise?* 2. *Can you identify the formula's conclusion(s)*? 3. *Can you identify the deontic modality (obligation, permission, other)*? 4. *Can you identify the formula's explicit ontological concepts?* 5. *Can you identify the formula's implicit ontological concepts?* 6. *Do you understand what the formula means?* 7. *Try to rewrite the formula in your own words. Did you succeed?*

We collected the average measure over the two formulæ. The results are shown in Table 3. The hand-processed model is where the evaluators agree almost unanimously over answering 'yes' to all questions, thus indicating high understandability; the control item, the XML file, is where instead there is a majority consensus on being not understandable. Our result also reflects that validators already knowledgeable of logic can somehow read the XML files, despite not fully; unsurprisingly, non-experts could not make any sense of it. Conversely, there is no consensus on the understandability of the automatically-processed model. Supposedly, better usability scores may be attained by training the legal evaluators, but we have not explored this possibility.

Threats to validity We measured understandability as the interrater agreement among a few testers: this measure can suffice to the present goal of having the human-readable model as a candidate within the methodology, but additional

Table 3. Output of the agreement (on 'yes' on 'no') on the readability experiment.

	Commented XML	Intermediate	Human-readable
all	60.6% (no)	40.9%(yes)/59.1%(no)	97.7%(yes)
non-experts	100% (no)	45.5%(yes)/54.5%(no)	95.5%(yes)

measures can provide a deeper evaluation of its usability. More evidence would be needed to assert that our hand-processed model is readable, but since our evaluators generally agree on its understandability, it can already be used to collect the answers that legal experts will give to questions in Table 1 and use them reliably. This will be a next step in the methodology, together with the analysis of the feedback collected during this research.

We stress that, yet, we did not apply the methodology fully: although we collected feedbacks on the legal quality of the formalization of the two GDPR's articles, their analysis is left as future work. Still, the research herein, on the understandability of the human-readable model is a necessary step to later perform the validation methodology.

5 Conclusions and future work

This paper proposed a methodology to perform an interdisciplinary validation of formal representation of a legal text, herein a RIO logic formalization of the GDPR Articles.

Although grounded in a domain-related implementation (the data protection domain), the methodology yields a more general spectrum, since any legal representation calls for a legal validation phase. To the best of the authors' knowledge, a general methodology to evaluate the quality and soundness of a legal representation is novel.

In our case, the generation of a human-readable model was conceived in two steps: an automated translation of the formulæ into a language that unwraps the underlying logic without changing its structure; and a (human-made) post-processing to highlight certain features of the rules that were still hard to detect by legal experts. The final result was then presented to a small group of testers, all from law, who provided feedback.

This work has outlined several interesting issues. First and foremost, inasmuch as human-readable an automated translation can be, it strongly depends on the degree of complexity of the deployed logic. In the use case presented here, for example, the automated translation was generally readable, but some particularly complex formulæ required an added debriefing session, normally not born by pure legal expert; hence the need for a post-processing. The feedback alternated between the correctness of the formulæ and the quality of the translation. The computer expert is therefore required to understand whether a certain feedback requires an improvement of the formula, or if a refined translation (whether a pre- or post-processing) could render a better understanding by the legal expert. The feedback was delivered through a set of answers to a built questionnaire. The answers from a group of legal experts were combined to discern problems in the translation from those in the formula, or whether there was an interpretative issue (not an error in itself) which could be rather used to improve the knowledge base.

Although confined in its early stages, this work is meant to open a new direction of research in the domain of formal modelling of legal texts, and envisions several follow-ups.

First, the methodology opens the possibility for a thorough validation of the DAPRECO knowledge base, thus improving not only the logic formulæ, but also refining the PrOnto ontology, in a recursive validation process.

Secondly, the results of this approach can also be used to find a clearer and more systematic way to write RIO logic formulæ in the LegalRuleML serialization. In fact, our automated processing highlighted several problematic issues, calling for a process of translation (e.g., in serializing the concepts, in choosing the names of the functions, in adding or omitting variables) that do not depend only on the sheer subjective experience of the modeller. Such a revision will possibly allow achieving better results in the automated translation. If the methodology is streamlined, it might be possible to eliminate the need for manual post-processing, thus leading to a more objective and verifiable human-readable model creation which is scalable to the whole DAPRECO Knowledge Base. Furthermore, a streamlined methodology would allow to embed an interdisciplinary validation process already at the very stage of the modelling of the provisions. This means that while legal provisions are represented (by the computer scientist) into a set of logic formulæ, they would be at one presented to the legal expert to provide feedback. Timely feedback during the creation of the knowledge base could empower the computer scientist to use such expert assessment while modelling subsequent provisions. Surely such a solution would work at its best when the generation of the model from the legal text is made in a semi-automated way, i.e., by means of Natural Language Processing (NLP) techniques, whereby the feedback from the legal expert could be used to improve the processing engine and immediately generate a refined version of the knowledge base. Such an approach can of course be extended to cover different models and domains of modelling, e.g., terms and conditions and privacy polices.

Thirdly, there is a need to define, together with the legal experts, a set of metrics that can express a validator's acceptable assessment on the legal quality of the formalization. In Subsection 4.1, we anticipated a few of possible metrics (i.e., conceptual completeness, conciseness, accuracy, consistency and clarity), but those were not thoroughly investigated yet. This may lead to a revision of the current human-readable model.

References

1. Athan, T., Governatori, G., Palmirani, M., Paschke, A., Wyner, A.: Legalruleml. In: Faber, W., Paschke, A. (eds.) Reasoning Web, Information Systems and Applications, incl. Internet/Web, and HCI, vol. 9203, pp. 151–188. Springer (2015)

2. Bartolini, C., Giurgiu, A., Lenzini, G., Robaldo, L.: Towards legal compliance by correlating standards and laws with a semi-automated methodology. In: BNAIC 2016: Modern Trends in Artificial Intelligence, Communications in Computer and Information Science, vol. 765, pp. 47–62. Springer International Publishing (2017)
3. Benenson, Z., Lenzini, G., Oliveira, D., Parkin, S., Uebelacker, S.: Maybe Poor Johnny Really Cannot Encrypt: The Case for a Complexity Theory for Usable Security. In: Proc. of the 2015 New Security Paradigms Workshop, NSPW 2015, Twente, The Netherlands, September 8-11, 2015. pp. 85–99 (2015)
4. Boella, G., Janssen, M., Hulstijn, J., Humphreys, L., van der Torre, L.: Managing legal interpretation in regulatory compliance. In: Verheij, B., Francesconi, E., Gardner, A. (eds.) Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law (ICAIL). pp. 23–32. ACM (June 2013)
5. Casanovas, P., Palmirani, M., Peroni, S., van Engers, T., Vitali, F.: Semantic web for the legal domain. Semantic Web 7(3), 213–227 (March 2016)
6. Casellas, N.: Ontology evaluation through usability measures. In: On the Move to Meaningful Internet Systems: OTM 2009 Workshops, Lecture Notes in Computer Science, vol. 5872, pp. 594–603. Springer (2009)
7. Francesconi, E.: A learning approach for knowledge acquisition in the legal domain. Approaches to Legal Ontologies, Law, Governance and Technology Series 1, 219–233 (2011)
8. Lévy, F., Nazarenko, A.: Formalization of natural language regulations through SBVR structured english. In: Morgenstern, L., Stefaneas, P., Lévy, F., Wyner, A., Paschke, A. (eds.) Theory, Practice, and Applications of Rules on the Web, Lecture Notes in Computer Science, vol. 8035, pp. 19–33. Springer (2013)
9. Mockus, M., Palmirani, M.: Legal ontology for open government data mashups. In: Parycek, P., Edelmann, N. (eds.) Proceedings of the 7^{th} International Conference for E-Democracy and Open Government (CeDEM). pp. 113–124. IEEE Computer Society (May 2017)
10. Palmirani, M., Cervone, L., Bujor, O., Chiappetta, M.: RAWE: A web editor for rule markup in LegalRuleML. In: CEUR Workshop Proceedings (2013)
11. Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., Robaldo, L.: PrOnto: Privacy ontology for legal compliance. In: Proceedings of the 18^{th} European Conference on Digital Government (ECDG) (October 2018), upcoming.
12. Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., Robaldo, L.: PrOnto: Privacy ontology for legal reasoning. In: Kö, A., Francesconi, E. (eds.) Electronic Government and the Information Systems Perspective, Information Systems and Applications, incl. Internet/Web, and HCI, vol. 11032, pp. 139–152. Springer (2018)
13. Ramakrishna, S., Górski, Ł., Paschke, A.: A dialogue between a lawyer and computer scientist. Applied Artificial Intelligence 30(3), 216–232 (2016)
14. Robaldo, L., Sun, X.: Reified Input/Output logic: Combining Input/Output logic and reification to represent norms coming from existing legislation. Journal of Logic and Computation 27(8), 2471–2503 (December 2017)
15. Wyner, A., Gough, F., Levy, F., Lynch, M., Nazarenko, A.: On annotation of the textual contents of scottish legal instruments. In: Proc. of the 30^{th} Int. Conf. on Legal Knowledge and Information Systems (JURIX). vol. 302, pp. 101–106. IOS Press (December 2017)