



# Derogations Analysis of European Legislation Through Hybrid AI Approach

Monica Palmirani<sup>(✉)</sup> and Davide Liga

CIRSFID-ALMA AI, University of Bologna, Bologna, Italy  
{monica.palmirani, davide.liga2}@unibo.it

**Abstract.** This paper presents a use case of hybrid AI approach applied to the European legislation with the aim to detect the derogations in the norms and to extract the main components. The result is modelled in Akoma Ntoso XML standard for supporting further applications, open data sharing and interoperability between different tools. We have conducted this research inside of the project ‘Drafting legislation in the era of AI and digitisation’ with the support of the EU Commission – Directorate General Informatics Unit B2 – Solutions for Legislation, Policy & HR.

**Keywords:** Hybrid AI · Akoma ntoso · Visualization

## 1 Drafting Legislation in the Era of Artificial Intelligence and Digitisation

The European Commission is recently providing a roadmap for the *digital-ready legislation*<sup>1</sup> with an interdisciplinary approach and it is investigating the “Drafting legislation in the era of artificial intelligence and digitisation” (workshop 2019)<sup>2</sup>. EU Commission, Directorate-General for Informatics is performing with the University of Bologna a study on “Drafting legislation in the era of artificial intelligence and digitisation” that includes three pilot-cases using AI techniques applied to support the legal drafting units. In this study we propose a third way (e.g., *Hybrid AI for Law*) with a legal and technical model for developing computable informatics legal systems compliant by-design (or *Legal Protection by-design* as Hildebrandt defined) with theory of law, integrating also Semantic Web approach and LegalXML annotation (Filtz 2021, Robaldo 2019).

## 2 Hybrid AI Approach and Methodology

We propose to use the so-called *Hybrid AI* where *human-in-the-loop*, *human-on-the-loop*, and *human-in-command* principles<sup>3</sup> are combined with different complementary

<sup>1</sup> <https://joinup.ec.europa.eu/collection/better-legislation-smoother-implementation/digital-ready-policy-making>.

<sup>2</sup> <https://ial-online.org/wp-content/uploads/2019/07/Invitation-EN.pdf>.

<sup>3</sup> High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence, 2019.

disciplines (law, philosophy, ethics), using symbolic and sub-symbolic AI techniques integrated with Semantic Web findings in order to add context and meanings to the pure *data-driven* or *code-driven* methodology. The *Hybrid AI* is very promising approach especially in legal domain where the context, values, concepts are fundamental for correctly apply the AI outcomes (Rodríguez-Doncel 2021, Fratrici 2021, Verheij 2020, Ashely 2017). Another important element that is fundamental for guaranteeing the legitimacy of the whole *digital law-making process* is the metadata concerning the workflow. Secondary we know about some specific critical technical issues that characterise the legislative domain when the AI non-symbolic alone is used: i) normative references knowledge and connected destination text are often not investigated, ii) article-level approach is fundamental respect the document-oriented method, iii) the temporal analysis of the documents permits to manage the point-in-time, iv) legal language peculiarities must be considered respect a sole frequency approach (e.g., TF-IDF term frequency-inverse document frequency method), v) contextual and semantic enrichment in legal domain is crucial (e.g., jurisdiction). For this reason, we use LegalXML Akoma Ntoso standard (now abbreviated in AKN) as background format of the legal document and on the top of this we add other AI solutions. In the light of the previous considerations, we have adopted a hybrid methodology mixing different techniques: unsupervised for discovering new knowledge to offer to the legal experts during the supervised phase. Additionally, we have also used symbolic and non-symbolic AI techniques. Finally, Akoma Ntoso is used as skeleton for creating the context and the semantic annotation. We have proceeded as following (Table 1).

**Table 1.** Methodology of hybrid AI.

A. Legal analysis	1. Legal analysis using Eur-Lex
B. Preparation of dataset	2. Selection of the document using Sparql end point of CELLAR <sup>a</sup>
	3. Use of metadata of ELI/CDM/Eurovoc (the ontologies defined at European level for managing legal metadata) from CELLAR
	4. Raw conversion of the documents in AKN
	5. Legal analysis of random sample of partitions and metadata
	6. Definition of the taxonomy of derogations
C. Exploration of the knowledge	7. Extraction of the partitions (article level)

(continued)

**Table 1.** (continued)

	8. Using KNIME <sup>b</sup> + Python <sup>c</sup> for exploring new knowledge
D. Experiment	9. Annotation of the fragments
	10. NLP analysis of the derogations (RegEx <sup>d</sup> , tree banks) and AI (tree kernel) for extracting the component of the derogations (e.g., destination, jurisdiction, scope, temporal elements, etc.)
	11. AKN conversion of the legal knowledge extracted
E. Evaluation	12. Visualization
	13. Validation
	14. Interpretation

<sup>a</sup>CELLAR: Publications Office, *Cellar: the semantic repository of the Publications Office*, Publications Office, 2018, <https://data.europa.eu/doi/10.2830/028321>.

<sup>b</sup><https://www.knime.com/>

<sup>c</sup><https://www.python.org/>

<sup>d</sup>See the standard specification of the language RegEx: ISO/IEC/IEEE 9945:2009 *Information technology – Portable Operating System Interface (POSIX®) Base Specifications, Issue 7*.

### 3 Derogations

A derogation is a legislative tool that makes it possible to create particular subcases starting from a basic obligation, permission, or right. The derogation is frequently connected with the action of ‘disapplication’ (e.g., “disapplication provision”<sup>4</sup>), limited to a specific interval of time, or related to some special categories of addresses regulated with the specifications of conditions. As we know from legal theory, this instrument is very relevant in preserving rule-of-law principles all the while making for flexibility in some circumstances like the COVID-19 pandemic. However, derogations are difficult to retrieve in the text and across the common legal databases. It is more difficult to track modifications over time. Legal experts and the legal drafters therefor struggle to follow the chain of derogations and thus obtain clear and transparent legal information. Here is an example of a modification of a derogation introduced in response to COVID-19:

---

#### Article 1

Regulation (EU) No 223/2014 is amended as follows:

(omissis) in Article 13(1), the following subparagraph is added:<sup>a</sup> ‘By way of derogation from the first subparagraph, the deadline for the submission of the annual implementation report for the year 2019 shall be 30 September 2020.’;

---

<sup>a</sup>Art. 1 of Regulation EU 2020/559, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32020R0559&from=EN>. Visited in June 2022.

---

<sup>4</sup> See here an example: <https://www.legislation.gov.uk/ukpga/2008/7/section/6/enacted>.

### 3.1 Preliminary Taxonomy of Derogations

Using EUR-Lex<sup>5</sup>, a legal expert defines 15 categories of derogations along four axes: i) Frequent linguistic formulations; ii) Temporal parameters; iii) Relationships between EU legislation and Member States; iv) Relationships between primary legislation and delegated acts. Then the legal informatics has extracted the fragment of the documents and the legal experts have grouped them in the following categories using the legal meaning:

- Frequent linguistic formulations:

---

by way of derogation from/to <partition>

---

without prejudice to the obligations defined by <partition>

---

derogation applied in accordance to/with

---

derogation from/to the <partition> referred to in <partition>

---

derogation from

---

derogation to application

---

derogation applied in accordance with

---

- Derogation and internal times

Derogation that includes temporal parameters that are part of the deontic norms:

---

By way of derogation from the first subparagraph, the deadline for the submission of the annual implementation report for the *year 2019 shall be 30 September 2020*

---

- Derogation and external times

Derogation that are limited by temporal parameters in the applications, enter into operation, or enter into force:

---

Upon Commission's initiative or in response to a request from a beneficiary country, a beneficiary country may be granted a *temporary derogation* from the provisions of this section where:

---

By way of derogation from the date of application referred to in the second paragraph of Article 61, Article 46 shall apply from **17 June 2018** insofar as necessary in order to allow a timely recognition of control authorities and control bodies.<sup>a</sup>

---

<sup>a</sup><https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018R0848&from=EN> Visited in June 2022.

---

<sup>5</sup> EUR-Lex "is your online gateway to EU Law. It provides the official and most comprehensive access to EU legal documents. It is available in all of the EU's 24 official languages and is updated daily" <https://eur-lex.europa.eu/content/welcome/about.html?locale=en>. Visited in June 2022.

- Derogation and jurisdiction

Derogation that are limited in the application to some countries:

---

By way of derogation from paragraphs 1 and 2, *in Cyprus, Croatia, Malta and Slovenia*, the amount referred to in those paragraphs may be set at a value lower than EUR 500, but not less than EUR 200 or, in the case of Malta, not less than EUR 50

---

- Exception

Exception is a special derogation where the subcase is strongly distinguished from the regular norm:

---

In any case, all operators and groups of operators, with the *exception* of those referred to in Articles 34(2) and 35(8), shall be subject to a verification of compliance at least once a year

---

- Shall not apply

‘Shall not apply’ is used as negative formulation:

---

Article 19 of this Directive *shall not apply* where a bundle within the meaning of Directive (EU) 2018/1972 includes elements of an internet access service as defined in point (2) of Article 2 of Regulation (EU) 2015/2120 of the European Parliament and of the Council

---

- Member States delegation of derogate

This formulation is used for delegate to the Member State the regulation of some subcases:

---

*Member States* shall regularly review derogations under this paragraph taking into account good practices in separate collection of waste and other developments in waste management

---

- Request of derogation from Member States

This formulation is used when the Member State requests more room for applying a derogation:

---

*At the request of a Member State*, the Commission may allow a derogation from the prohibition set out in Article 13(1) of Regulation (EC) No 1967/2006, provided that a number of conditions set out in Article 13(5) and(9) are fulfilled

---

- Delegated Acts

This formulation is used for adopting delegated Acts. This is a derogation using different level of source of law:

---

*The Commission shall adopt delegated acts* in accordance with Article 264 concerning the special rules referred to in paragraph 1 of this Article regarding derogations from the requirements provided for in Article 229(1) and Articles 233 and 237 and imposing additional requirements for the entry into the Union of the following:

---

Finally, we frequently find the so-called *reflexive derogation* that act internally to the same act often correlated to temporal period of efficacy.

---

By way of derogation from **point (a)**, from *1 January 2021 to 31 December 2021* the obligation to decommission the unique identifier of medicinal products which the wholesaler intends to distribute outside of the Union shall not apply to products which he intends to distribute in the United Kingdom

---

The *external derogation* has the destination another legal document:

---

2. By way of derogation from **Article 13(6) of Regulation (EC) No 1235/2008**, at the verification of a consignment the relevant Member State's competent authority shall endorse the certificate of inspection by validating box 20 in TRACES<sup>a</sup>

---

<sup>a</sup><https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020R0977>.

### 3.2 Dataset of Derogations

The dataset is made up of legislative acts in the span of time from 2010 to 2020 for a total of 15,328 documents. All the documents were provided in Formax<sup>6</sup> format by the European Publication Office with also the related metadata (e.g., ELI<sup>7</sup>/CDM<sup>8</sup>/Eurovoc<sup>9</sup>). Each document is a package with a main document and possible further attachments and annexes. The documents are converted into Akoma Ntoso and using the taxonomy we came up with 13,587 partitions involved in the derogation, using a preliminary ‘indicator’ taxonomy for extracting the text involved in the experiment. The goal should detect several important elements in the text: destination of the derogation, action proposed, jurisdiction, temporal elements, conditions, quantities. In Akoma Ntoso we can model directly the text of the derogation and integrate the semantic role of each element in the metadata specifically designed for hosting the information concerning the modifications, including derogation.

### 3.3 Akoma Ntoso Conversion

During the conversion in AKN from Formex we detected the part of the wording involved in the derogation: citations of the main obligation that is derogated from, temporal parameters, and the particular scope being derogated from. Additionally, we have modelled the derogation in Akoma Ntoso metadata in order to reuse them for further statistical elaborations.

---

<sup>6</sup> Formex is the acronym of “Formalized Exchange of Electronic Publications”. “Formex describes the format for the exchange of data between the Publication Office and its contractors. In particular, it defines the logical markup for documents which are published in the different series of the Official Journal of the European Union. Formex V4 is based on the international standard XML (Extensible Markup Language—W3C Recommendation, February 10, 1998). It entered into force on May 1st, 2004.” <https://op.europa.eu/it/web/eu-vocabularies/formex>.

<sup>7</sup> ELI is the acronym of “European Legislation Identifier” that defines Unique Naming convention for legislation, common data model of metadata, interoperable mechanism (API) for query the information between the member states. It has been introduced with the “Council conclusions of 6 November 2017 on the European Legislation Identifier” in all the member states and in the European institutions. The ELI patterns for the European legislation is defined here: [https://eur-lex.europa.eu/eli-register/eu\\_publications\\_office.html](https://eur-lex.europa.eu/eli-register/eu_publications_office.html), <https://op.europa.eu/it/web/eu-vocabularies/eli>. Visited in June 2022.

<sup>8</sup> CDM is the acronym of “Common data model” that is used by the European Publication Office for creating a sharable ontology: <https://op.europa.eu/it/web/eu-vocabularies/cdm>.

<sup>9</sup> Eurovoc “is the EU’s multilingual and multidisciplinary thesaurus. It contains keywords, organized in 21 domains and 127 sub-domains, which are used to describe the content of documents in EUR-Lex.” <https://eur-lex.europa.eu/browse/eurovoc.html?locale=en>. Visited in June 2022.

Formex	Akoma Ntoso
<p>&lt;ALINEA&gt;By way of derogation from the second paragraph, Member States may choose not to apply the provisions of point ORO.FTL.205(e) of Annex III to Regulation (EU) No 965/2012 and continue to apply the existing national provisions concerning in-flight rest until &lt;DATE ISO="20170217"&gt;17 February 2017&lt;/DATE&gt;.&lt;/ALINEA&gt;</p>	<pre>&lt;alineia eId="body_art_2_al_3"&gt;   &lt;content eId="body_art_2_al_3_content"&gt;     &lt;mod       eId="body_art_2_al_3_content_mod_1"&gt;       &lt;p&gt;By way of derogation from the second         paragraph, Member States may choose not to apply         the provisions of &lt;ref eId="ref_1"           href="href="/akn/eu/act/regulation/2012-02-             17/965-2012!/main/&gt;annex_III"&gt; point           ORO.FTL.205(e) of Annex III to Regulation (EU)           No 965/2012 &lt;/ref&gt; and continue to apply the           existing national provisions concerning in-flight           rest until&lt;date date="2017-02-17"             refersTo="#derogationTime"&gt;17 February             2017&lt;/date&gt;.&lt;/p&gt;         &lt;/mod&gt;       &lt;/content&gt;     &lt;/alineia&gt;</pre>

We first define all elements that in the sentence relate to derogation action. Thus, we extract the following element that find an adequate representation in:

- destination of the derogation is detected in the text using <ref href > (normative references connected with the derogation in order to produce a graph of all the derogations and the relative norms). This information is reported also in the metadata in the tag <destination> inside of the modification information <scopeMod>;
- conditions (e.g., only for the bank, only for the COVID-19 pandemic situation) are detected in the text and then reported in the metadata tag <condition>;
- jurisdiction (e.g., only for Denmark) is detected in the text using <location> even if it is not really a physical place but the concept of the state sovereignty and the consequent jurisdiction;
- temporal parameters (e.g., for six months) are detected in the text and stored in different tags of the metadata according to the role played by this information: <duration> for specifying how long the derogation acts, <force> for recording the enter into force time, <efficacy> for recording the time of enter into operation;
- limitation in the application of some specific domain (e.g., only for tax law) we can store this information in the <domain> tag.

We then model everything in AKN in order to fix the knowledge and to reuse it for the search engine, the semantic web filter, or other sophisticated application.



```

<scopeMod type="exceptionOfScope">
<source href="body_art_2_al_3_content_mod_1"/>
<destination href="/akn/eu/act/regulation/2012-02-17/965-2012!/main/annex III"/>
<force>
<date date="2014-02-20"/>
</force>
<duration>
<date date="2017-02-17" refersTo="#endDate"/>
</duration>
<condition> Member States </condition>
<domain> continue to apply the existing national provisions concerning in-flight
</domain>
</scopeMod>

```

### 3.4 Technical Aspects

To retrieve the relevant part of the text we combined the use of regular expressions together with the powerful SpaCy library, in Python. More precisely, a software has been created which is capable of reading Akoma Ntoso structures by using two python libraries dedicated to the navigation of xml files (i.e. lxml and xml.etree), searching for all the elements of the body of the legal document to match all those elements containing the most frequent linguistic formulas by which derogations appears in the text. This first part of the process employed regular expressions and aims at extracting the sentences where derogations are found. After the navigation of the Akoma Ntoso files and the matching process of the derogation linguistic formulas, a SpaCy<sup>10</sup> pipeline has been feed with all retrieved sentences. This pipeline (which is currently under development) is built to detect as much information about the derogations as possible (e.g. time, places, conditions). After having found derogations, and having extracted the relative information, the software returns an Akoma Ntoso file which is marked up with the information about derogations.

Regarding the first phase, the matching process performed by using regular expressions produces a list of matches for the content of each element found in the xml structure of the body of the Akoma Ntoso files. It is important to note that this first part has been designed to deal with the physiological fragmentation of the content (fragmentation which can be more or less relevant depending on the verbosity of the information within the xml structure). In fact, those xml tags which can contain textual content sometimes have inner inline elements which can add a further layer of complexity during the extraction of the sentences. In this part, we made sure that all elements' texts and tails were considered, in order to reconstruct the content of the sentences as much as possible. This choice of design makes it possible to work also with irregular structures, or with structures which might possibly be not well-formed, or whose internal text is fragmented by inner inline tags. In this way, we tried to feed the regular expressions with sentences that are as much complete as possible.

<sup>10</sup> SpaCy is an open-source library of Natural Language Processing techniques capable to manage the task called Named Entity Recognition that permits to detect in the text persons, organizations, roles, dates, locations, concepts. For applying SpaCy to legal language domain it is necessary a customizaiton. <https://spacy.io/>.

In the second phase, the sentence containing the retrieved provision is transferred into a SpaCy pipeline in order to find as much information about the derogation as possible. Also in this case, having a sentence as much complete as possible is crucial to achieve better results. In fact, the steps of the pipeline are performed by navigating the dependency tree, the part of speech, and the tags of the sentences. The pipeline, which is currently under development, is partially completed, and has been designed as follows:

- 1 Searching for normative references (i.e. references to legal documents, such as “Regulation 96/2016”) and add them as “act\_reference” entities;
- 2 Searching structures of legal documents (i.e. references to portion of legal documents, such as “Second paragraph of Article 3”), adding them as “part\_reference” entities, and connecting them to any related normative reference possibly found in the first step;
- 3 Searching for EU member states and adding them as entity;
- 4 Searching for time references, and adding them as temporal entities;
- 5 Adding as entities the derogation formulas (which, as said before, are found using regular expressions, an example of derogation formula is “by way of derogation from”);
- 6 Adding deontic operators such as “should” (which can be crucial to determine the presence of other elements related to the derogation within the same sentence, and which can be used to take into consideration the deontic and argumentative dimension of the derogation);
- 7 Connecting the derogation formula (found in step 5) to its corresponding “destination”: this means connecting the derogation formula to the relative part\_reference (including the act\_reference, if any);
- 8 Finding any delegation and adding this information into the SpaCy document (i.e. finding all those linguistic formulas which imply a delegation of power within the derogation);
- 9 Find conditions (i.e. finding as many conditions as possible, related to the derogation) by navigating specific linguistic indicators (such as “only if”, “unless”, and so on).

In other words, all the provision extracted in the first phase are then passed to this SpaCy pipeline to search for the information about the derogation itself. The result of the SpaCy pipeline is a complex object which contains all the information we need to finally markup the original Akoma Ntoso with the new information about the derogation (including all the metadata which have been found). Although the pipeline is still under development, we managed to extract hundreds of derogations and produce marked up Akoma Ntoso accordingly.

### 3.5 The Derogation Graph

Starting from the Akoma Ntoso files marked up with annotations (as described before), we created a graph to represent and visualize derogations using d3js, a powerful library which is capable of generating interactive visualizations. Graphs are generally composed of two elements (van Kuppevelt 2020, Coupette 2021, Chalkidis 2021): nodes

and links connecting nodes. It is not the first time that the scholars create graphs for representing the legal knowledge, but concerning the legislation there are few experiments that stressed three main problems: i) the computational issues because the calculation of the network and the relationships between nodes need a specific algorithm capable to work with hundreds nodes (Bommarito 2019, Katz 2014); ii) the visual model for managing the complexity of the graph with large numbers of nodes (Koniaris 2021); iii) the transparency in the navigation that could be useful for the legal experts and in meantime understandable and auditable with valuable legal metadata (Alschner 2020; de Jong 2022).

For coping with these three main issues, we extracted a graph where each node is a legal act (which stores the information of the act itself, and where the unique ID is the FRBR Akoma Ntoso URI) and where each link is a relationship of derogation which can be established either with other nodes or with the same node (for reflexive relations). We used the AKN annotation and specifically the metadata modification `<scopeMod>` elements. The destination of the derogation is used for creating a connection between derogation and document derogated. The other elements (which are not necessarily present) are considered as information about the single connection, so their information is stored within the relative link. More precisely, “duration” provides information about the duration of the derogation, “place” provides information about the geographical zones where the derogation is valid, “domain” provides information about the domains of application of the derogation, “condition” provides information about the conditions for which the derogation is valid.

For each node of the graph (i.e., for each legal document) we also collected its relative Eurovoc classifications, which can be found within the AKN tag `<classification>`, and more precisely in the “value” attribute of any tag `<keyword>` whose “dictionary” attribute is equal to “eurovoc”. In the image above, we stored seven pieces of information for each node. The *unique AKN ID*, the *CELEX number*<sup>11</sup>, the *title* of legal act, and all the times in which the single node derogates others or itself (*derogator\_times* and *selfderogated\_times*, respectively), or is derogated by others (*derogated\_times*). Finally, we stored the Eurovoc related to each legal document. Regarding the links, they can store not only the *source* and *destination*, but also all the information that we extracted from the files as explained before, namely: *places*, *conditions*, *domain*, *duration*. Also, it is important to underline that the information of single nodes are displayed whenever the mouse is over them, as described in the following image:

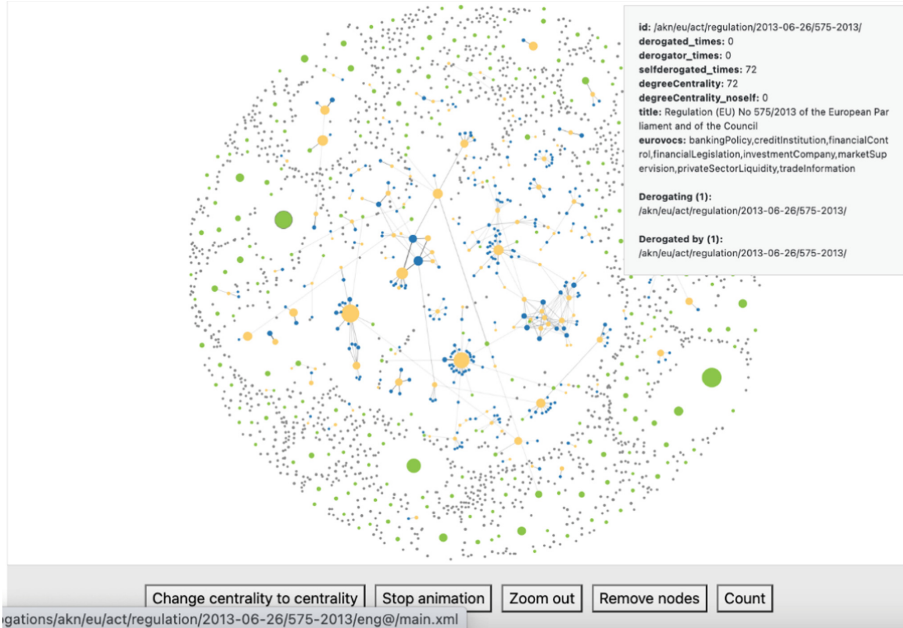
Importantly, each node has been assigned a *color* and a *weight* depending on 3 variables:

- 1) how many times the node is derogating other nodes (not considering itself);
- 2) how many times the node is derogating itself;
- 3) how many times the node is derogated by other nodes.

<sup>11</sup> *CELEX number* “is a unique identifier assigned to a document. It is independent of the language of the document.” It is the unique identifier adopted before the ELI and AKN naming convention and it is a legacy number for detecting legislative document in CELLAR independently to the various generations of the technology evolution. See <https://eur-lex.europa.eu/content/help/eur-lex-content/celex-number.html>.

The weight, which determines the dimension of the node, is achieved by following formula:

$$w = k * (d + r + t)$$



**Fig. 1.** Network analysis of the derogations.

where  $k$  is a fixed coefficient intended to produce the most appropriate graphical result,  $d$  is the times in which the node is derogating other nodes (excluding itself),  $r$  is the times in which the node is self-derogated, and  $t$  is the times in which the node is derogated (by other nodes or by itself). We call  $d + r + t$  “centrality”.

The color is decided as follows: *yellow* nodes are nodes which are mainly derogated by other nodes ( $t$  is greater than  $r$  and  $d$ ), *green* nodes are nodes which are mainly derogating themselves by other nodes ( $r$  is greater than  $t$  and  $d$ ), *blue* nodes are nodes which are mainly derogating other nodes ( $d$  is greater than  $t$  and  $r$ ). Finally, *grey* nodes are nodes which are not derogating nor being derogated ( $d = t = 0$ ). To make the graph less verbose, we included in the grey nodes only those whose  $r$  is greater than 4 (i.e., they derogated themselves more than 4 times).

By clicking on the single node, it is possible to visualize the relative xml file (if it is available), which will be visualized in a new tab, within the same browser with all the metadata for a better explicability to the legal expert (see Fig. 1). More precisely, we used eXist-db to store and visualize each AKN document.

## 4 Conclusions

The derogation analysis provided much important information that could be used for better regulation:

- 6% of derogations are delegated to Member State legislation. This is interesting because we can detect the parts of derogations from the EU legislation that involve national law (relationship between EU and Member State law);
- A minimal part is connected with temporal conditions;
- Another minimal percentage is connected with ‘delegated acts’. Also, this is relevant in the relation between primary and secondary law in EU sources of law.

Three patterns (see Fig. 2) win out, linguistically speaking, in virtue of how useful they turn out to be in LEOS<sup>12</sup> (a specialized editor developed by the EU Commission for legal drafting) modelling for harmonizing derogations and markup (this is useful for the search engine):

- “By way of derogation from/to” - 26%;
- “shall not apply” - 20%;
- “By way of exception” - 11%.

We have also used Eurovoc classification to understand the topics which are more affected by derogations. We have discovered that the domains with more derogations are: import license, tariff, food and agriculture. The most frequent self-derogation falls in the financial/market domain regulation. Finally, the most frequent derogating acts belongs to the food, agriculture, regulation of the market topics, even if the recent emergency due to the COVID records a significant percentage of derogations.

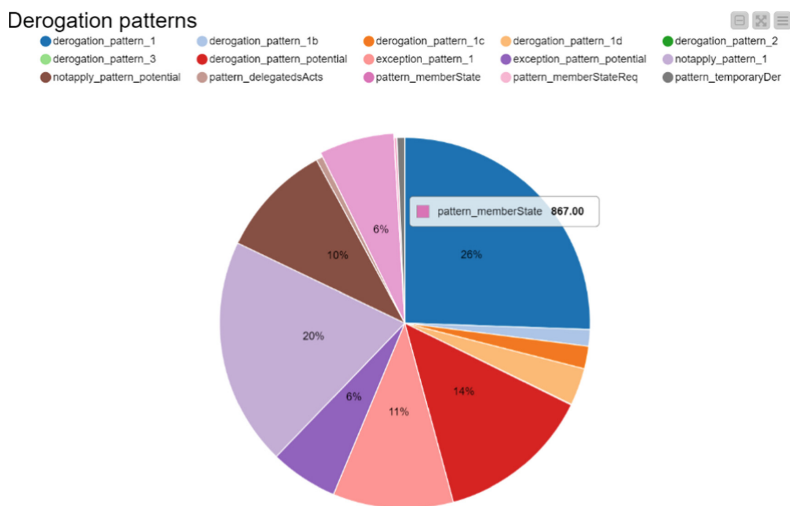


Fig. 2. Statistical about the pattern.

<sup>12</sup> LEOS “(Legislation Editing Open Software) is an open source software designed to help those involved in drafting legislation, which in itself is a complex knowledge-intensive process, by supporting efficient online collaboration.” <https://joinup.ec.europa.eu/collection/justice-law-and-security/solution/leos-open-source-software-editing-legislation>. Visited in June 2022.

**Acknowledgements.** This work was supported by the ‘Drafting legislation in the era of AI and digitisation’ with the support of the EU Commission – Directorate General Informatics Unit B2 – Solutions for Legislation, Policy & HR. The European Publication Office provided the documents and the datasets.

## References

- Alschner, W.: Legal Analytics. AI and the Law in Canada. In: Scassa, T., Martin-Bariteau, F. (eds.) LexisNexis Canada (2020). Available at SSRN: <https://ssrn.com/abstract=3662196>
- Ashley, K.: Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age. Cambridge University Press (2017)
- Bommarito II, M.J., Katz, D.M.: A mathematical approach to the study of the United States Code. *Physica A: Statistical Mechanics and its Applications*, 389 (2019)
- Boulet, R., Mazzega, P., Bourcier, D.: A network approach to the French system of legal codes, part I: analysis of a dense network. *Artificial Intelligence and Law*, 19 (2011)
- Chalkidis, I., et al.: LexGLUE: A Benchmark Dataset for Legal Language Understanding in English (13 Oct 2021)
- Coupette, C., Beckedorf, J., Hartung, D., Bommarito, M., Katz, D.M.: Measuring law over time: a network analytical framework with an application to statutes and regulations in the United States and Germany. *Front. Phys.* **9**, 658463 (2021)
- de Jong, H., van Dijk, G.: Network analysis in legal history: an example from the Court of Friesland. *Tijdschrift voor Rechtsgeschiedenis-Revue d Histoire du Droit-The Legal History Review*, 1–13 (2022)
- Filtz, E., Kirrane, S., Polleres, A.: The linked legal data landscape: linking legal data across different countries. *Artificial Intelligence and Law* **29**(4), 485–539 (2021). <https://doi.org/10.1007/s10506-021-09282-8>
- Fitsilis, F.: Artificial Intelligence (AI) in parliaments – preliminary analysis of the Eduskunta experiment. *The Journal of Legislative Studies* **27**(4), 621–633 (2021)
- Forst, R., Günther, K.: *Normative Ordnungen*. Suhrkamp Verlag (2021)
- Francesconi, E.: On the future of legal publishing services in the semantic web. *Future Internet* **10**(6), 48 (2018)
- Katz, D.M., Bommarito II, M.J.: Measuring the complexity of the law: the United States Code. *Artificial Intelligence and Law*, 22 (2014)
- Koniaris, M., Anagnostopoulos, I., Vassiliou, Y.: Network analysis in the legal domain: a complex model for European Union legal sources (v3) (2021). <https://arxiv.org/abs/1501.05237>
- Liga, D., Palmirani, M.: Classifying argumentative stances of opposition using Tree Kernels. In: ACAI 2019: Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, pp. 17–22 (2019)
- Palmirani, M.: Legislative change management with Akoma Ntoso. In: *Legislative XML for the semantic Web*, pp. 101–130. Springer (2011)
- Palmirani, M., Vitali, F.: Akoma Ntoso for legal documents. In: *Legislative XML for the semantic Web*, pp. 75–100. Springer (2011a)
- Palmirani, M.: Akoma Ntoso for making FAO resolutions accessible. In: *Knowledge of the Law in the Big Data*, pp. 159–169. IOS Press (2019)
- Palmirani, M., Cervone, L.: Measuring the Complexity of the Legal Order over Time. In: Casanovas, P., Pagallo, U., Palmirani, M., Sartor, G. (eds.) AICOL -2013. LNCS (LNAI), vol. 8929, pp. 82–99. Springer, Heidelberg (2014). [https://doi.org/10.1007/978-3-662-45960-7\\_7](https://doi.org/10.1007/978-3-662-45960-7_7)

- Robaldo, L., Villata, S., Wyner, A., Grabmair, M.: Introduction for artificial intelligence and law: special issue “natural language processing for legal texts.” *Artificial Intelligence and Law* **27**(2), 113–115 (2019). <https://doi.org/10.1007/s10506-019-09251-2>
- Rodríguez-Doncel, V., Palmirani, M., Araszkievicz, M., Casanovas, P., Pagallo, U., Sartor, G. (eds.): *AICOL/XAILA 2018/2020. LNCS (LNAI)*, vol. 13048. Springer, Cham (2021). <https://doi.org/10.1007/978-3-030-89811-3>
- Sovrano, F., Palmirani, M., Vitali, F.: Deep Learning Based Multi-Label Text Classification of UNGA Resolutions. *ICEGOV* (2020)
- van Kuppevelt, D., van Dijck, G., Schaper, M.: Purposes and challenges of legal network analysis on case law. In: Whalen, R. (ed.) *Computational legal studies: The promise and challenge of data-driven legal research* Edward Elgar Publishing (2020)
- Verheij, B.: Artificial intelligence as law. *Artif Intell Law* **28**, 181–206 (2020)