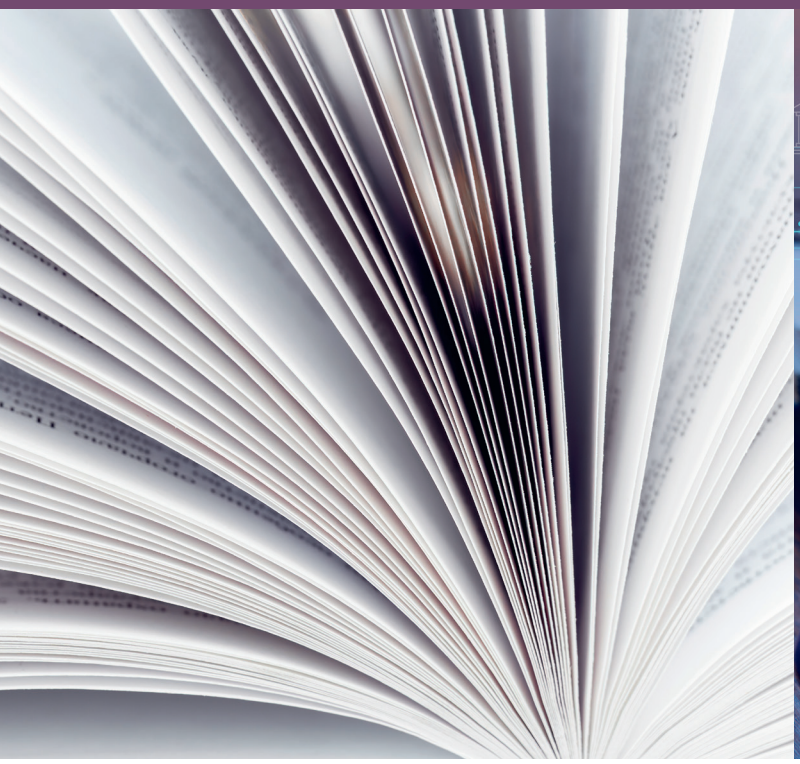
SEMANTIC CATEGORISATION OF JUDICIAL DECISIONS IN THE CASE LAW DATABASES WITH RECOMMENDATIONS

Foster transparency of judicial decisions and enhancing the national implementation of the European Convention on Human Rights



COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

SEMANTIC CATEGORISATION OF JUDICIAL DECISIONS IN THE CASE LAW DATABASES WITH RECOMMENDATIONS

Foster transparency of judicial decisions and enhancing the national implementation of the european convention on human rights

Monica Palmirani

Council of Europe

The opinions expressed in this work are the responsibility of the author(s) and do not necessarily reflect the official policy of the Council of Europe.

Cover design and layout: Documents and Publications Production Department (SPDP), Council of Europe

This publication has not been copy-edited by the DPDP Editorial Unit to correct typographical and grammatical errors

The project Foster Transparency of Judicial Decisions and Enhancing the National Implementation of the European Convention on Human Rights (TJENI) is funded by Iceland, Liechtenstein and Norway through the EEA and Norway Grants Fund for Regional Cooperation.

Iceland Liechtenstein Norway grants

Norway grants

Contents

Introduction

This report presents an analysis of the semantic categorization of court decisions in the context of digitalization, where different technologies, including the latest generation of artificial intelligence, is combined with human expert knowledge.

Some risks could arise considering the very sensitive environment of the judicial system. In December 2018, the European Commission for the Efficiency of Justice (CEPEJ), under the Council of Europe, adopted the first European guidelines on ethical principles relating to the use of artificial intelligence (AI) in judicial systems[1]. In 2019, the European Commission put out the Ethics Guidelines for Trustworthy Artificial Intelligence (AI)P[2], drawn up by the High-Level Expert Group on Artificial Intelligence (AI HLEG). Also, in light of the Convention for the protection of individuals with regard to the processing of personal data (Convention 108+)[3], the semantic categorisation of court decisions underscores some critical issues relating to data protection during processing.

If the risks are clearly and properly identified and addressed, they can be mitigated through an interdisciplinary hybrid method that combines different technologies on a law-by-design approach and uses organizational tools to evaluate the effectiveness and correctness of solutions over time. Indeed, with the evolution of society, legislation, and the case law, the task of categorising judicial decisions should be dynamic and always under *human control*.

This report remarks on these aspects and provides final recommendations, including tools (e.g., a checklist) for conducting a good analysis.

The report contains a glossary with the used terminology that can be used for a better understanding of the report.

1. Executive Summary

The semantic categorisation of judicial decisions has been a fundamental feature of the legal tradition from the outset (e.g., *Arbor actionum*).[4] Since the first collection of case law, thematic indexes, thesauruses, reviews, and glossaries have been used to make it easier to compare and retrieve material. The categorisation or classification of judicial decisions is thus a legal method that has been in use before their digitalisation.

Digital databases emerged in 1980s (with CD-ROMs and the like), and their evolution in digital online collections (from the mid-1990s to the 2000s) intensified their use for categorising legal knowledge, especially case law. Some case law databases are officially managed by the courts; others are edited by private publishers. In either case, there are several principles that need to be followed in making sure that the authoritative version of online legal content can be distinguished from unofficial sources, which could also be generates using AI (e.g., GPT4), raising the problem of the authenticity and quality of legal information. The source metadata of judicial material is therefore fundamental in avoiding unofficial or made-up legal information.

1. European Commission for the Efficiency of Justice, (December, 2018), European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, See link : https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c
2. European Commission, (April 8th, 2019), Ethics Guidelines for Trustworthy AI, See link: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419
3. Council of Europe, (1981), Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, See link: https://rm.coe.int/16808ade9d. It has been modernised by a 2001 additional Protocol.
4. The *Arbor actionum* is a classification of legal suits in Justinianian law. It was conceived in Bologna in the 12th century by the glossator Giovanni Bassiano. See http://mosaico.cirsfid.unibo.it/images/22%7D.

The categorisation or classification feature consists in extracting, analysing, and standardizing legal knowledge and mapping it onto a taxonomy in order to facilitate the task of retrieving information, ranking judicial decisions by relevance, and comparing of judicial materials by topic. The new generation of Artificial Intelligence techniques (e.g., Machine Learning), based on a statistical approach, together with natural language processing (NLP) and Semantic Web approaches, can provide some useful solutions for automatically classifying the relevant parts of decisions.

There is no doubt that digital technologies offer great opportunities in expanding our capabilities when it comes to searching for, discovering, navigating, processing, correlating, and analysing content in the immense and dynamic collections of court decisions. Some of the capabilities are as follows:

1. sorting decisions by topic;
2. sorting legal content by relevance;
3. grouping similar decisions;
4. visualising the web through which decisions are linked over time;
5. linking all decisions at different levels of the judiciary;
6. grouping case law by legal basis, ratio decidendi, logic reasoning, argumentation, and the facts of the case.

There are now several criteria for classifying court decisions:

1. By ***domain topic*** (e.g., civil law, IP law, criminal law). Glossaries, thesauri, and ontologies can be used to classify a decision's domain (its subject area). An example of such classification by topic is through the keywords used by the *European Journal*, as well as through ECLI metadata.
2. By ***ratio decidendi***: categorising the decisions into final ones and those that can be reviewed by courts.
3. By ***judicial reasoning***: with main focus on the argumentation used by judges to reach a conclusion in a case.
4. By ***citation***. Citations to statutory and judicial decisions make it possible to classify decisions by constructing a web of links through which they are connected.
5. By ***relevance***. As difficult as it may be to determine the relevance of legal decisions, they can nonetheless be ranked by governing principle, social impact, legal argument, and other criteria, asking, for example, whether the decision introduces a novel principle, has an impact on society, or develops an innovative line of argument.[5]

Box n. 1 - Emerging AI techniques like Machine Learning and NLP, along with Semantic Web approaches, ease the automating classification of judicial decisions. Digital technologies enhance capabilities in sorting, grouping, visualizing, and linking decisions by domain, ratio decidendi, judicial reasoning, citation, and relevance. Criteria for classifying court decisions encompass domain topics, ratio decidendi, judicial reasoning, citations, and relevance, each serving specific categorization purposes. The building, deployment and use of case law repositories, especially when functioning based on the resort to AI systems for the filtering and ordering of decisions, should therefore bear in mind the need to comply with such fundamental rights as protected by the European Convention on Human Rights. With regard to the right to a fair trial (Article 6) in particular, the independence and impartiality of the judge must be ensured: hence, the need to guarantee unwarranted influences caused by the resort to AI. Moreover, the principle of fairness recognized by Article 6 ECHR also entails the necessity that such systems follow the ethical principles and standards for ensuring fair AI. The semantic categorisation of judicial decisions may involve the processing of personal data of the parties involved in the cases, such as their names, addresses, health conditions, etc. According to its Article 3 ("Scope"), Convention 108+ applies to this type of processing and requires that it respects the rights and freedoms of the data subjects.

5. On the concept of "relevance" see Marc Van Opijnen and Cristiana Santos. 2017. On the concept of relevance in legal information retrieval, Artif. Intell. Law 25(1) (March 2017): 65–87. https://doi.org/10.1007/s10506-017-9195-8.

2. Legal and Ethical Issues

Classification of the case law by legal experts existed even before database collections we digitized. This was done on the basis of legal criteria[6]. Several techniques are used in digitizing legal decisions and automating the work done with them, such as extracting salient parts of decisions and reusing them for other purposes such as calculating statistics, predicting future case law, correlating the decisions of different instances (first and appeal), clustering similar case law, and detecting analogies. The new generation of artificial intelligence (including sub-symbolic and non-symbolic AI) uses annotated corpora of case law to detect hidden legal knowledge for the legal expert (e.g., clustering of similar cases) and enable other novel applications such as predicting the likelihood that a party to a suit will prevail on appeal.

However, these techniques need to be integrated with a robust semantics to avoid the risk of confusing the general common language with specific legal terminologies (as happens with AI hallucination), misreading legal discourse (its figures of speech, metaphors, similitudes, etc.), or failing to capture the methods of legal argumentation (deductive, inductive, abductive, analogical, evidence-based, etc.).[7] Integration at the semantic level provides meaning in context (e.g., European legislation), as well as the rules and principles of legal theory and of procedural business logic. Additionally, it is an emerging research area called neuro-symbolic AI.[8]

There are, however, some risks related to the semantic categorisation of the court decisions, like the following.

1. ***Crystallization***. Classification depends on where a legal system is historically, on its temporal development. For example, the classification of the concept of "stalking" is evidently a new crime, but other legal concepts evolve over time (e.g., EU citizenship in determining jurisdiction over a contract governed by private international law). If the classification model does not account for change over time, it risks crystallizing legal concepts at given points in time.
2. ***Limiting innovation***. If the classification model is not dynamic, we risk limiting innovation in judicial interpretation. If the classification mechanism works better with quantitative big data and struggles with qualitative data, a novel decision will have a lower ranking than settled landmark decisions.
3. ***Polarization and the filter-bubble effect***. If the classification algorithm agglomerates similar decisions and foregrounds them in searching, we risk a polarization effect where similar decisions are placed into a "filter bubble" that hides dissimilar or contrasting decisions from view.
4. ***Bias***. Automatic semantic categorisation could lead to bias and discrimination by reiterating legal concepts from past law. This happens when the algorithms entrench a series of historical semantic categorisations reflecting a society based on different rights, legal principles, and regulations. Here we face the risk of reinforcing prejudices in areas that have since been regulated differently (e.g., gender balancing, hate speech, fake news).
5. ***Visualization distortion***. Even the way in which the categorized data and documents are visualised can be biased, giving the end-user a distorted perception of what has been categorised, depending on such factors as the criteria chosen for grouping documents (e.g., by year or topic), the colour scheme (e.g., red for drawing attention), and the overall layout and the design elements (e.g., less important decisions set in smaller type).[9]
6. ***Emphasizing relevance***. An algorithm may distort a decision's relevance (its ranking in the list of most widely read decisions) by assigning recursive importance depending on how frequently the decision has been viewed. The more frequently a decision is viewed by end-users, the more likely is it that the system will select it for viewing the next time a user enters a search query.
7. ***Limiting autonomy***. Legal experts should be able to consult a collection of decisions without filters, as that will support the end-user's autonomy in accessing legal information without barriers, and the tool

6. CJEU, "Digest of the case-law", https://curia.europa.eu/jcms/jcms/Jo2_7046/en/.
7. Bongiovanni Giorgio, Postema Gerald, Rotolo Antonino, Sartor Giovanni, Valentini Chiara, Walton Douglas. (2018). Handbook of Legal Reasoning and Argumentation. Springer.
8. Hitzler P., Sarker M.K (eds). 2022. Neuro-Symbolic Artificial Intelligence: The State of the Art. Vol. 342 of Frontiers in Artificial Intelligence and Applications. Amsterdam: IOS Press.
9. An example of a user-friendly graphical layout designed to visualise information without bias is this portal collecting the judgments and rulings issued by the Constitutional Court of Italy: http://bach.cirsfid.unibo.it/ldms-cortecostituzionale/#/pronunce. In this portal, developed by the University of Bologna with the Court's authorisation, users can view the legislation and case-law cited in each judgment and ruling, each within its own category.

that makes it possible to access all decisions should be easy to use. This principle is also referred to as the "human-in-command"[10] or "under user control".[11] If this principle is to be effective, the user interface needs to provide information about the categorisation process, making it possible for legal experts to make their own assessment in that regard.

8. ***Silence***. The automatic semantic categorisation algorithm needs to avoid "silence" while searching, which is what would happen if it cannot find any topic or criteria on which basis to classify a decision.

9. ***Prediction***. Semantic categorisation is used in making predictions on the basis of criteria such as past trends, similarity among decisions and arguments, or the frequency with which a decision is viewed or cited. Although it is definitely a worthwhile goal to harmonize decisions, we shouldn't forget that, as society changes, including in response to drastic events (such as the COVID pandemic), so does legislation and the case law, introducing new relevant opinions to deal with a factual landscape that can no longer be relied on to be stable or to exhibit the same regularity.

10. ***Lack of authenticity***. Semantic categorisation can also be done using general-purpose market tools developed by Big Tech like GPT-4 and Bard. However, unless this categorisation is validated by humans with legal expertise, it could not be considered authoritative. Big Tech often uses classification to improve performance in their system, pursuing their specific business goals. Semantic categorisation by an expert authority (e.g., a court, a government agency, the Ministry of Justice) seeks to provide legal information consistently principles of neutrality, impartiality, transparency, equality, fair access, quality, and accuracy. A private publisher is instead seeking to sell an advanced legal informatics service to legal experts (e.g., lawyers),[12] and here, too, it is essential to maintain quality standards to ensure that legal sources are accurate. For this reason, considering the multiple players acting in the legal information domain pursuing different goals, it is fundamental to use a provenance ontology to provide the classification metadata with additional information about the authority, the date of classification, the office in charge, and so on.[13]

Box n. 2 - Artificial Intelligence utilizes annotated corpora for legal knowledge extraction and prediction. However, risk arises from semantically integrating AI with legal context, potentially crystallizing concepts, limiting innovation, causing polarization, bias, visualization distortion, relevance emphasis, and limiting autonomy. Further risks include prediction issues, lack of authenticity, and the impact of big tech tools on legal classification, emphasizing the need for provenance ontology in metadata. Article 5 of Convention 108+ ("Legitimacy of data processing and quality of data") sets out the principles of data quality, such as lawfulness, fairness, proportionality, accuracy, relevance, and security. The semantic categorisation of judicial decisions may affect the quality of data in several ways, such as by introducing errors, biases, inconsistencies, or distortions in the data or the results.

3. Methodological Principles

In order to minimise the risks linked to the semantic categorisation that are listed above, the following several principles shall be taken into account during the design of the categorisation.

1. **Transparency**. The annotation, filtering, and search criteria need to be clearly understandable by end-users, especially by the decision-maker (e.g., judges, lawyers, civil servants). The annotation process, the algorithm used, the work done by legal experts, and their level of expertise need to be clearly described, so that it

10. This idea is defined in HLEG 16 as "the capability to oversee the overall activity of the AI system (including its broader economic, societal, legal and ethical impact) and the ability to decide when and how to use the system in any particular situation".
11. In the "CEPEJ European Ethical Charter on the Use of Artificial Intelligence (AI) in Judicial Systems and their Environment" this idea of human control is understood as "precluding a prescriptive approach and ensuring that users are informed actors and in control of their choices".
12. See the "Guide on the Use of Artificial Intelligence-Based tools by Lawyers and Law Firms in the EU," issued by the Council of Bars and Law Societies of Europe: https://www.ccbe.eu/fileadmin/speciality_distribution/public/documents/IT_LAW/ITL_Reports_studies/EN_ITL_20220331_Guide-AI4L.pdf.
13. Three examples of an ontology used to track the provenance of information are (i) PROV-O https://www.w3.org/TR/prov-o/, put out by the W3C consortium; (ii) DCAT-AP, used by the European Commission for Open Government Data (https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe/release/11); and (iii) PREMIS. https://www.loc.gov/standards/premis/ used by many libraries like Government Publication Office of USA. See Se also Angela Dappert, Rebecca Squire Guenther, Sébastien Peyrard, 2016, Digital Preservation Metadata for Practitioners (Cham: Springer), https://doi.org/10.1007/978-3-319-43763-7_10.

can be audited with credible evidence. Also, the user interface needs to be designed in such a way as to avoid distorting information or drawing attention to only a part of the document collection. In other words, users should at any moment be able to access all the documents and have clear explanations of the ranking criteria.

2. **Explicability**. As far as the main parts of the automatic annotation process is concerned, its results should be explicable so as to make it possible for legal experts (e.g., judges, lawyers) to have all the information they need to check that the annotation is correct. Access to the information requires that the license for the data collection and for the metadata be clearly published.

3. **A user-centred system**. The semantic categorisation system should be user-centred, with different levels of explanation depending on the kind of user being addressed. Moreover, it is very important that results be visualised in such a way as not to give rise to misperception (e.g., an illusion of completeness) or distortion (e.g., a statistics bar chart that uses absolute instead of relative values).

4. **Anonymisation**. Decisions need to be classified in such a way that it is not possible to reidentify the data subjects or parties involved in the case, all of whom need to be anonymised. The rules for anonymising or pseudonymising such data are specific to the judicial system, and each country has its own specifically tailored domestic regulations, consistent with the rules on open government data and on data sharing.

5. **Standardization**. Semantic categorisation needs to adhere to technical standard like a common vocabulary (e.g., EUROVOC),[14] formal ontologies designed using well-known patterns (e.g., ELI and ECLI,[15] UNDO)[16] or vocabularies (e.g., WordNet)[17] recognised in the Semantic Web community. Standardization makes it possible to ensure that solutions are technically sound, platforms are interoperable, and the semantic model is resilient over time.

6. **Documentation**. Semantic categorisation needs to be documented with all necessary information about (i) the type of annotation (automatic, manual, hybrid); (ii) the technology adopted; (iii) the classification criteria; (iv) the frequency of annotation (e.g., quarterly); (v) the user's legal expertise (e.g., a different level for judges, court clerks, annotators); (v) the authoritativeness of the validation; (vii) the validation method (e.g., double-checking); (viii) the metric for measuring the technical accuracy and quality; and (ix) the license for the documents, metadata, and code.

7. **Temporal modelling**. Decisions will have different categorisations at different points in time depending on how the statutes and the case law change in response to a changing society, as well as on the changing technology. For this reason, it is important to periodically update the semantic categorisation criteria and to model the technical system so as to enable multilevel semantic annotation according to temporal parameters. For example, the concept of "European citizen" has been changing over time, and the annotation should be synchronized with the historical period (e.g., Brexit for the UK, accession for Malta). Similarly, a categorization of relevance concerning the "right to be forgotten" should be updated to reflect new case law and legislation. Classification is therefore dynamic, and the labelling could change over time. This means that we have to design a user interface that is clear and user-friendly, in order to not give rise to confusion in end-users and to inform them about changes in legislation and the case law and about any other changes that may affect categorization.

8. **Organization**. Classifications need to be published and constantly updated, informing the end-user about the steps and procedure used in the classification process, as well as the type of legal expertise and periodical checking.

9. **Ethics compliance**. Especially if the classification is done using artificial intelligence techniques, the technical solution adopted needs to be checked for compliance with an ethical code so as to avoid prejudice or bias (e.g., gender bias), discrimination (e.g., ethnic discrimination). The Council of Europe CEPEJ and European

14. EUROVOC is a multilingual legal ontology for classifying the legislation published by the Publications Office of the European Union: https://op.europa.eu/en/web/eu-vocabularies.
15. ELI is the European Legislation Identifier: https://eur-lex.europa.eu/eli-register/about.html?locale=en. ECLI is the European Case-Law Identifier: https://e-justice.europa.eu/175/EN/european_case_law_identifier_ecli?init=true. See van Opijnen, Marc and Ivantchev, Alexander, "Implementation of ECLI: State of Play" (October 10, 2015). JURIX 2015: The Twenty-Eighth Annual Conference on Legal Knowledge and Information Systems, Available at SSRN: https://ssrn.com/abstract=2706768.
16. United Nations Ontology: https://unsceb-hlcm.github.io/onto-undo/.
17. WordNet is a large lexical database for English: https://wordnet.princeton.edu/. See M.T. Sagri and D. Tiscornia, "Metadata for Content Description in Legal Information," 14th International Workshop on Database and Expert Systems Applications, 2003. Proceedings, Prague, Czech Republic, 2003, pp. 745–749, doi: 10.1109/DEXA.2003.1232110.

Commission have put out guidelines for evaluating ethical compliance when AI is used,[18] and many AI regulation strategies[19] include a risk assessment and accountability approach for evaluating the ethical issues raised by AI and the principles by which to assess its use.[20]

10. **Legal analysis**. The semantic annotation criteria need to undergo a legal analysis defining a precise qualitative and quantitative methodology. Particular attention should be dedicated to the standards set out in the Convention 108+, GDPR and the AI Act. In general, the blueprint for developing a legal-by-design semantic categorisation methodology [21] should be Convention 108+ [22] and the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment.[23]

11. **Technical soundness**. The technical solution adopted needs to be valid, sound, accurate, secure, and updated to reflect the state of the art. For this reason, the organization responsible for this task should periodically check the robustness of the solution to see if some risks or weakness come to light in view of practical usage, research, or legislation. The frequency of use of some legal terms in the search engine could involuntarily misrepresent the relevance of some decisions. This mechanism should be balanced with other qualitative and quantitative criteria.

12. **Quality validation**. The technical solution needs to be checked for quality to see whether the results meet legal and ethical standards. It also needs to undergo a risk assessment analysis as required under proposed legislation (and in particular under the AI Act proposal for what concerns the Fundamental Rights Impact Assessment). Additionally, if datasets are annotated by hand to prepare them for machine-learning applications, a code of practice needs to be drafted for a robust annotation methodology.[24]

13. **Multilingual categorisation**. Semantic categorisation is language-dependent, and in courts where multilingual documents are key to democratic participation and to equal access through all official languages, it is essential that keywords be correctly translated.[25] There are many studies on legal translation,[26] as well as studies addressing the interpretation problems involved in this task. If semantic categorisation is automatically done by machine algorithms (e.g., machine learning, large language models) it becomes crucial to be able to detect mistranslation.

14. **Crowdsourcing annotations**. Some web portals make it possible to annotate decisions using the expertise of lawyers, legal practitioners, law professors, and graduate and postgraduate students. "Grassroots" methods of tagging (e.g., folksonomy) are quite easy to implement, but they carry some risks (as outlined in Section 4 below). This process can be helpful in detecting mistakes or in quickly bringing the categorisation mechanism up to date when legislation or the case law change. However, using the external experts to annotate decisions is quite risky because the validation process cannot count on a homogenous application of the categorisation method. Additionally, although different interpretations could enhance a collection of decisions, it is doubtful whether this will result in a neutral, impartial, or explicable process. For this reason, it is essential that the provenance of legal information be tracked so as to guarantee authenticity and authoritativeness. The

18. "CEPEJ European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment" (2018), https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c., "Ethics Guidelines for Trustworthy Artificial Intelligence (AI)", a document drafted by the High-Level Expert Group on Artificial Intelligence (AI HLEG), 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419.
19. The European AI strategy is called "EU AI Act: First Regulation on Artificial Intelligence": https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence. The UK AI strategy is called "A Pro-Innovation Approach to AI Regulation": https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1176103/a-pro-innovation-approach-to-ai-regulation-amended-web-ready.pdf. The Canadian AI strategy is laid out in "The Artificial Intelligence and Data Act (AIDA)—Companion Document": https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document.
20. See MICKLITZ, Hans-Wolfgang, POLLICINO, Oreste, REICHMAN, Amnon, SIMONCINI, Andrea, SARTOR, Giovanni, DE GREGORIO, Giovanni (eds.). (2022). Constitutional Challenges in the Algorithmic Society. Cambridge: Cambridge University Press. https://hdl.handle.net/1814/74296.
21. Guiding Principles for Automated Decision-Making in the EU. ELI Innovation Paper.
22. Convention for the Protection of Individuals with regard to the Processing of Personal Data: https://www.coe.int/en/web/data-protection/convention108-and-protocol.
23. Available at https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.
24. See Braun, D. (2023). "I Beg to Differ: How Disagreement Is Handled in the Annotation of Legal Machine Learning Data Sets." *Artif Intell Law*. https://doi.org/10.1007/s10506-023-09369-4.
25. Malta has two official languages: Maltese and English. Switzerland has three official languages: French, German, and Italian.
26. See Van der Jeught, S. (2018). "Current Practices with regard to the Interpretation of Multilingual EU Law: How to Deal with Diverging Language Versions?" *European Journal of Legal Studies*, 11 (1): 5–38; see also European Parliament (2017). Legal aspects of EU multilingualism (http://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2017)595914); European Commission, Directorate-General for Translation, Study on Language and Translation in International Law and EU Law : Final Report, Publications Office, 2012 (https://data.europa.eu/doi/10.2782/64020).

use of LLMs (e.g., ChatGPT) with querying (e.g., prompting) by legal experts may result in a partial vision of a specific case, and this could modify the balancing of the classification mechanism.

Box n. 3 - The successful implementation of semantic categorization in legal systems has to rely on a spectrum of legal and ethical principles. These encompass ensuring transparency in processes, facilitating the explicability of outcomes, adopting a user-centric design approach, anonymising data for privacy, standardizing methodologies, maintaining thorough documentation, accommodating temporal adaptability, ensuring organizational clarity, upholding ethics compliance, conducting robust legal analysis, ensuring technical soundness, validating quality, considering multilingual aspects, cautiously leveraging crowdsourcing, and adhering to evolving legal standards and technological advancements for comprehensive effectiveness. These principles promote compliance to Article 5 of Convention 108+ ("Legitimacy of data processing and quality of data"), which requires that personal data are processed in a way that ensures their quality and reliability, accuracy, and relevance. The proposed principles ensure data quality, transparency, explicability, standardization, documentation, temporal modelling, technical soundness, and quality validation. Moreover, the proposed principles relate to Article 8 of Convention 108+ ("Transparency of processing"), which requires that data subjects are informed about the processing of their data, such as the purpose, the logic, the consequences, and the rights of access and rectification. The passage mentions several principles that are related to transparency of processing, such as user-centred system, anonymisation, ethics compliance, and legal analysis.

4. Semantic Categorisation

4.1. Design Approaches

Semantic categorisation can be implemented using different methods. All of them present critical issues, and for this reason the emerging state-of-the-art trend is to mix different approaches in order to make up for weakness and enhance the prospects for success.

One of the methods that could address many of previously listed issues is the Open Government Data approach, specifically Linked Open Data and FAIR,[27] and governed by the Public Sector Information Directive.[28] Also, the Data Governance Act[29] mentions data sharing as a good practice for reinforcing government services, the democratic system, digital citizenship, and the digital economy. Linked Open Data is a method for representing digital knowledge that makes it possible to reuse decisions classified with a standard technical method (Resource Description Framework, or RDF). This method runs into several barriers, including from domestic law, when it comes to making the text of judicial decisions available as Open Data, but for classification metadata, sharing this information within the community presents fewer legal and ethical issues. Sharing the annotations and classifications used to categorise decisions (sharing the corpus of annotated metadata) supports research and AI applications in the legal domain.

There are seven main tasks in semantic categorisation:

1. **Defining the legal categorization methodology**. This step is fundamental in defining the principle in legal theory on which basis judicial decisions are classified, as well as in defining the process for monitoring legal and ethical issues. This step also defines the framework for semantic categorization and organisation, as well as the validation process, which is fundamental in providing explanations of the criteria adopted.

2. **Extraction of legal knowledge from decisions**. This step aims to extract legal knowledge on the basis of the technical framework defined in step 1.[30]

3. **Classification of legal knowledge**. This step classifies, groups, and compares the extracted information.

27. For FAIR we mean the acronym coming from some pillar characteristics of the open data collection: findable, accessible, interoperable and re-usable. See Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* (2016). "The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci Data* 3: 160018. https://doi.org/10.1038/sdata.2016.18. Jones, Sarah, and Grootveld, Marjan. (2017). How FAIR Are Your Data? Zenodo. https://doi.org/10.5281/zenodo.1065991.
28. Directive (EU) 2019/1024, https://eur-lex.europa.eu/eli/dir/2019/1024/oj (recast).
29. Regulation (EU) 2022/868, http://data.europa.eu/eli/reg/2022/868/oj.
30. For a survey of the last thirty years of advancing research in AI and Law see Villata, S., Araszkiewicz, M., Ashley, K. et al. (2022). "Thirty Years of Artificial Intelligence and Law: The Third Decade." *Artif Intell Law*, 30: 561–591. https://doi.org/10.1007/s10506-022-09327-6z.

4. **Representation of legal knowledge**. When the legal knowledge is extracted and validated, the results should be represented digitally using a standard vocabulary and methodology in order to facilitate the task of sharing and managing that knowledge over time using Semantic Web technologies. The representation method also affects future applications and the technological neutrality with respect to vendors. ECLI,[31] ELI, AKN4EU,[32] PERMIS, DCAT-AP,[33] PROV-O,[34] and EUROVOC are just some standards that could be used.

5. **Usage of the legal knowledge**. Semantic categorization can be used for different purposes, like information retrieval and AI prediction. This task is separate from proper semantic categorization and needs a different assessment process, that shall be based on the basis of the CEPEJ European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and AI Act proposal.

6. **Visualization of legal knowledge**. Different software packages can be used to visualise the results of categorization,[35] and the main risk is that the visual representation may distort the data (as described above).

7. **Explaining the results**. Clearly communicating results is key to ensuring transparency and the autonomy of decisions, while making it possible to detect possible bias unwittingly injected into the system.

All these steps need to be analysed carefully through a legal and ethical analysis using the checklist in attachment.

There are different approaches we can use to go about semantic categorisation and to implement the different technical methods:

1. ***Manual annotation***. Judicial decisions are annotated manually by legal experts only.

2. ***Unsupervised annotation***. Decisions are annotated automatically using only an automated system (e.g., through deep learning).

3. ***Supervised annotation***. Decisions are classified automatically based on a previous dataset (e.g., baseline) manually annotated by legal experts.

4. ***Semi-supervised annotation***. Decisions are classified automatically based on a partially manual annotation.

We recommend a methodology that in any event includes a documented and stable process of human annotation and validation by legal experts and using a robust code of practice in order to ensure a homogenous use of legal semantic knowledge. It is good practice to have (i) an annotation *manual*; (ii) *indicators of technical evaluation* (e.g., F1); and (iii) *criteria for legal and technical validation* (e.g., end-user validation).

4.2. A Survey of Techniques

Following is a survey of some types of annotation and semantic categorisation techniques.

4.2.1. Controlled Vocabulary, Taxonomies, Thesauri

The use of controlled vocabularies, taxonomies, and thesauri consists in creating a standard for codifying reality (e.g., a country), organising linguistic terms (taxonomies), and defining legal-concept hierarchies (thesauri). It is not easy to build these instruments, since it takes great effort to find the best terminology, one that is understandable, representative, meaningful, and resilient over time. This is the approach most widely used by the courts' official portals and by open data services. EURLex, CURIA, Find Case Law (of the UK National Archives),[36]

31. See Agnoloni, Tommaso, Lorenzo Bacci, Ginevra Peruginelli, Marc van Opijnen, John Van Den Oever, Monica Palmirani, Luca Cervone, Octavian Bujor, Arantxa Arsuaga Lecuona, Alberto Boada García, Luigi Di Caro and Giovanni Siragusa. (2017). "Linking European Case Law: BO-ECLI Parser, an Open Framework for the Automatic Extraction of Legal Links." International Conference on Legal Knowledge and Information Systems.
32. AKN4EU (https://op.europa.eu/it/web/eu-vocabularies/akn4eu) is a customization of the OASIS LegalDocML standard for representing European legal documents. LegalDocML is also mentioned as an interoperable standard for legal sources in the ISA^2 strategy.
33. A metadata catalogue for describing open datasets and their distribution: https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe/release/11.
34. An ontology for defining the provenance of a dataset: https://www.w3.org/TR/prov-o/.
35. Python and D3.js libraries: https://d3js.org/. See also Stancin, Igor and Alan Jović. (2019). "An Overview and Comparison of Free Python Libraries for Data Mining and Big Data Analysis." 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO): 977–982.
36. This web portal aims to publish open decisions in Akoma Ntoso XML format annotated with Linked Open Data metadata: https://caselaw.nationalarchives.gov.uk/.

and HUDOC [37] all use a controlled vocabulary of keywords to assign a thematic categorization to decisions. The European Commission has a good example of the use of taxonomies,[38] thesauri,[39] and controlled vocabularies[40] for creating a standardised way to name basic legal terms (taxonomy), legal concepts (thesauri), and common elements (e.g., authority). However, these instruments need to be updated frequently to track changes in legislation, case law, and society; otherwise, the classification may become obsolete, opening the door to mistakes. In meantime, there is also a need to maintain past legal knowledge and heritage for historical and comparative purposes. A controlled vocabulary can be updated using tools like VocBench.[41] Another fundamental task is to manage the multilingual legal vocabularies and synchronise them with different legal terminologies from a linguistic point of view (e.g., synonyms, hypernyms, hyponyms, acronyms, abbreviations, etc.).[42]

4.2.2. Facets and Folksonomy

Facets classification (or the folksonomy approach) consists in accepting tags that end-users assign in classifying judicial decisions. This method is widely used in some applications (e.g., social media, online news), but it could give rise to problems in legal theory due to the use of nonlegal terminologies and common-sense keywords. In addition, the frequency with which a specific community (e.g., tax lawyers) visits a given web portal (e.g., a revenue and customs website) may polarise tagging in that specific domain within that profession.[43] This methodology can be used to integrate formal closed vocabularies or formal ontologies so as to capture trends among end-users, but the inputs need to be carefully evaluated by legal experts.

4.2.3. Formal Legal Ontology

Also helpful are formal legal ontologies,[44] like ELI and ECLI,[45] or Akoma Ntoso,[46] or ontologies specific to a given legal domain (e.g., PrOnto for privacy,[47] IPROnto[48] for intellectual propriety rights, criminal procedural law,[49] etc.). These ontologies can be used to model relationships between legal concepts and achieve a more sophisticated representation of legal knowledge. For example, in civil law, the concept "damages" could be specified in subclasses like "lost profits" and "consequential damages"; "lost profits" is in turn related to "income" like "salary", "pension", or "business income". A legal procedure could require specific actions from the actors (e.g., attorney, plaintiff, defendant), and these relationships could be modelled so as to better categorise decisions and to use terms to infer from them more knowledge about the case law. Legal ontologies are used in this sense by the European Publications Office to favour open data sharing.[50] Legal relationships are defined at ontology level, and they can be used to enrich the search engine using different techniques, including legal knowledge graph (LKG) technology.[51] But the knowledge graph based on metadata RDF model assertions can

37. HUDOC User Manual: https://www.echr.coe.int/documents/d/echr/HUDOC_Manual_ENG. HUDOC search: https://hudoc.echr.coe.int/.
38. Taxonomies: https://op.europa.eu/en/web/eu-vocabularies/taxonomies.
39. Thesauri: https://op.europa.eu/en/web/eu-vocabularies/thesauri.
40. Controlled vocabularies: https://op.europa.eu/en/web/eu-vocabularies/controlled-vocabularies.
41. VocBench: https://op.europa.eu/en/web/eu-vocabularies/vocbench.
42. Ajani, G., Boella, G., Caro, L.D., Robaldo, L., Humphreys, L., Praduroux, S., Rossi, P., Violato, A. (2016). "The European Taxonomy Syllabus: A Multilingual, Multi-Level Ontology Framework to Untangle the Web of European Legal Terminology." Appl. Ontology 11(4): 325–375.
43. For an analysis of the theoretical implications of folksonomy in the legal domain, see Costantini, F. (2014). "#Folksonomies & #Law: From 'Quid Juris?' to 'Quid Jus?' to 'Cur Jus?'" In Hoekstra, R., Breuker, J., Guarino, N., Kok, J. N., Liu, J., Lopez de Mantaras, R., Mizoguchi, R., Musen, M., Pal, S. K., & Zhong, N. (eds.), JURIX 2014. Frontiers in Artificial Intelligence and Applications, proceedings of the twenty-seventh annual conference, pp. 203–204. Amsterdam: IOS Press, 2014.
44. See Cleyton Mário de Oliveira Rodrigues, Frederico Luiz Gonçalves de Freitas, Emanoel Francisco Spósito Barreiros, Ryan Ribeiro de Azevedo, and Adauto Trigueiro de Almeida Filho (2019). "Legal Ontologies Over Time: A Systematic Mapping Study." Expert Syst. Appl. 130, C (Sept. 2019): 12–30. https://doi.org/10.1016/j.eswa.2019.04.009.
45. See footnote 12.
46. Akoma Ntoso OASIS LegalDocumentML (LegalDocML) standard: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml.
47. Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., Robaldo, L. (2018). "PrOnto: Privacy Ontology for Legal Reasoning." In Kő, A., Francesconi, E. (eds) Electronic Government and the Information Systems Perspective. EGOVIS 2018. Lecture Notes in Computer Science, vol. 11032. Cham: Springer. https://doi.org/10.1007/978-3-319-98349-3_11.
48. Intellectual Property Rights Ontology (IPROnto): https://dmag.ac.upc.edu/ontologies/ipronto/.
49. See Melissa Zorzanelli Costa, Giancarlo Guizzardi, João Paulo A. Almeida (2022). "On Capturing Legal Knowledge in Ontology and Process Models Combined." In Enrico Francesconi, Georg Borges, Christoph Sorge, editors, Legal Knowledge and Information Systems – JURIX 2022: The thirty-fifth annual conference, Saarbrücken, Germany, 14–16 December 2022. Vol. 362 of Frontiers in Artificial Intelligence and Applications, pp. 267–272. IOS Press.
50. Francesconi, E., Küster, M.W., Gratz, P., Thelen, S. (2015). The Ontology-Based Approach of the Publications Office of the EU for Document Accessibility and Open Data Services. In Kő, A., Francesconi, E. (eds) Electronic Government and the Information Systems Perspective. EGOVIS 2015. Lecture Notes in Computer Science, vol. 9265. Springer: Cham. https://doi.org/10.1007/978-3-319-22389-6_3.
51. Elena Montiel-Ponsoda, Víctor Rodríguez-Doncel, Jorge Gracia (2017). "Building the Legal Knowledge Graph for Smart Compliance Services in Multilingual Europe." TERECOM@JURIX 2017: 15–17.

be a good classification tool under some conditions, as when it is used for objective categorization criteria (e.g., date of publication, citations, references). The knowledge graph can infer connections navigating through relationships, but the absence of relationships doesn't mean that there is no information. Especially if a temporal model is not well designed (e.g., a legal concept is valid for old case law but is not applicable today) the LKG could navigate nodes that are not appropriate for the context (e.g., a legal case could be classified as related to Brexit even if it is dated 2000). For this reason, we use LKG only to extract an approximate categorisation of a corpus of decisions by relying more on relationships between legal concept definitions than on rules and reasoning. In order to achieve greater sophistication in extracting the legal reasoning relating to the grounds of a decision or to its *ratio decidendi*, we need a more complex technical tool (e.g., defeasible logic or rule-base systems).[52] The legal ontology is in any case a fundamental pillar and a good practice in the legal domain when it comes to modelling legal terms, concepts, and knowledge in a standard format that can be managed by human beings and machines and can also be reused by other applications (e.g., AI applications). There are some paradigms, such as Linked Open Data, that can be followed to model a good legal ontology from a theoretical and technical standpoint.[53] Some good examples of the use of legal ontologies and legal-rule modelling can be found in Francesconi and Governatori 2023,[54] as well as in Filtz, Kirrane & Polleres 2021[55] and Vanderstichele 2021[56] (which also integrates the legal design principles for transparent communication). Also, in legal ontologies we face the problem of multilingualism, which needs to be addressed to guarantee the semantic equivalence of relationships between legal concepts in all languages.[57]

4.2.4. Text Mining and NLP

One of the classic approaches used to extract legal knowledge from the text of decisions is text mining,[58] which uses a combination of NLP technologies.[59] This is the approach the European Commission's Joint Research Centre adopted to classify documents and monitor European policies.[60] However, it is difficult nowadays to clearly sperate NLP tools from AI techniques because they are integrated. We can see several tools that can be used to categorise decisions:

1. Named Entity Recognition (NER),[61] for identifying specific qualified parts of a text (e.g., agent, authority, organization, time, date) and using them for better classification.

2. Treebank, for analysing the syntax of sentences (as by tagging parts of speech: verb, noun, adjective, etc.).

3. Structural topic models,[62] for identifying a specific prevalent topic running through a document's narrative.

---

52. Athan, T., Governatori, G., Palmirani, M., Paschke, A., Wyner, A. (2015). *Legal-RuleML: Design Principles and Foundations*, pp. 151–188. Springer International Publishing.
53. For a robust methodology for modelling legal ontologies, see Giovanni Sartor, Pompeu Casanovas, Mariangela Biasiotti, Meritxell Fernández-Barrera. 2011. *Approaches to Legal Ontologies: Theories, Domains, Methodologies*. Springer.
54. Francesconi, E., Governatori, G. (2023). "Patterns for Legal Compliance Checking in a Decidable Framework of Linked Open Data. *Artif Intell Law* 31: 445–464. https://doi.org/10.1007/s10506-022-09317-8.
55. Filtz, E., Kirrane, S. & Polleres, A. (2021). "The Linked Legal Data Landscape: Linking Legal Data across Different Countries. *Artif Intell Law* 29: 485–539. https://doi.org/10.1007/s10506-021-09282-8.
56. Vanderstichele, Geneviève (2021). "Knowledge Graphs as an Example of Legal Design to Model Legal Analytics for Adjudication with respect to the Rule of Law (March 17, 2021). In Marcelo Corrales Compagnucci, Helena Haapio, Margaret Hagan and Michael Doherty (eds.), *Legal Design: Integrating Business, Design, & Legal Thinking with Technology* (Edward Elgar). SSRN: https://ssrn.com/abstract=3806781.
57. Doncel VR, Ponsoda EM. 2020. "LYNX: Towards a Legal Knowledge Graph for Multilingual Europe." *Law in Context: A Socio-legal Journal* 37(1): 175-8. https://journals.latrobe.edu.au/index.php/law-in-context/article/view/129.
58. Katz, Daniel Martin and Hartung, Dirk and Gerlach, Lauritz and Jana, Abhik and Jana, Abhik and Bommarito, Michael James (2023). "Natural Language Processing in the Legal Domain" (January 24, 2023). Available at SSRN: https://ssrn.com/abstract=4336224 or http://dx.doi.org/10.2139/ssrn.4336224.
59. For a general introduction to semantic text analysis see Hradec, J., Ostlaender, N., Macmillan, C., Acs, S., Listorti, G., Tomas, R. and Arnes Novau, X. (2019). Semantic Text Analysis Tool. SeTA: Supporting Analysts by Applying Advanced Text Mining Techniques to Large Document Collections. Luxembourg: Publications Office of the European Union. https://doi.org/10.2760/577814.
60. Glaser, I.; Matthes, F. (2020). "Classification of German Court Rulings: Detecting the Area of Law." In proceedings of the 2020 Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL).
61. Cardellino, C., Teruel, M., Alemany, L., and Villata, S. (2017). "A Low-Cost, High-Coverage Legal Named Entity Recognizer, Classifier and Linker." In ICAIL-17 Proceedings, pp. 9–18. New York: ACM. Leitner, E., Rehm, G., and Moreno-Schneider, J. (2019). "Fine-Grained Named Entity Recognition in Legal Documents." In Maribel Acosta, et al. (eds.), *Semantic Systems: The Power of AI and Knowledge Graphs*. Proceedings of the 15th International Conference (SEMANTiCS2019), no. 11702 in Lecture Notes in Computer Science, pp. 272–287. Karlsruhe, Germany, 10–11 September 2019. Springer. https://github.com/elenanereiss/Legal-Entity-Recognition.
62. See Roberts, M.E., B.M. Stewart and D. Tingley. (2019). "stm: R Package for Structural Topic Models." Journal of Statistical Software, 91(2): 1–40. Wendel, L., Shadrova, A., & Tischbirek, A. (2022). "From Modeled Topics to Areas of Law: A Comparative Analysis of Types of Proceedings in the German Federal Constitutional Court." *German Law Journal*, 23(4): 493–531. https://doi.org/10.1017/glj.2022.39.

4. Sentiment analysis, for detecting emotions in discourse (as in transcripts of witness testimonies). This technique now uses a neural network approach.

5. Word2Vec/Doc2Vec/Bag of Words, for detecting the main topic in a document, as well as the most frequent words and their correlation. This technique now uses neural network approach.

4.2.5. Network Analysis

Network analysis is a method that uses graph theory to create a map of decisions.[63] Nodes are decisions, and relationships are edges that connect nodes on the basis of a criterion (e.g., citations) and of mathematical weights (called degrees) to discover more knowledge/information (e.g., clustering of decisions within a citation). This tool offers great potential in categorising decisions that are similar or connected by some criteria, but it also carries some risks: (i) incomplete annotations can hide important case law; (ii) the links are often sensitive to changes in the law, and as a result some nodes can lose relevance over time; (iii) the connection between decisions needs to be worked out by looking at their context, and cannot come down to a matter of calculation; (iv) quantitative information (e.g., the frequency of citations) may not be as relevant with new topics and contentious issues;[64] and (v) some outcomes on appeal affect the relevance of citations with retroactive effect (e.g., decisions on constitutional issues, Supreme Court decisions). For this reason, is it important to have a clear and transparent method for correctly communicating doubts where a judicial opinion might be open to challenge.[65]

4.2.6. AI and Law [66]

For three decades now, AI and Law has been offering techniques for classifying court decisions by taking a symbolic approach (e.g., rule-based reasoning, case-based reasoning) to the emerging large language models.[67]

Machine learning is a non-symbolic AI technique based on probabilistic algorithms, and it is an important technique that emerged in the last decade in artificial intelligence in the legal domain. This family of technologies is applied to both extract and classify legal knowledge from judicial decisions.[68] However, there are some issues that we need to take into account in designing these applications:[69]

i. *Paragraph vs. structure*. Machine learning (ML), supervised or unsupervised, works at sentence level and does not take the document's structure into account. ML cannot semantically connect portions of provisions (e.g., obligation-exception, duty-penalty). For this reason, it is essential that ML be integrated with structural semantic annotation.

ii. *Text vs. context*. ML often works without additional information about a provision's context (e.g., jurisdiction, temporal parameters); this amounts to ignoring elements that are key to the legal domain (e.g., derogations depend on certain conditionals, a clear example being sunset clauses).

---

63. For a survey of the state of the art, see MOODLEY, K., HERNANDEZ-SERRANO, P., ZAVERI, A., SCHAPER, M., DUMONTIER, M., & VAN DIJCK, G. (2020). "The Case for a Linked Data Research Engine for Legal Scholars." European Journal of Risk Regulation, 11(1): 70–93. https://doi.org/10.1017/err.2019.51.
64. We can see this with the *Schrems I* and *II* cases (C-362/14 ECLI:EU:C:2015:650; C-311/18 ECLI:EU:C:2020:559), whose novelty and uniqueness means that they have few citations in other case-law, and not all in support of the opinion. So the functions served by citations is fundamental (e.g., distinguishing, supporting, overruling, pointing to an analogy). A good classification of the functions served by citations can be found in the Akoma Ntoso XML vocabulary, inspired by the Shepard classification (supports, isAnalogTo, applies, extends, restricts, derogates, contrasts, overrules, dissentsFrom, putsInQuestion, distinguishes): http://docs.oasis-open.org/legaldocml/akn-core/v1.0/akn-core-v1.0-part1-vocabulary.html.
65. See the different warning levels implemented in the LexisNexis database to mitigate risks and provide legal practitioners with transparent information: https://www.lexisnexis.com/community/insights/legal/b/product-features/posts/shepard-s-citation-guide-part-2-shepardize-your-legal-research. Also, at Wolters Kluwer Germany a search engine based on a knowledge graph was designed for German court-case data, in which regard see Junior, Ademar Crotti, Fabrizio Orlandi, Damien Graux, Murhaf Hossari, Declan O'Sullivan, Christian Hartz and Christian Dirschl. (2020). "Knowledge Graph-Based Legal Search over German Court Cases." Extended Semantic Web Conference.
66. Araszkiewicz M, Bench-Capon T, Francesconi E, Lauritsen M and Rotolo A. (2022). "Thirty Years of Artificial Intelligence and Law: Overviews. *Artificial Intelligence and Law*, 30 (4): 593–610. Monica Palmirani, Fabio Vitali, Willy Van Puymbroeck, Fernando Nubla Durango. (2022). "Legal Drafting in the Era of Artificial Intelligence and Digitisation." European Commission, Directorate-General for Informatics. https://bit.ly/EC-DIR-GEN-informatics-2022.
67. On the state of the art in symbolic and non-symbolic AI and Law, see Governatori, G., Bench-Capon, T., Verheij, B. et al. (2022). "Thirty Years of Artificial Intelligence and Law: The First Decade. Artif Intell Law 30: 481–519. https://doi.org/10.1007/s10506-022-09329-4; Sartor, G., Araszkiewicz, M., Atkinson, K. et al. (2022). "Thirty Years of Artificial Intelligence and Law: The Second Decade." Artif Intell Law 30: 521–557. https://doi.org/10.1007/s10506-022-09326-7.
68. Kevin D. Ashley. (2019). "Automatically Extracting Meaning from Legal Texts: Opportunities and Challenges." 35 Ga. St. U. L. Rev.
69. For more details, see Palmirani, Monica (2022). "A Smart Legal Order for the Digital Era: A Hybrid AI and Dialogic Model." *Ragion Pratica*, 2/2022: 633–655. https://www.rivisteweb.it/doi/10.1415/105387.

iii. *Prediction vs. relevance*. ML works mostly by applying probabilistic techniques based on data series, and if a trend becomes widespread in the legal system, it is likely to be repeated by the statistical model even if the legislation has changed. For this reason, in the legal domain, it is also very important to consider the relevance of the legal phenomenon being analysed (e.g., new legislation). This peculiar aspect should be included in the ML model using particular techniques (e.g., assigning weights to events) that have already been adopted in some industrial sectors where recent data are more important than past data.

iv. *Internal vs. external content*. ML does not consider normative and legal citations (normative cross-references) as qualified parts of a legal provision. For ML, a citation is just a sequence of characters. This makes it necessary to recall the portion of the text cited and inject it in the dataset.

v. *Static vs. dynamic content*. The content linked up by way of normative citations changes as the legal system changes over time (e.g., art. 3 will not be the same forever). ML cannot understand this semantic aspect, and for this reason we need to integrate each normative citation with the corresponding point-in-time version of the text.

vi. *Partiality vs. completeness*. Another critical aspect is the need to use a balanced and completed dataset to achieve the best representation of the reality. If a collection of decisions is incomplete or unbalanced, this problem needs to be mitigated using some technique. If the collection is mostly in the domain of commercial law, the classification algorithm is affected accordingly.

There is a whole range of subtypes of algorithms and techniques in this area. Here we will only list those that are most significant in the legal domain:

- SVM/FVM/TF-IDF[70]. Support Vector Machines use feature vectors to process documents; TF-IDF measures term importance in text analysis;
- Decision-Tree, bag of words[71]. Decision Trees organize data through a hierarchical structure of decision nodes; bag of words counts word occurrences for analysis;
- K-nearest Neighbour, which determines classification based on nearby instances in feature space;
- Neural Network[72], which models knowledge based on interconnected nodes and can be used for complex pattern recognition;
- Question-Answering[73], which process questions to generate appropriate responses;
- LLMs (large language models)[74], which process vast text data for diverse applications.

---

70. For a use-case of French Supreme Court decisions, see Sulea, O.-M., Zampieri, M., Malmasi, S., Vela, M., Dinu, L. P., and Van Genabith, J. (2017). "Exploring the Use of Text Classification in the Legal Domain." In Proceedings of the 2nd Workshop on Automated Semantic Analysis of Information in Legal Texts (ASAIL 2017).
71. See Ashley, K. (2017). *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge: Cambridge University Press.
72. Kostrzewa Ł, Nowak R. (2022). "Polish Court Ruling Classification Using Deep Neural Networks." Sensors, 2022 (6):2137. https://doi.org/10.3390/s22062137; Giulia Grundler, Piera Santin, Andrea Galassi, Federico Galli, Francesco Godano, Francesca Lagioia, Elena Palmieri, Federico Ruggeri, Giovanni Sartor, and Paolo Torroni. (2022). "Detecting Arguments in CJEU Decisions on Fiscal State Aid." In Proceedings of the 9th Workshop on Argument Mining, pp. 143–157
73. van Kuppevelt, D., & van Dijck, G. (2017). "Answering Legal Research Questions about Dutch Case Law with Network Analysis and Visualization." In A. Wyner, & G. Casini (eds.), *Legal Knowledge and Information Systems*, Vol. 302, pp. 95–100. Frontiers in Artificial Intelligence and Applications. IOS Press. https://doi.org/10.3233/978-1-61499-838-9-95; Joe Collenette, Katie Atkinson, and Trevor Bench-Capon. (2023). "Explainable AI Tools for Legal Reasoning about Cases: A Study on the European Court of Human Rights." Artif. Intell., 317, C (April 2023). https://doi.org/10.1016/j.artint.2023.103861; Francesco Sovrano, Monica Palmirani, Biagio Distefano, Salvatore Sapienza, and Fabio Vitali. (2021). "A Dataset for Evaluating Legal Question Answering on Private International Law." In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law (ICAIL 2021), 230–234. New York: Association for Computing Machinery. https://doi.org/10.1145/3462757.3466094.
74. For a survey see Kevin D. Ashley. (2022). "Prospects for Legal Analytics: Some Approaches to Extracting More Meaning from Legal Texts." 90 U. Cin. L. Rev. See also Francesca Lagioia, Jack Mumford, Daphne Odekerken, Hannes Westermann (2022). Proceedings of the 6th Workshop on Automated Semantic Analysis of Information in Legal Text Co-located with the 19th International Conference on Artificial Intelligence and Law (ICAIL 2023), Braga, Portugal, 23 September 2023. CEUR Workshop Proceedings 3441, CEUR-WS.org 2023; Classification of US Supreme Court Cases using BERT-Based Techniques; Silveira, R., Fernandes, C. G., Neto, J. A. M., Furtado, V., and Pimentel Filho, J. E. (2021). Topic Modelling of Legal Documents Via LEGAL-BERT. Proceedings. http://ceur-ws. org ISSN, 1613:00; Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., and Androutsopoulos, I. (2020). LEGAL-BERT: The Muppets Straight Out of Law School. In Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 2898–2904. Online. Association for Computational Linguistics.

4.3. Comparison Analysis

The best approach to mitigating weaknesses from specific solutions is to use multiple techniques in combination.

** (In the table below we use the Likert score ranking methodology:[75] 1 very poor, 2 poor, 3 acceptable, 4 good, 5 very good).

	Non-Crystallisation	Transparency	Explicability	Expressiveness	Accuracy	Standardisation	Multilingualism	Open Data	Cost of maintenance
Closed Vocabulary	3	5	5	2	4	5	5	5	4
Folksonomy	5	2	2	1	1	1	1	1	5
Legal Ontology	4	5	5	4	3	5	4	5	5
NLP	4	3	3	3	4	3	4	4	3
Network Analysis	3	3	3	3	4	3	4	4	3
AI symbolic	4	4	4	5	3	3	3	3	2
ML supervised (e.g, Forest Tree)	3	3	3	3	4	3	4	3	3
ML unsupervised (e.g. DNN, RNN)	4	2	2	2	4	2	3	3	3
ML semi-supervised (e.g., K-Means	4	3	3	3	4	3	4	3	4
LLM	3	2	2	2	4	2	5	2	NA

	Extraction	Classification	Representation	Explanation	Visualization
Closed Vocabulary			X	X	X
Folksonomy			X	X	X
Legal Ontology			X	X	X
NLP	X	X			
Network Analysis	X		X	X	X
AI symbolic	X	X	X	X	
ML supervised (e.g., Forest Tree)	X	X			X
ML unsupervised (e.g., DNN, RNN)	X	X			X
ML semi-supervised (e.g., K-Means	X	X			X
LLM	X	X			

75. Brown, S. (2010). "Likert Scale Examples for Surveys." Iowa State University Extension. https://www.extension.iastate.edu/Documents/ANR/LikertScaleExamplesforSurveys.pdf.

Box n. 4 - Article 5 of Convention 108+ ("Legitimacy of data processing and quality of data") emphasizes the quality and fairness of data processing, insisting on transparency and accountability. The technical tools mentioned above support this by providing mechanisms for transparent classification and processing of jurisprudential data. For instance, Controlled Vocabulary and Legal Ontology enable a standardized, transparent, and organized way of classifying legal concepts, ensuring consistency in decision-making processes. Moreover, according to Article 8 of Convention 108+ ("Transparency of Processing") data subjects should be informed about the processing of their data, such as the purpose, the logic, the consequences, and their rights. The semantic categorisation should be designed and implemented in a way that is clear and understandable for the data subjects, and that they are provided with adequate information about the processing and the results that may concern them.

5. Recommendations

Semantic categorization of court decisions is nowadays mostly automatized using different techniques that can be combined into a hybrid approach.

- Analysing the **law-and-ethics-by-design** approach to the categorization solution.
- Defining the **organizational process** to keep the categorization up to date over time, guaranteeing quality, authoritativeness, neutrality, impartiality, and transparency.
- Laying out the method, criteria, and technical approach used to categorise court decisions, while also communicating the outcome of that work, so that the categorisation is **transparent and explicable**. A user-centred approach (divided by target audience) is fundamental in offering unbiased visualization and communication.
- Making sure the **technical solution** is quantitatively and qualitatively sound. Supervised and semantic annotation are preferable in order to minimize risks.

We recommend the following ten-step methodological approach:

1. **Define the legal categorisation criteria** through a deep analysis based on the theory of law and on a comparative law methodology, while also relying on domain-specific expertise (e.g., civil procedural law). A permanent task force should be established and a legal-ethical checklist should be used to document decisions.

2. **Anonymise decisions** so as not to reveal hidden connections between judicial cases and the parties involved.

3. Define a **legal ontology** for detecting relationships between legal terms and concepts. The concepts are related to topics and conditions (e.g., jurisdiction, temporal interval, emergency, etc.). The legal ontology should be free of bias, discriminatory concepts, and prejudice.

4. Establish a **baseline of decisions** annotated relying on legal experts (supervised methodology) and do strict monitoring.

5. Analyse a robust **technological pipeline** of technical solutions to be adopted (e.g., NLP+Semantic Web+AI). A technical checklist is to be used to document decisions.

Implement the **technical solution** involving the end-user involved at every step and following the human in command principle.

6. **Test the solution** intensively with different classes of end-users, on the human in commend principle.

7. Define an **organization workflow** on which basis to periodically tweak the categorization process, considering that the work is dynamic and never crystallised. The legal, ethical, and technical checklist is to be periodically revised.

8. **Document** the details of the entire workflow and the technical procedure.

9. **Communicate and visualise the solution** using warnings and additional information, adopting a cautionary principle so as not to undermine the autonomy of the legal expert.

Glossary

Term	Definition	Synonymy
Semantic categorization	The activity that assign qualification to a text of decision or its part. The activity could be done manually or automatic, or semi-manually	Semantic classification, Semantic annotation
Decision	Judicial decision	Judgment, sentence, case law
Jurisprudence	The study and analysis of judicial decisions, precedent, and the reasoning behind court judgments, shaping legal interpretations and principles	case law, precedents
AI, Artificial Intelligence	A machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments	
RDF - Resource Description Framework	W3C standard for structuring data by expressing relationships using subject-predicate-object triples, aiding in web data interoperability	
FAIR	Findable, accessible, interoperable and re-usable	
ECLI	European Case Law Identifier - a standard code for uniquely identifying court judgments across European jurisdictions	
ELI	European Legislation Identifier - a standardized system for uniquely identifying and referencing legislative documents across the European Union	
DCAT-AP	Data Catalogue Application Profile - a standard specifying metadata for describing public sector datasets in Europe	
AK4EU	an XML-based standard, utilizing Akoma Ntoso, facilitates machine-readable legal document exchange within the EU's legislative procedure, encompassing legislative acts and proposals	
PROV-O	a W3C recommendation, providing vocabulary to represent and interchange provenance information, tracing entities and activities in the Semantic Web	
LLM	Large Language Models - AI systems with extensive neural networks, capable of analysing and processing text through vast training on diverse data sources	
Prompting	providing specific instructions or text cues to guide LLMs in generating contextually relevant and desired responses, shaping their output for various tasks or queries	
ML	Machine Learning - involves algorithms enabling systems to learn patterns from data, improving predictions or behaviours without explicit programming, crucial in AI development	Sub-symbolic AI
DNN	Deep Neural Networks - complex AI models composed of multiple connected layers, enabling advanced pattern recognition and hierarchical data abstraction for various tasks	Deep Learning
RNN	Recurrent Neural Networks - a type of artificial neural network designed to process sequential data by retaining information in loops, suitable for time-series analysis or sequential tasks	
NER	Named Entity Recognition - an NLP technique identifying and classifying named entities (names, roles, agents, etc.) within unstructured text for information extraction and organization	
	a neural network architecture in Natural Language Processing, predicts a word from its context by averaging word embeddings, facilitating efficient representation learning from surrounding words.	

Annex - Some examples

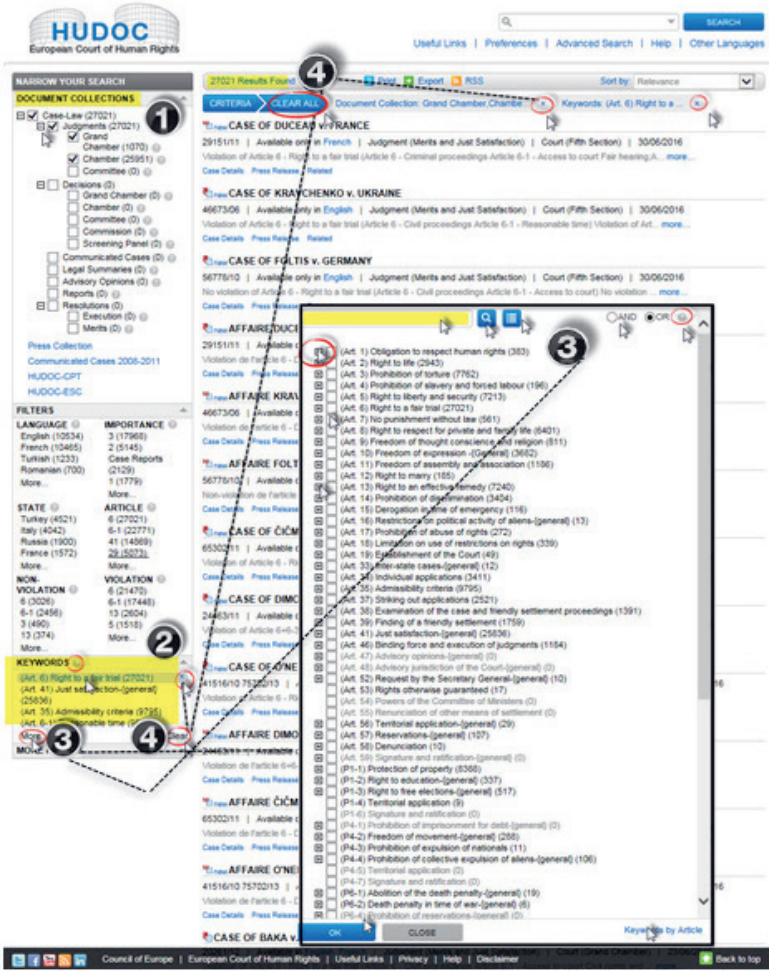


Figure 1 – HUDOC classification criteria based on a closed vocabulary.

Figure 1 - e-Justice portal with ECLI metadata classification. See example from the Netherlandish collection of decisions.

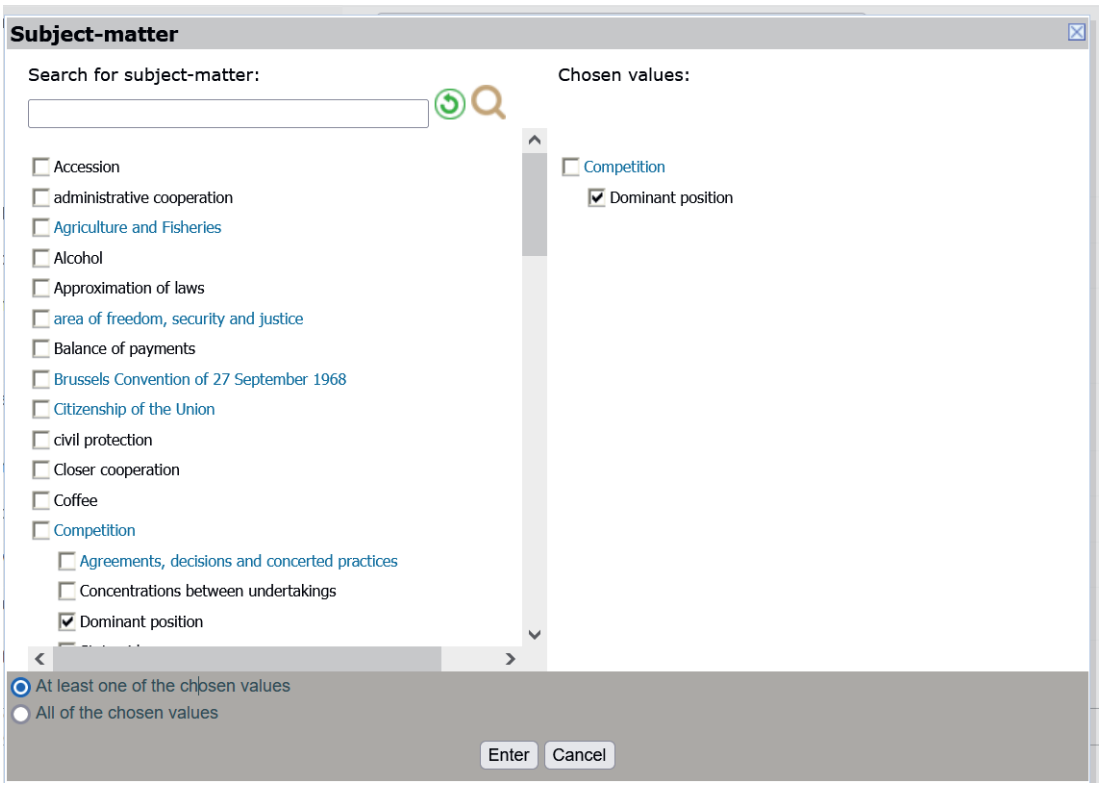


Figure 3 – CURIA classification criteria based on a closed vocabulary.

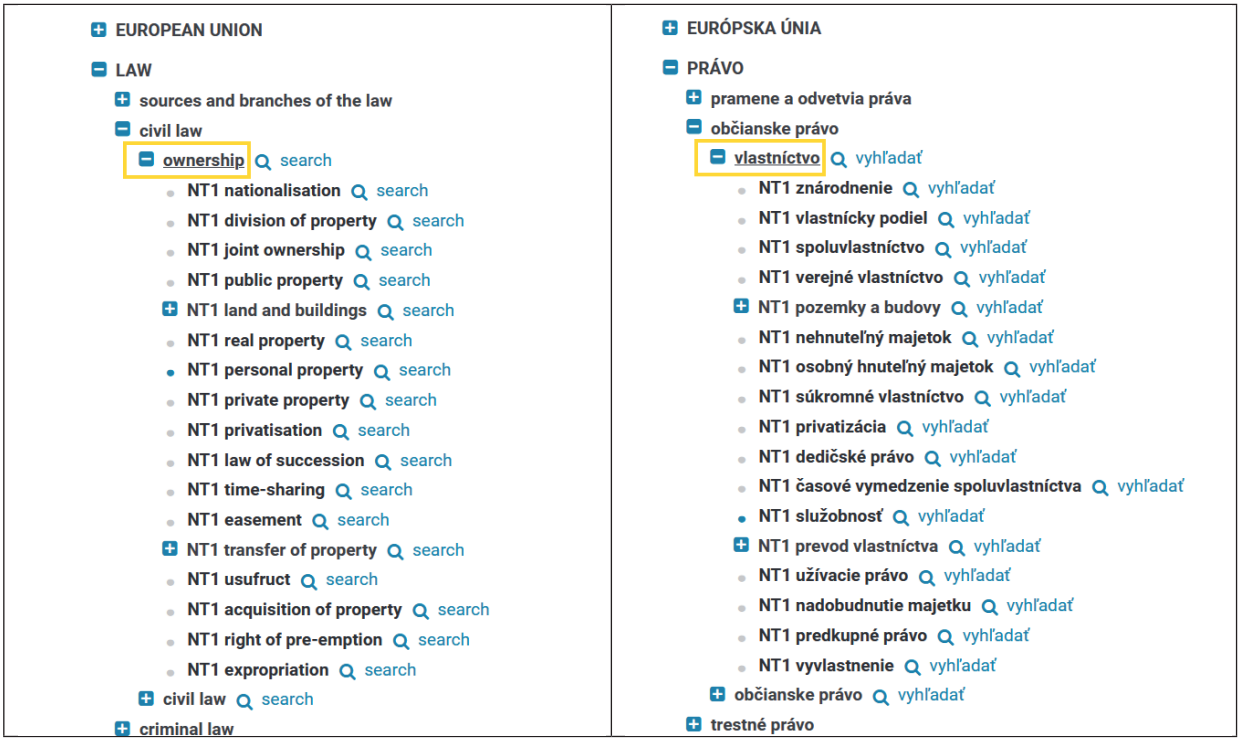


Figure 4 – EUROVOC classification criteria based on a closed multilingual vocabulary.

Figure 5 – National Archives UK Akoma Ntoso representation of decisions based on the Linked Open Data approach.

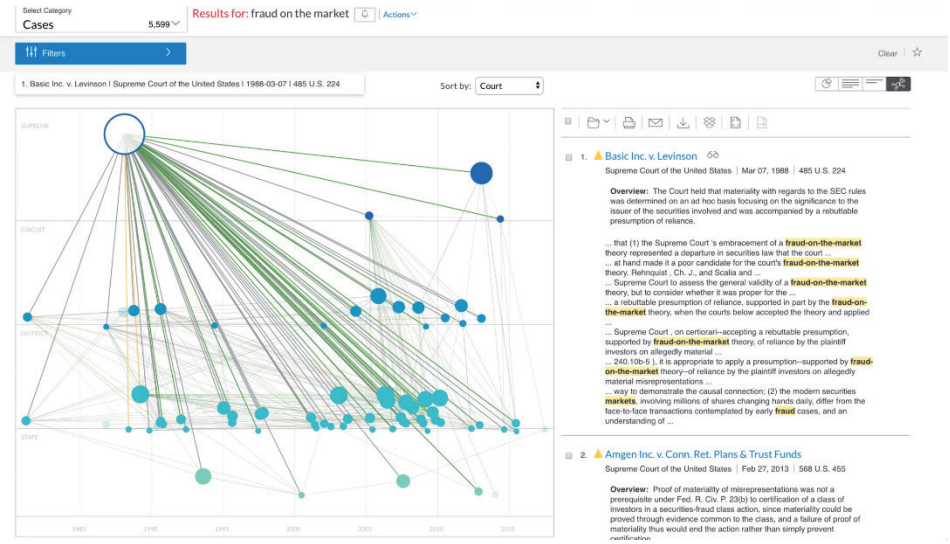


Figure 6 – LexisNexis integration of Ravel search results visualization.[76]

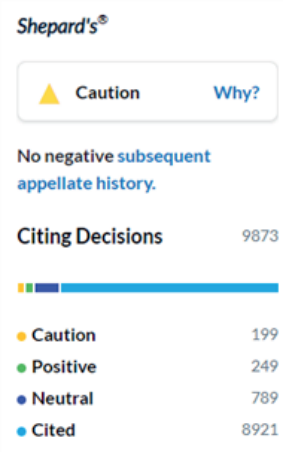


Figure 7 – LexisNexis alert mechanism.[77]

76. Ambrogi, Bob. (2018). "Lexis Advance Will Now Fully Integrate Ravel Visualizations In Search Results." *LawSites*, 28 June 2018. https://www.lawnext.com/2018/06/lexis-advance-will-now-integrate-ravel-visualizations-search-results.html.
77. ""Shepardize" Your Legal Research: Shepard's Citation Guide Part 2." LexisNexis, 23 Sept. 2022. https://www.lexisnexis.com/community/insights/legal/b/product-features/posts/shepard-s-citation-guide-part-2-shepardize-your-legal-research.

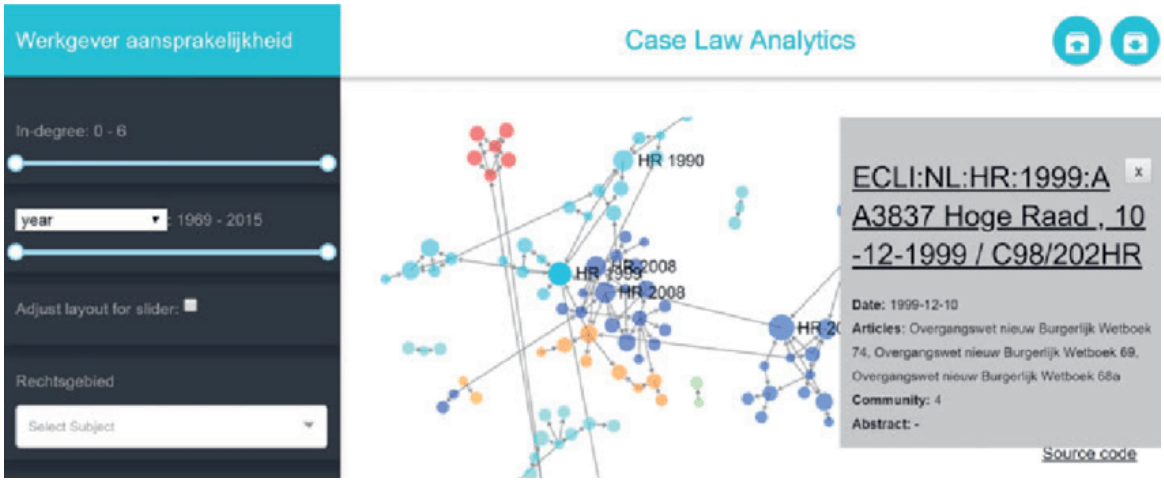


Figure 8 – Case law analytics dashboard: a demo by the University of Maastricht.[78]

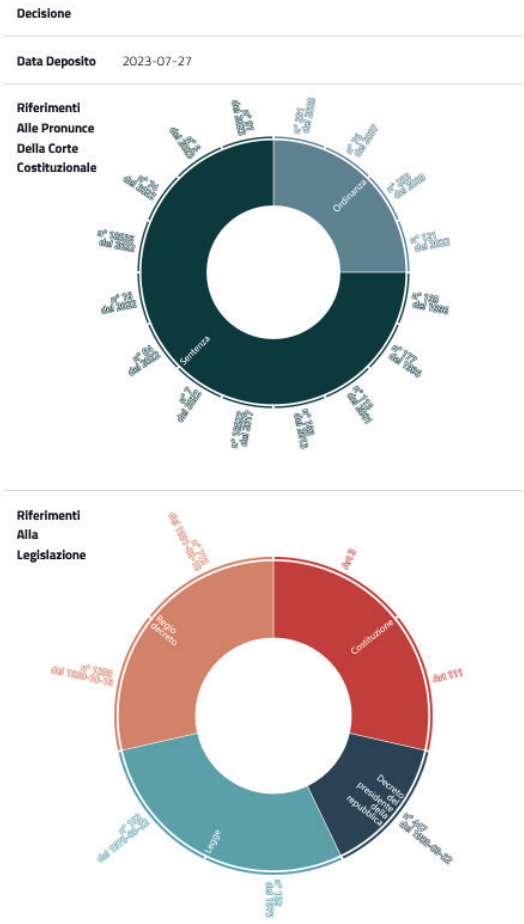


Figure 9 – Case law analytics dashboard: a demo by the University of Bologna.[79]

78. Demo at https://nlesc.github.io/case-law-app/.
79. Pronunce della Corte Costituzionale (CIRSFID): http://bach.cirsfid.unibo.it/ldms-cortecostituzionale/#/pronounce.

PREMS 007624

ENG

www.coe.int

The Council of Europe is the continent's leading human rights organisation. It comprises 46 member states, including all members of the European Union. All Council of Europe member states have signed up to the European Convention on Human Rights, a treaty designed to protect human rights, democracy and the rule of law. The European Court of Human Rights oversees the implementation of the Convention in the member states.

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE