




Analysis of a German Legal Citation Network

Tobias Milz ^a, Michael Granitzer ^b and Jelena Mitrović ^c
Chair of Data Science, University of Passau, Innstraße 41, 94032 Passau, Germany
{tobias.milz, michael.granitzer, jelena.mitrovic}@uni-passau.de

Keywords: Legal Language Processing, NLP, Network Analysis, PageRank, Neo4j, TF-IDF, Node Similarity

Abstract: The paper introduces the creation and analysis of a German legal citation network. The network consists of over 200.000 German court cases from all levels of appeal and jurisdiction and more than 50.000 laws. References to court decisions and laws are extracted from within the decision text of the court cases and added as links to the network. We apply network-based analysis techniques to support common legal information retrieval tasks such as identification of important court decisions and laws and case similarity searches. Furthermore, we demonstrate that the German case citation network displays scale-free behaviour, similar to that of the U.S. and Austrian Supreme Courts as shown by previous research.

1 INTRODUCTION


Investments into supporting technologies for the legal industry (also known as Legal Tech) have set a new record high in 2019. The market value was estimated to be at 17.32 billion U.S. dollars and is expected to reach 25 billion by 2025¹. This growing interest is also apparent in the legal research community, which, especially in recent years, has given a surge of attention to the use of AI in the various areas of technology-assisted legal research. One of these areas is concerned with the development of NLP based information retrieval systems, which has been of particular importance to this domain since the early 1960s (Widdison, 2002). This can be attributed to the ever growing amount of text-based information the law profession produces.


Research in the legal realm reveals challenges as it not only differs based on the underlying legal system (civil law, common law, customary law, etc.) but also between the countries with the same or similar legal foundation. NLP based systems have the task to adjust accordingly and investigate which studies can be translated and replicated across these different manifestations of the legal domain.


In the early nineteenth century, the use of legal citation indexes started to become an important infor-

mation retrieval tool in case law (Geist, 2009). Knowing which precedents are still being cited, or which new cases are gaining a lot of attention was crucial information - information that was otherwise difficult to obtain due to the increasing amount of new cases. This practice could be seen as an early example of citation analysis. Today, the study of citation networks is an established research area across many domains including citation behaviour analysis of scientific publications (Milz and Seifert, 2018). Consequently, legal scholars have long been theorising the benefits of using citation networks and network analysis algorithms (Neale, 2013) for legal research purposes. For example, case citation networks can reveal critical information on precedents (Cross et al., 2010) based on the characteristics of the citations towards the precedent candidates. Although this research has received some consideration from legal scholars in the past, as section 2 will outline, we believe it is still underutilised, especially for civil law systems and the Germanic law in particular.

In this publication, we describe the development and analysis of a German legal citation network. In contrast to most previous research, however, we also include citations to laws as opposed to common case-to-case citation networks. We analyse the results from a quantitative standpoint and ignore in-depth legal interpretations as this is out of the scope of this study. However, by presenting our approach and findings, we hope to create a baseline for future analysis and stir the interest of legal scholars to interpret the results.

^a  <https://orcid.org/0000-0003-3159-7666>

^b  <https://orcid.org/0000-0003-3566-5507>

^c  <https://orcid.org/0000-0003-3220-8749>

¹ <https://www.statista.com/statistics/1155852/legal-tech-market-revenue-worldwide/>

Our paper is structured as follows: In section 2 we summarise some of the related work in this field and indicate how our work can provide new insights into this area of research. In section 3 we introduce the data collection and citation extraction process before presenting the details of the resulting citation network in section 4. In section 5 and 6 we present the results of the network analysis and conclude our overall contribution.

2 RELATED WORK

As indicated in section 1 the use of case citation networks for legal information retrieval has a long history. Some of the most important studies in this area are mentioned here in chronological order.

One of the first and most extensive efforts on legal network analysis research to date was conducted by Thomas A. Smith in 2005 (Smith, 2005). In this study, he acquired a citation network containing over 4 million U.S. federal and state cases and compared its properties with the world wide web and a physics literature citation network. His work provides a comprehensive overview of use cases for legal citation network analysis. In 2007 (Fowler et al., 2007) analysed a network containing 26,681 opinions from the U.S. Supreme Court. They presented a method for identifying current and future precedent cases using a ranking system based on citation analysis. In 2009 (Geist, 2009) analysed 80,195 opinions from the Austrian Supreme Court of Justice. In his thesis, he provides evidence for the benefits of using citation counts as a metric for case relevance rankings.

Another analysis of the citation practices of the U.S Supreme Court was undertaken by (Cross et al., 2010). However, in this study, the researchers used the network to also investigate the citation behaviour of the individual Justices. In 2011 (Staffan Malmgren, 2011) analysed a citation network containing 14,327 decisions from the European Court of Justice and the General Court. He compares the precision of multiple network algorithms such as PageRank or node degree values for identifying relevant cases to a specific topic. An analysis of cross-country citation behaviour was conducted by (Gelter and Siems, 2012). They analysed the citations from decisions of the highest courts from ten different European countries. In 2013 (Neale, 2013) created a citation network by extracting citations from 594,540 Canadian court opinions. Among other findings, this study conducted a statistical analysis of citations over time and identified that most cases will not get cited anymore after 3 to 15 years with the exception of decisions from the

Supreme Court of Canada, which still get cited after 50 years.

To the best of our knowledge, one of the first larger studies on German case citation networks was conducted by (Corinna Coupette, 2019). However, the author only considered cases and citations from and to the federal constitutional court ("Bundesverfassungsgericht"). In 2020 (Katz et al., 2020) compared yearly snapshots of law citation networks of the U.S. and Germany from 1994 to 2018. In contrast to previous research, the authors concentrated on the analysis of laws and the references (citations) between them. One aggravating problem for the analysis of German legal data is the lack of freely available datasets. A problem that has been recognised by (Ostendorff et al., 2020) who in response made an effort to collect data from various sources (see section 3.1) and created the largest openly available German legal dataset to date. In their work, they also introduced a prototype of a citation network and suggested future work to investigate the network's properties. Similarly, other recent datasets have become available (Urchs et al., 2020; Urchs et al., 2021) to provide annotated legal datasets for legal argumentation identification. The most recent analysis of German legal citation networks was conducted by (Rönneburg, 2021). He created a small citation network based on about 55.000 court cases. One aspect of this study is the thematic similarity comparison between cases within automatically detected clusters. Among the two investigated clusters, the legal expert could confirm a large semantic similarity.

3 PROBLEM STATEMENT

As highlighted by the size of the datasets in section 2, the largest legal research studies have been conducted on the law systems of the U.S. and Canada. Research on specific European jurisdictions is still to receive a similar amount of attention. Furthermore, previous research on this topic has focused primarily on case-to-case citation networks from specific courts (e.g. US Supreme Court, European Court of Human Rights or Austrian Supreme Court) and case-to-law citations have not been considered at all. However, the inclusion of laws in the citation network provides a more detailed representation of the court's decision process. This can help to provide more robust network-based similarity measures (see section 5.3.1) and furthermore, allow us to identify the most important and influential laws (see section 5.2.2). Consequently, we saw the opportunity and need for a large scale, descriptive analysis of a German legal citation

network that considers courts and cases from all levels of appeal. The goal of this study is to analyse interesting properties of the network (e.g. case relevance, court relations and law importance) from a quantitative point of view (i.e. excluding legal knowledge).

3.1 DATA COLLECTION

One reason for the discrepancy between the amount of North American and European empirical legal research is the lack of openly available datasets of European court cases. In contrast, "Court Listener"² is a freely accessible dataset for U.S legal data that contains more than 3.7 million precedential court decisions. In the case of Germany, the only databases that reach that level of complexity are commercially operated (e.g. Juris GMBH³, C.H. Beck Journal⁴). Although, the Federal Ministry of Justice and Consumer Protection⁵ has published about 55,000 court cases from German federal courts and most states publish some records on their individual websites, not all of them allow users to scrape their content. As a result, until recently, there was no centralised openly available source for legal data from different jurisdictions.

This fact was recognised by (Ostendorff et al., 2020) who published an openly available dataset called "Open Legal Data"⁶ containing over 200,000 court decisions and 50,000 laws from over 1,000 different courts which they collected directly from the courts and the states and federal websites. However, as some of the data had been scraped from the individual state websites, newly published court decisions are not always added automatically. Other free legal data services such as Rewis⁷ and openJur e.V.⁸ on the other hand have either fewer cases or do not allow access to their data. Consequently, Open Legal Data represents the largest and most accessible legal dataset available at the moment and has therefore been taken as the foundation for our citation network.

3.2 CITATION EXTRACTION

In order to create the links between the nodes in the citation network, it is paramount to be able to effectively extract the citations within the decision text of the cases. Unfortunately, this task is not trivial as references to previous court decisions or laws are not

²<https://www.courtlistener.com/>; 26.04.202.

³<https://www.juris.de/>; 26.04.21

⁴<https://beck-online.beck.de/>; 26.04.21

⁵<https://www.rechtsprechung-im-internet.de/>; 26.04.21

⁶<https://de.openlegaldata.io/>; 24.04.21

⁷<https://rewis.io/>; 26.04.21

⁸<https://openjur.de/>; 26.04.21

always distinctly identifiable. Additionally, there is only one annotated dataset for German legal cases available at the moment (Leitner et al., 2020) and it is unsuitable for training or testing citation extraction methods. We will not go into too much detail about the intricacies of German law referencing but want to highlight some difficulties that we encountered.

The leading factor for the complications in identifying case-to-case citations is the lack of a well-structured unique identifier for court decisions. The introduction of the European Case Law Identifier (ECLI) in 2011 might solve this problem in the future, but older and current cases are still being referenced via a file number. Unfortunately, the file numbers are not as well structured as the ECLI and lack uniformity, which aggravates the use of automatic extraction methods.

In regards to the law-reference extraction, we simplified this task by only considering citations that are indicated by the article sign ("§"). Among those, there can be inconsistencies that lead to ambiguity, but we found that these only occur in instances where multiple laws are referenced in one citation. This reference pattern is easily recognisable by two leading article signs ("§§") instead of one. However, in our dataset, this pattern only occurs in 7.34% of all citations and a manual inspection of 100 of these patterns showed only one instance of ambiguity.

Overall we managed to extract 1,279,105 case-to-case citations (avg. 6.34 citations per court decision) and 2,234,934 case-to-law citations (avg. 11.1 citations per court decision).

4 LEGAL CITATION NETWORK

The result of the data collection and citation extraction step is our legal citation network. In contrast to previous research, we present a case-to-case citation network that also connects cases to the laws that are referenced within their decision text.

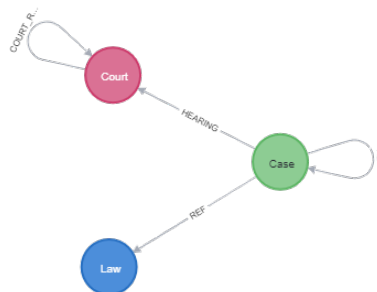


Figure 1: Neo4j graph representation of the legal citation network.

Figure 1 shows the graph representation of the network within a Neo4j database. The nodes of the graph are "Case", "Court" and "Law". Table 1 provides an overview of the most important properties of each node. Exactly one directed edge exists between two Case nodes (n) and (m) if (n) references (m) in the decision text at least once. For our analysis it is of importance that a reference occurs, but not how often. Hence, multiple references from the same decision text to the same node are disregarded. Similarly, an edge between a Case node (n) and Law node (l) is created if (l) is cited at least once within the decision text of (n). Each Case node is also connected to the corresponding Court node that made the decision. Additionally, an edge between two Court nodes (c) and (d) is created whenever an indirect citation between courts is identified. This occurs, when a Case node (n), belonging to Court (c), is referencing a Case node (m), belonging to Court (d).

Table 1: Node properties of the legal citation graph.

Node	Property	Example
Case	DecisionText	Der Antrag des Antragstellers, § 1 Abs. 5 Corona VV HE 4 im Wege der einstweiligen Anordnung...
Case	File Number	IX ZR 70/20
Case	Decision Date	25.03.2021
Law	Article	§ 242
Law	Statute	BGB
Law	Law Text	Der Schuldner ist verpflichtet, die Leistung so zu bewirken...
Court	Name	Finanzgericht Hamburg
Court	State	Hamburg
Court	Jurisdiction	Finanzgerichtsbarkeit

As we can only add an edge between nodes that exist in the database, extracted references to laws or court decisions that are not yet included in the database are ignored. The total number of nodes and edges in the database are listed in table 2. This shows that 59.9% of extracted references to laws could be added as edges, while only 16.3% of references to other court decisions are represented in the graph. Those missed edges are mainly due to the lack of data in the graph, as we made a particular effort to reduce the chances of false-positive citation identification. A copy of the database can be downloaded below⁹.

⁹<https://osf.io/8d2v4/>

Table 2: Legal citation graph statistics.

Name	Type	Amount
Case	Node	201,823
Law	Node	50,814
Court	Node	1,119
CASE_REF	Edge	208,585
LAW_REF	Edge	1,340,506
COURT_REF	Edge	7,612

5 RESULTS OF THE ANALYSIS

Network algorithms and properties are valuable tools for exploring the behaviour of an interconnected system. Even without being experts in the legal domain, we can investigate these attributes to reveal and infer information about the nature of German law.

5.1 SCALE-FREE NETWORK

Previous research has shown that the case citation networks of the U.S. Supreme Court (Smith, 2005), the Austrian Supreme Court (Geist, 2009) and the European Court of Justice (Staffan Malmgren, 2011) exhibit scale-free characteristics. The original hypothesis for this fact stems from the similarity of case citation networks to the World Wide Web. Both are directed and dynamically growing networks with large hubs that tend to follow the principle of preferential attachment (Barabasi, 2003). In the World Wide Web, this principle has the effect that as the network grows, websites with more links ("hubs") are more likely to receive new links than websites with fewer links.

Seeing the same behaviour in the German case citation network would reveal that there is a very small cluster of court decisions that hold a substantial amount of legal influence. Similar to the works above, we tested this hypothesis by inspecting the in-degree distribution (number of incoming citations) of our network. As evident by the histogram in figure 2, our network displays the same typical power-law degree distribution as other scale-free networks¹⁰. More than 70% of court decisions are not cited at all and 92.6% of cases are cited less than five times, meaning that only a small number of cases receive most of the citations. Therefore, we can assume that the German case citation network is following a scale-free behaviour. The same conclusion can be made when looking at the case-to-law in-degree histogram in figure 3. Evidently, case and law references are not equally distributed but there are in fact "hub"-like decisions and laws that are more likely to be cited.

¹⁰As shown by (Lux et al., 2007), this type of diagramming might not provide conclusive evidence but is sufficient for this estimate and the scope of this work.

Case In-Degree Histogram

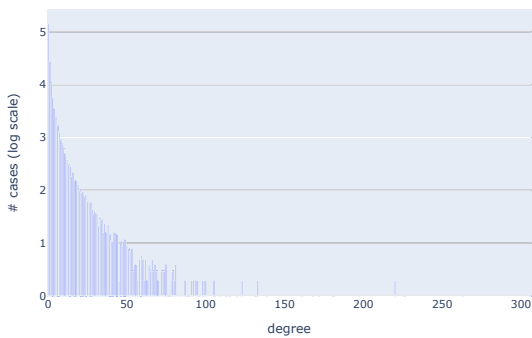


Figure 2: In-Degree distribution with log scale. The horizontal axis shows the number of incoming citations (in-degree value) and the vertical axis represents the corresponding number of court cases. This type of distribution suggests a scale-free network behaviour.

Law In-Degree Histogram

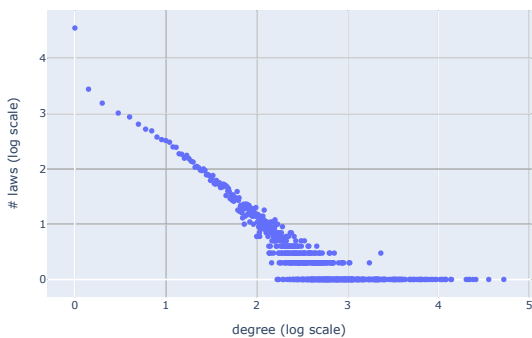


Figure 3: In-Degree distribution of laws with log scale. This scatter plot shows that most laws have not been cited at all, while a small number of laws receive most of the citations. Similar to figure 2, this type of distribution suggest a scale-free network behaviour.

5.2 CENTRALITY

One of the driving forces behind the development of legal citation indexes and networks was the desire to facilitate the search for important or influential court decisions and precedents. Using citation counts as a measure for excellence is a common practice in many areas, including scientific research. Consequently, efforts into analysing ranking scores based on different citation metrics have been made (Geist, 2009; Staffan Malmgren, 2011; Fowler et al., 2007; Rönneburg, 2021). In-degree and PageRank scores have shown to be strong indicators for identifying precedents or otherwise influential cases.

5.2.1 PAGERANK

In figure 4 we identified the twenty most important court decisions based on their overall (current) PageRank rating. The stacked bars show the change of their PageRank value over time (2005-2021). The graph indicates that most decisions become highly influential after the third year of their appearance. One exception seems to be the decision "9 AZR 44/09", which denotes its highest PageRank value in the same year of its existence.

Top 20 Court Decisions by PageRank

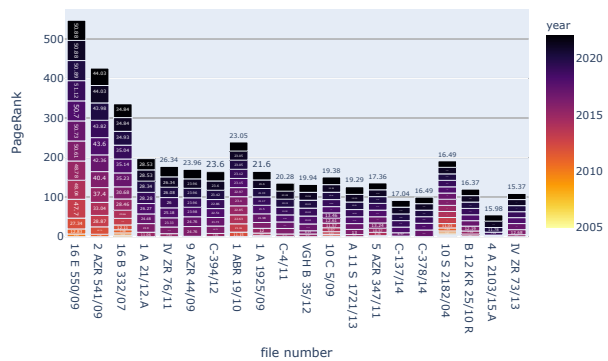


Figure 4: Stacked bar chart displaying the most influential court decisions based on their PageRank value. Each bar shows the development of the PageRank value over time.

Cases with the highest PageRank by year

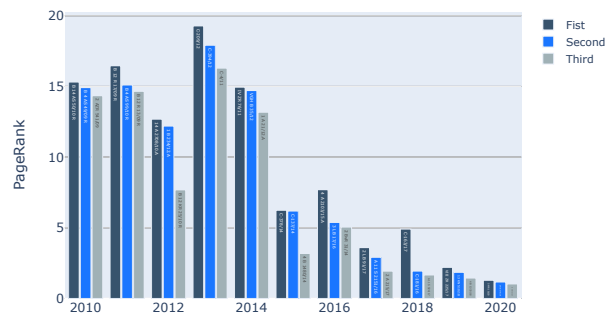


Figure 5: Top three court decisions based on PageRank by year. Older cases have a larger PageRank value as they have more chances of being cited.

Similarly, figure 5 shows the top decisions based on PageRank, but with respect to the year of the case. In this particular diagram, we focused on the top three decisions of the last ten years. This visualisation shows that in some years the most important deci-

sions were made by the European Court of Justice. For example, the case "C-209/12" which is about immigration and freedom of settlement, has the highest PageRank value of all decisions from 2013.

5.2.2 IN-DEGREE

Although (Staffan Malmgren, 2011) have shown that rankings using PageRank can be more accurate than using the in-degree value, it can still convey interesting information. In figure 6 we have identified the most cited court decisions of the last 40 years. It is noticeable that some decisions maintain their importance for consecutive years until a new "breakout" case receives more citations.

Most Cited Case by Year

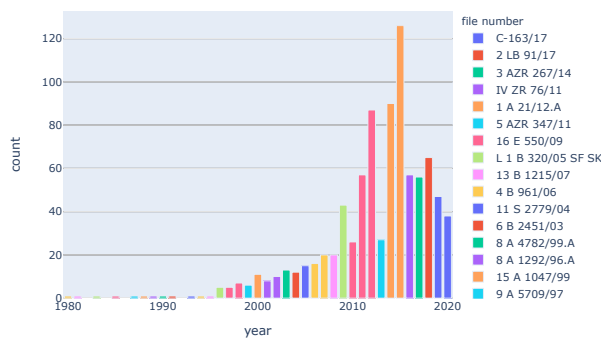


Figure 6: Court decisions with the highest amount of incoming citations per year. Some of the top cited cases seem to stay important for a consecutive year, but do not receive the same amount of attention after that.

The in-degree measure can also help to visualise the citing behaviour between courts. Figure 7 shows the most cited federal courts in Germany. The chart shows that decisions from the federal social court ("Bundessozialgericht") are cited most often, while the federal constitutional court ("Bundesverfassungsgericht") received only 3.2% of all citations.

Another avenue of utilising the in-degree measure is to find the most cited laws. In particular, it might be interesting to visualise the most important laws of each statute. As an example, figure 8 shows the most cited laws of the German civil code ("BGB").

5.3 SIMILARITY

One important task of legal research includes the discovery of related cases based on topic or field of law. Previous research into legal document similarity (Kumar et al., 2011; Wagh and Anand, 2017; Mandal

Most Cited Federal Courts

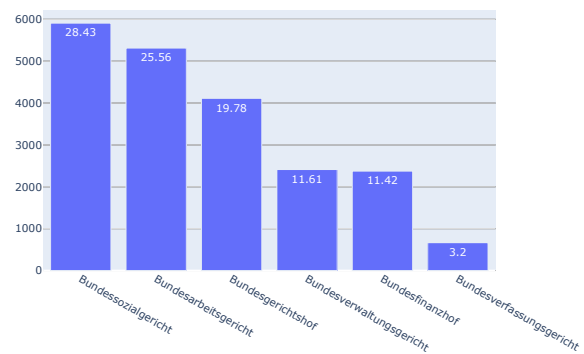


Figure 7: Among the federal courts, the social, labour and justice courts receive the highest amount of citations. The administrative and fiscal court receive each about 11% of all citations, while the constitutional court is cited least often.

Most Cited Laws (top 20) in BGB

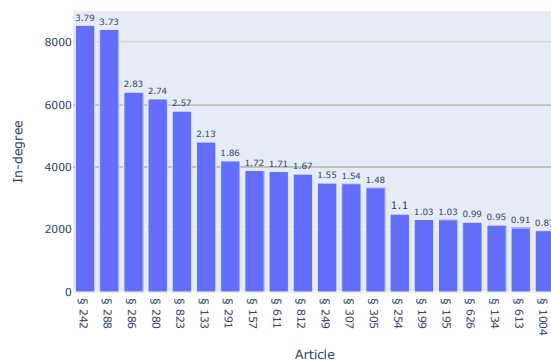


Figure 8: Most cited laws of the German civil code. Article 242 has been referenced most often receiving 3.79% of all citations towards the civil code (BGB). Article 242 states: "Performance in good faith: An obligor has a duty to perform according to the requirements of good faith, taking customary practice into consideration."

et al., 2017; Bhattacharya et al., 2020) has shown that network-based similarity measures can perform, but are also very perceptive to the sparsity of the network. However, these findings are based on case-to-case networks, the introduction of links to laws should improve the accuracy of these measures, especially for sparser case networks. However, the lack of labelled data is again aggravating the testing of this hypothesis.

5.3.1 NODE SIMILARITY

To identify similar court decisions we compute the Jaccard similarity score between all pairs of nodes within the Neo4j graph. The algorithm considers a

Table 3: Comparison between node similarity and TF-IDF based text-similarity.

Case1	Case2	Node Similarity	TF-IDF Similarity
VI-3 Kart 18/09 (V)	VI-3 Kart 17/09 (V)	1.00	1.00
VI-3 Kart 18/09 (V)	VI-3 Kart 26/09 (V)	1.00	1.00
VI-3 Kart 18/09 (V)	VI-3 Kart 27/09 (V)	1.00	0.99
VI-3 Kart 18/09 (V)	VI-3 Kart 28/09 (V)	1.00	0.99
VI-3 Kart 18/09 (V)	VI-3 Kart 29/09 (V)	1.00	0.99
W 7 M 19.30082	W 7 M 19.30083	1.00	0.96
W 7 M 19.30082	5 L 1635/14.TR	0.88	0.69
W 7 M 19.30082	17 L 1610/14.A	0.81	0.71
W 7 M 19.30082	7 L 1224/14.A	0.58	0.55
W 7 M 19.30082	3 E 187/17	0.50	0.53
L 8 R 208/05	L 8 R 361/06	1.00	0.98
L 8 R 208/05	L 8 R 44/06	0.86	0.97
L 8 R 208/05	L 8 R 47/06	0.86	0.13
L 8 R 208/05	L 8 R 62/07	0.80	0.97
L 8 R 208/05	L 3 R 98/05	0.35	0.97

pair of nodes as similar if they share the majority of their neighbours. We compare the results of the algorithm with a simplified text-based similarity calculation. For this, we determine the cosine similarity between the respective TF-IDF vectors of the decision text pairs (without stop words). Table 3 shows some examples of this comparison. As indicated by these examples, the text-based and the network-based similarity scores seem to be in agreement more often than not. In order to quantify this agreement, we calculated the Pearson correlation between the similarity measures of 1,000 semi-randomly chosen court decisions and their five most similar counterparts. To avoid disconnected or unique cases we randomly selected these samples from a list of cases that contain at least ten citations and one similar case (i.e. at least one pair of nodes with a 0.5 Jaccard similarity score). The result of 0.64 implies a positive correlation, which suggests that network-based similarity can provide an additional avenue for case-similarity search methods.

6 CONCLUSION & FUTURE WORK

In this paper, we introduced the development of a German legal citation network containing over 50,000 laws and 200,000 decisions from more than 1000 courts from all levels of appeal by using the Open Legal Data dataset (Ostendorff et al., 2020). In North America the analysis of case-to-case citation networks for legal information retrieval purposes has a much longer research history, compared to most European countries. One factor for this discrepancy orig-

inates from the lack of openly available legal data in Europe and Germany in particular. In this context, we highlighted the difference between the extensiveness of North American open legal data platforms and their German counterparts. In extension, we also underlined that the lack of annotated data aggravates the development of citation extraction approaches. However, we found that a rule-based approach using regular expressions can work sufficiently well. We stored the resulting network in a Neo4j database to enable efficient querying and the use of graph algorithms.

Using in-degree histograms, we show that the German case-to-case and case-to-law citation networks exhibit scale-free behaviour. We use the PageRank algorithm and in-degree scores to reveal the most influential court decisions, laws and courts (we have taken the federal courts as an example). Lastly, we indicate that node similarity measures have a positive correlation to text-based similarity scores.

Future research in this domain could explore the impact of citation behaviour on the outcome of a court decision. For this, it would be necessary to enrich the database with a label for the decision’s outcome. Fortunately, German court decisions are well structured and should support automatic classification efforts. Furthermore, community detection algorithms could be used to help identify laws or cases that are often cited together. Lastly, if there was more access to annotated citations from court decisions and laws, improvements could be made to the citation extraction task and more links could be added to the network. This could also include expanding the network to include links between laws.

ACKNOWLEDGEMENTS



The project on which this report is based was funded by the German Federal Ministry of Education and Research (BMBF) under the funding code 01—S20049. The author is responsible for the content of this publication.

REFERENCES

- Barabasi, A.-L. (2003). The New Science of Networks. *J. Artificial Societies and Social Simulation*, 6.
- Bhattacharya, P., Ghosh, K., Pal, A., and Ghosh, S. (2020). Methods for Computing Legal Document Similarity: A Comparative Study. In *LDA 2019 Workshop in the JURIX 2019 Conference*.
- Corinna Coupette (2019). *Juristische Netzwerkforschung*. Mohr Siebeck.
- Cross, F., Spriggs, J., Johnson, T., and Wahlbeck, P. (2010). Citations in the U.S. Supreme Court: An Empirical Study of Their Use and Significance. *University of Illinois law review*, pages 489–575.
- Fowler, J., Johnson, T., Spriggs, J., Jeon, S., and Wahlbeck, P. (2007). Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Political Analysis*, 15.
- Geist, A. (2009). Using Citation Analysis Techniques for Computer-Assisted Legal Research in Continental Jurisdictions. *SSRN Electronic Journal*.
- Gelter, M. and Siems, M. (2012). Networks, Dialogue or One-Way Traffic? An Empirical Analysis of Cross-Citations Between Ten of Europe’s Highest Courts. *Utrecht Law Review*, 8(2).
- Katz, D. M., Coupette, C., Beckedorf, J., and Hartung, D. (2020). Complex societies and the growth of the law. *Scientific Reports*, 10(1).
- Kumar, S., Reddy, P. K., Reddy, V. B., and Singh, A. (2011). Similarity Analysis of Legal Judgments. In *Proceedings of the Fourth Annual ACM Bangalore Conference*, COMPUTE ’11, New York, NY, USA. Association for Computing Machinery.
- Leitner, E., Rehm, G., and Schneider, J. M. (2020). A Dataset of German Legal Documents for Named Entity Recognition. *CoRR*, abs/2003.13016.
- Lux, M., Granitzer, M., and Kern, R. (2007). Aspects of Broad Folksonomies. In *18th International Conference on Database and Expert Systems Applications (DEXA 2007)*. IEEE.
- Mandal, A., Chaki, R., Saha, S., Ghosh, K., Pal, A., and Ghosh, S. (2017). Measuring Similarity among Legal Court Case Documents. In *Proceedings of the 10th Annual ACM India Compute Conference on ZZZ - Compute ’17*, New York, New York, USA. ACM Press.
- Milz, T. and Seifert, C. (2018). Analysing Author Self-citations in Computer Science Publications. In *Database and Expert Systems Applications*, pages 289–300, Cham. Springer International Publishing.
- Neale, T. (2013). Citation Analysis of Canadian Case Law. *Journal of Open Access to Law*, vol. 1, no. 1, pp. 1–51.
- Ostendorff, M., Blume, T., and Ostendorff, S. (2020). Towards an Open Platform for Legal Information. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020*.
- Rönneburg, L. (2021). Analyse eines juristischen Entscheidungscorpus mit Methoden der Netzwerkforschung und Sprachtechnologie. *University of Hamburg, Bachelor Thesis*.
- Smith, T. A. (2005). The Web of Law. *SSRN Electronic Journal*.
- Staffan Malmgren (2011). Towards a theory of jurisprudential relevance ranking Using link analysis on EU case law. *Stockholm University, Master Thesis*.
- Urchs, S., Mitrovic, J., and Granitzer, M. (2020). Towards Classifying Parts of German Legal Writing Styles in German Legal Judgments. In *2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*. IEEE.
- Urchs, S., Mitrović, J., and Granitzer, M. (2021). Design and Implementation of German Legal Decision Corpora. In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence*. SCITEPRESS - Science and Technology Publications.
- Wagh, R. and Anand, D. (2017). Application of citation network analysis for improved similarity index estimation of legal case documents : A study. In *2017 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC)*, pages 1–5.
- Widdison, R. (2002). New Perspectives in Legal Information Retrieval. *International Journal of Law and Information Technology*, 10(1):41–70.