

Making legal knowledge accessible to machines: challenges and opportunities



Sabrina Kirrane

Joint work with:

Erwin Filtz, María Navas-Loro, Axel Polleres, Rodríguez-Doncel, and Cristiana Santos,

Semantic Technologies for Scientific, Technical and Legal Data @ ESWC 2023



'The Godfather' Leaves Ge Warns of Da

For half a century, Geoff
 technology at the heart of
 Now he worries it will

Daniel Feldman
 @d_feldman

A lawyer used ChatGPT to do "legal research" and cited a number of nonexistent cases in a filing, and is now in a lot of trouble with the judge

6. As the use of generative artificial intelligence has evolved within law firms, your affiant consulted the artificial intelligence website Chat GPT in order to supplement the legal research performed.

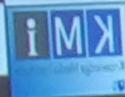
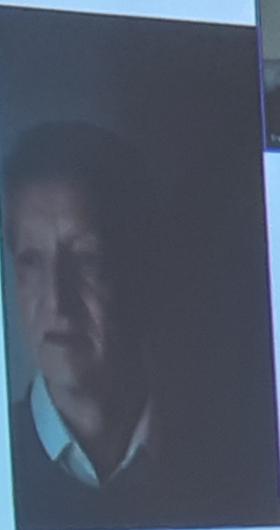
7. It was in consultation with the generative artificial intelligence website Chat GPT, that your affiant did locate and cite the following cases in the affidavium in opposition submitted, which this Court has found to be nonexistent.

Case 1:22-cv-01461-PKC Document 32-1 Filed 05/25/23 Page 2 of 6

Varghese v. Chive Southern Airlines Co Ltd, 905 F.3d 1339 (11th Cir. 2019)
 Shallosa v. Equinox, 2013 U.S. App. (9th Cir. 11/27/13) 11th App. Cir. 2013
 Pateros v. Jeter, No. 193 F. Supp. 3d 121 (D.C. 2015)
 Martinez v. Delta Airlines, Inc., 2019 WL 4039462 (Tex. App. Sept. 25, 2019)
 Estate of Dutton v. S.M. Royal Dutch Airlines, 2017 WL 2479823 (Okla. Ct. App. June 9, 2017)
 Miller v. United Airlines, Inc., 174 F.3d 368 (2d Cir. 1999)

6:02 AM · May 27, 2023 · 5M Views

9,301 Retweets · 1,341 Quotes · 30.9K Likes · 2,530 Bookmarks



European Legal Data Landscape

Legal data across Europe

- Legal databases
 - For different jurisdictions
 - Have different access policies
 - Run by governments and non-governmental organizations
- Automated data processing is challenging
 - Missing/ incomplete and heterogeneous metadata
 - Access policies, file formats, etc.



European Legal Data Landscape

Legal data across Europe

- European Law Identifier (ELI), European Case Law Identifier (ECLI)
 - Proposed by the Council of the European Union ~10 years ago
 - Standardized identifiers for legislative (ELI) and judiciary (ECLI) documents

<http://data.europa.eu/eli/dir/2014/92/oj>

<https://www.ris.bka.gv.at/eli/bqbl/I/2016/35/20160608>

 - Minimum set of metadata for legal documents (publication date, title,...)
 - Metadata is published in a semantic web format to be reused
- Additional data
 - Named Authority Lists (language, country codes,...)
 - Thesauri



Legal Data in Austria

The Starting Point

- Central point for legal information
 - Federal/ state law (law gazettes, consolidated versions)
 - Judicial branch
 - All documents publicly available
- ELI and ECLI partially implemented
- Documents only partially interlinked
- Keyword based search process with additional filters
- RIS API



RECHTSINFORMATIONSSYSTEM DES BUNDES RIS

[Bundesrecht] Landesrecht Bezirke Gemeinden Judikatur Kundmachungen, Erlässe Gesamtabfrage

Bundesrecht konsolidiert

Suchworte

Titel, Abkürzung

Paragraf von bis

Artikel von bis

Anlage von bis

Kundmachungsorgan

Typ

Index

Unterzeichnungsdatum

Suche nach Fassung

Fassung vom

Suche nach zeitlichem Geltungsbereich

Inkrafttretensdatum von bis

Außerkrafttretensdatum von bis

Neu/geändert im RIS seit

Suche starten Zurücksetzen

Webseiten

- Bundeskanzleramt
- Bundesministerium für Digitalisierung und Wirtschaftsstandort
- Oesterreich.gv.at
- Parlamentarische Materialien
- Fundstellennachweis zu den Materialien
- SozDok – Dokumentation des SV-Rechts
- EU-Recht

Informationen

- Coronavirus - Rechtliches BMISGPK
- COVID-19 - Ausgewählte Rechtsnormen
- Index des Bundesrechts
- Stand der Aktualität
- Abkürzungsverzeichnis

Enhancements: Modelling The AustroVoc Thesaurus

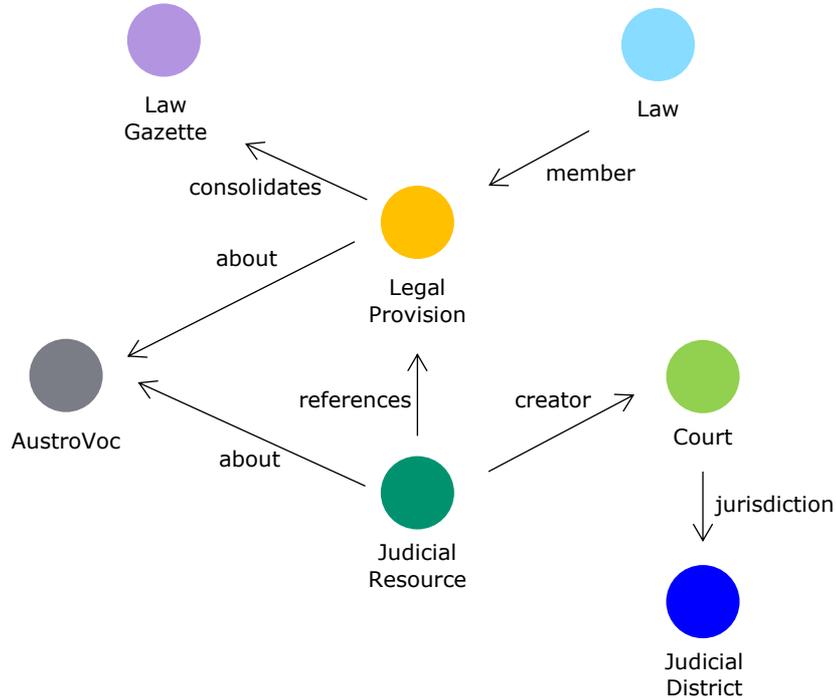
Bundesrechtindex

- **Bundesrechtindex**
 - **Verfassungs- und Organisationsrecht, Medienrecht, allgemeine international Angelegenheiten**
 - **+** Verfassungsrecht
 - **+** Hoheitsgebiet, Grenzen
 - **+** Internationale Angelegenheiten
 - **+** Staatsvertragsdurchführung, Kriegsfolgen
 - **+** Organisationsrecht
 - Unabhängigkeitserklärung, Rechtsüberleitung, Übergangsrecht, Rechtsbereinigung
 - **+** Medienrecht
 - Vereinbarungen gemäß Art. 15a B-VG
 - Kundmachungswesen
 - **+** Völkerrechtliche Verträge
 - **Zivil- und Strafrecht**
 - **-** Privatrecht allgemein
 - Allgemeines bürgerliches Gesetzbuch (ABGB)
 - Familienrecht
 - Erwachsenenschutz
 - Erbrecht einschließlich Anerbenrecht
 - Wohn- und Mietrecht
 - Konsumentenschutz
 - Schadenersatz, Haftpflicht
 - Urheberrecht
 - Internationales Privatrecht
 - Priester, Ordensleute
 - Grundbuch
 - Urkunden
 - Sonstiges
 - **+** Handels- und Wertpapierrecht
 - **+** Zivilprozess, Außerstreitiges Verfahren
 - **+** Insolvenzrecht, Exekutionsrecht
 - **-** Strafrecht
 - Strafgesetzbuch
 - Jugendgerichtsbarkeit

➤ AustroVoc thesaurus

- Austrian specific legal terms and entities
- Law index turned into a thesaurus
- Laws and court decisions reference law index
- Resource types
- Court types

Enhancements: Modelling A Legal Knowledge Graph



➤ Legal Knowledge Graph

- Model to represent Austrian legal information

➤ New classes

- Law Gazette
- Law
- Legal Provision
- Judicial Resource
- Court
- Judicial District

Enhancements: Population Legal Entity Extraction



1. Gemäß [§ 18 Abs 1 erster Satz BStG 1971](#) gebührt dem Enteigneten für alle durch die Enteignung verursachten vermögensrechtlichen Nachteile Schadloshaltung ([§ 1323 ABGB](#)). Die dem Enteigneten gebührende Entschädigung muss alle durch die Enteignung verursachten vermögensrechtlichen Nachteile erfassen, wobei der Verkehrswert der entzogenen Liegenschaft den wichtigsten Faktor für dessen Bemessung darstellt ([6 Ob 161/10k mwN](#)). Auch eine nachträgliche Entwicklung der tatsächlichen Verhältnisse, die im Zeitpunkt der Enteignung als wahrscheinlich vorausgesehen werden konnte, kann die Höhe des zu ersetzenden Verkehrswerts (hier: nach [§ 18 BStG](#)) beeinflussen ([RIS-Justiz RS0053483](#)). Für die Bewertung eines Grundstücks sind nach ständiger Rechtsprechung des [Obersten Gerichtshofs](#) daher neben der bestehenden Widmung auch realistisch beurteilte künftige Verwendungsmöglichkeiten samt ihrer Auswirkung auf den Marktwert entscheidend, sofern die reale Möglichkeit einer solchen Verwendung bereits im Zeitpunkt der Enteignung gegeben war und nicht bloß für eine noch unbestimmte Zukunft erhofft worden ist ([RIS Justiz RS0058043 \[T3\]](#); [RS0057977](#); [RS0057981](#); zuletzt [3 Ob 46/11b](#); [1 Ob 201/13k](#) je mwN [krit zur Einstufung als Bauerwartungsland bei bloßer Möglichkeit der Umwidmung: [Kerschner](#), [Bauerwartungsland](#), [ZfB 2013/38](#)]).

- Corpus of 50 Austrian Supreme Court decision texts
 - Manually annotated
 - 7 different entities
 - Reference
 - Person
 - Court

- Comparative analysis of entity extraction approaches
 - Rules
 - Conditional Random Fields (CRF)
 - Flair
 - BERT
 - DistilBERT

Enhancements: Population Legal Entity Extraction

„[...] zugunsten des obsiegenden Klägers (RIS-Justiz **RS0079624** [T14]). Ein berechtigtes Interesse des obsiegenden Beklagten an der Urteilsveröffentlichung ist dann gegeben, wenn der Rechtsstreit eine gewisse Publizität erlangt hat (RIS-Justiz **RS0079511**), etwa wenn [...]“

RS [0-9] * → **LegalRule**

„[...] vgl. **Mayrhofer/Tangl in Fenyves/Kerschner/Vonkilch, Klang3 § 6 Abs 1 Z 2 KSchG Rz 1** [...]“

- Rules
 - Different rules for all entities
 - Mainly regular expression based
- Conditional Random Fields (CRF)
 - Features of a token,
 - For instance position and casing, are used to calculate the probabilities of tokens following each other
- Machine Learning
 - Flair
 - Contextualized character level embeddings
 - BERT
 - Transformer based architecture
 - DistilBERT
 - A teacher-student setting

Enhancements: Population

Legal Entity Extraction

- 50 Austrian Supreme Court decisions

„[...] zugunsten des obsiegenden Klägers (RIS-Justiz **RS0079624** [T14]). Ein berechtigtes Interesse des obsiegenden Beklagten an der Urteilsveröffentlichung ist dann gegeben, wenn der Rechtsstreit eine gewisse Publizität erlangt hat (RIS-Justiz **RS0079511**), etwa wenn [...]“

„[...] vgl. **Mayrhofer/Tangl in Fenyves/Kerschner/Vonkilch, Klang3 § 6 Abs 1 Z 2 KSchG Rz 1 [...]**“

	Approach	Case Reference	Contributor	Court	Legal Provision	Law Gazette	Legal Rule	Literature	
F-Scores	Rules	Rules (lenient)	0.9824	0.8426	0.9801	0.9090	0.9460	1	0.8674
	ML / DL	CRF	0.9787	0.9328	0.9616	0.9459	0.9473	0.9997	0.8866
		BERT	0.9712	0.9583	0.9616	0.9489	0.9396	0.9986	0.8448
		DistilBERT	0.9772	0.9551	0.9586	0.9521	0.9437	0.9989	0.8626
	Δ		0.0112	0.1157	0.0215	0.0431	0.0077	0.0014	0.0418

Enhancements: Population

Legal Event Extraction

THE FACTS

I. THE CIRCUMSTANCES OF THE CASE

5. The applicant was born in 1965 and lives in Szajol.

6. On 6 October 1990 he married Ms N.R. The couple had one daughter, R.Ó, born on 23 March 1994.

7. On 24 March 2006 the couple divorced and agreed on custody of the child and other parental rights. Their agreement, which was approved by the Szolnok District Court, placed the daughter with her mother and granted the applicant contact every other weekend, during the school holidays, Easter Sunday, Whit Monday, the first holiday following Christmas Eve and three weeks during the summer holidays. The applicant was also to pay child allowance, amounting to 20% of his salary, but at least 20,000 Hungarian forints (HUF approximately 66 euros (EUR)) per month.

8. Ms N.R. moved out of the former family home and took up residence in Budapest.

9. The applicant could exercise contact on 25 and 26 June and 24 July 2006. However, his daughter was reluctant to stay with him during the summer holidays and preferred to spend her holidays with her grandparents.

10. The applicant's further attempts to exercise contact failed, seemingly because in the mother's view it was up to the child to decide whether she wanted to see her father. The applicant applied to the Szolnok County Guardianship Authority to have the contact agreement enforced. In a decision of 11 September 2006 it ordered Ms N.R. to comply with the agreement, warning her that failure to do so could result in her being fined and ordered to reimburse the applicant's travel costs. Given the strained relationship between the parents, a notary initiated child protection proceedings (védelembe vétel) on 27 December 2006. Subsequently, on a number of occasions Ms N.R. was fined for not complying with the agreement. It appears from the case file that the applicant could exercise contact between 10 July 2007 and 10 November 2008.

11. In 2007 Ms N.R. brought an action seeking to change the applicant's contact rights and an increase in the amount of child allowance. In a counterclaim the applicant requested that the child be placed in his custody.

12. The applicant also filed a criminal complaint against Ms N.R. alleging child endangerment. The parties' attempt to stabilise the relationship between the applicant and his daughter during the suspension of the custody proceedings were unsuccessful and the Szolnok District Court ordered that any meetings between them take place at a child protection centre.

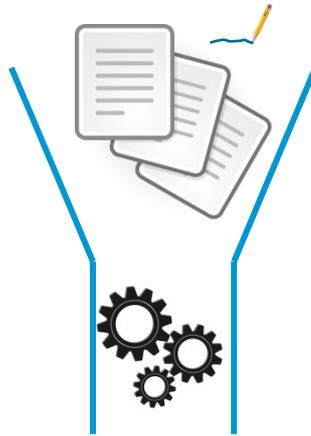
13. In a judgment of 10 June 2008 the Szolnok District Court reduced the applicant's contact with his daughter to every first and third Saturday of the month from 9 a.m. to 6 p.m. It held that the previously agreed form of contact was unlikely to be implemented and would only lead to further proceedings before the guardianship authority, which would be to the detriment of the child. The court based its judgment on an expert opinion finding that the child should have had an adaptation period to re-establish her relationship with her father. The court dismissed the applicant's claim concerning custody, stating that the child's wishes had to be taken into account, given that she was now fourteen years of age.

14. On 19 November 2008 the Jász-Nagykun-Szolnok County Regional Court upheld the first-instance judgment in essence but amended the applicant's contact rights to every other Saturday between 9 a.m. and 4 p.m. until 31 May 2009, and all weekend-long visits every other weekend as of 1 June 2009.

15. In 2009 the applicant failed to turn up at numerous scheduled meetings for months, for which he was fined HUF 10,000 (approximately EUR 35).]

- Court decisions are long
- „What happened when?“
- Quick overview over case
- Components
 - Dates („when“)
 - Subjects („who“)
 - Core („what“)

Enhancements: Population Legal Event Extraction



- 1) Comparative analysis of temporal taggers
 - TempCourt corpus
 - 30 documents of ECHR, ECJ and USSC
- 2) Comparative analysis of event extraction approaches
 - EventsMatter corpus
 - 30 documents of ECHR

1965	The applicant was born.
June 1999	The applicant refused to see his son.
15 January 2000	The District Court upheld the judgment.

TempCourt: Evaluation of Temporal Taggers on a new Corpus of Court Decisions. The Knowledge Engineering Review, 2019.

Events matter: Extraction of events from court decisions. JURIX, 2020.

Enhancements: Population

Legal Event Extraction

- EventsMatter corpus: 30 court decisions of ECHR

„The applicant **refused to see his son** in June 1999.“

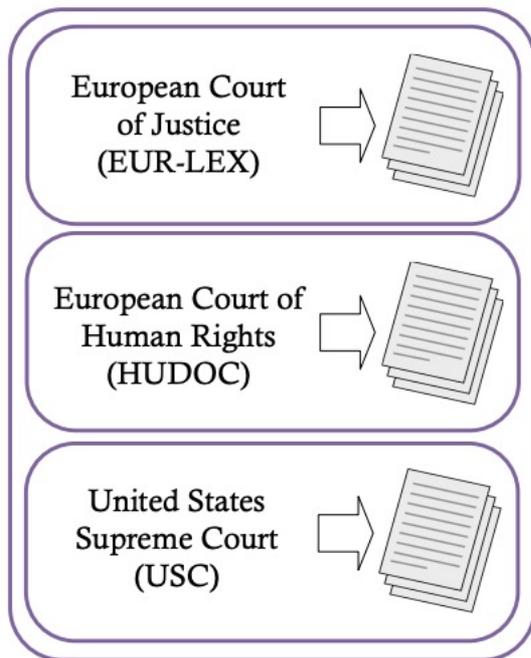
„On 15 January 2000 the District Court **upheld the Judgment.** “

		Approach	What	When	Who
F-Scores	Rules	Rules (lenient)	0.2195	0.6977	0.6857
	ML/DL	CRF	0.8050	0.8658	0.7834
		BERT	0.6583	0.9022	0.9044
		DistilBERT	0.6237	0.8823	0.8998
	Δ		0.5855	0.2045	0.2187

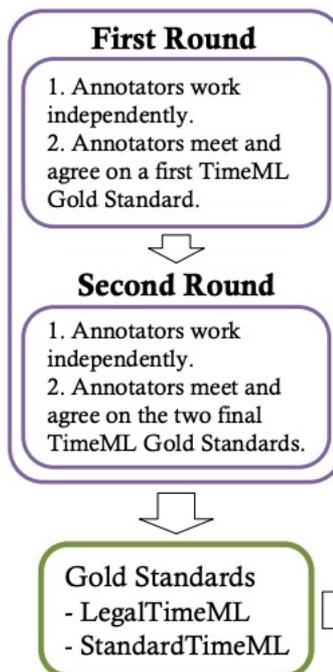
Enhancements: Population

Legal Temporal Entity Extraction

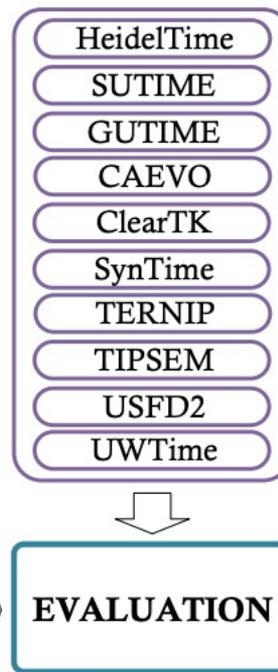
DOCUMENT COLLECTION



ANNOTATION



TAGGING



Enhancements: Population Legal Temporal Entity Extraction

Table 1 Structure of ECJ, ECHR and USC decisions.

Section	ECJ	ECHR	USC
A	Involved parties	Involved parties	Involved parties
B	Case summary	Procedure	Syllabus
C	Legal framework	Circumstances of the case	Main Opinion
D	Circumstances of the case	Legal framework	Concurring and dissenting opinions
E	Court assessment	Court assessment	
F	Judgment	Judgment	

Table 5 Overview of temporal taggers. (*) Not all the types are covered.

Temporal Tagger	Approach	Identification	Normalization	Events	Relations
HeidelTime (HE)	rule-based	✓	✓	-	-
SUTime (SU)	rule-based	✓	✓	-	-
GUTime (GU)	hybrid	✓	✓	✓	✓
CAEVO (CA)	hybrid	✓	✓	✓	✓
ClearTK (CL)	machine-learning	✓	-	✓	✓
SynTime (SY)	rule-based	✓	-	-	-
TERNIP (TE)	rule-based	✓	✓	-	-
TIPSem (TI)	hybrid	✓	✓	✓	✓
USFD2 (US)	hybrid	*	*	-	*
UWTime (UW)	hybrid	✓	✓	-	-

Enhancements: Population

Legal Temporal Entity Extraction

Table 3 Corpus statistics

Corpus	# Doc.	# Tokens	Doc. Size (Avg. KB)	Doc. Size (Avg. Tokens)	Sentence length (Avg. Tokens)
ECHR	10	7,252	4	725	13
ECJ	10	53,044	32	5,304	32
USC	10	50,874	25	5,087	18
Total	30	111,170	20	3,705	21

Table 4 Statistics of corpora annotated with TimeML in literature.

Corpus	# Doc.	# Tokens	Doc. Size (Avg. Tokens)
TimeBank ²⁰	183	78,444 (61,000 ²¹)	428.7
AQUAINT ²²	73	34,154	467.9
TempEval-3 Eval. (<i>Platinum</i>) (47)	20	~6,000 ²³	~300
WikiWars (42)	22	119,468	5,430.4
Time4SMS (42)	1,000	20,176	20.2
Time4SCI (42)	50	19,194	383.9

Enhancements: Population

Legal Temporal Entity Extraction

Table 7 Evaluation results for the ECHR corpus for each temporal tagger, both for identification (two first columns, *lenient* and *strict*) and normalization (two last columns, *lenient* and *strict*). The first row (in white) corresponds to results against the *StandardTimeML* gold standard, while the second (in gray) corresponds to the *LegalTimeML* gold standard.

A	lenient			strict			lenient+ value			strict+ value		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
HE	0.99	0.99	0.99	0.84	0.84	0.84	0.78	0.78	0.78	0.78	0.78	0.78
	0.88	0.99	0.93	0.71	0.80	0.75	0.67	0.75	0.71	0.64	0.72	0.68
SU	0.88	0.87	0.88	0.85	0.84	0.84	0.78	0.78	0.78	0.76	0.75	0.75
	0.76	0.85	0.80	0.71	0.80	0.76	0.66	0.74	0.79	0.64	0.72	0.68
GU	0.96	0.93	0.94	0.95	0.92	0.93	0.86	0.84	0.85	0.86	0.84	0.85
	0.84	0.92	0.88	0.83	0.92	0.87	0.74	0.82	0.78	0.74	0.82	0.78
CA	0.88	0.87	0.87	0.83	0.82	0.82	0.78	0.78	0.78	0.75	0.75	0.75
	0.75	0.85	0.80	0.70	0.79	0.74	0.65	0.74	0.69	0.64	0.72	0.67
CL	0.92	0.78	0.85	0.34	0.32	0.35	-	-	-	-	-	-
	0.80	0.77	0.78	0.33	0.32	0.33	-	-	-	-	-	-
SY	0.98	0.93	0.96	0.83	0.79	0.81	0	0	0	0	0	0
	0.86	0.93	0.90	0.70	0.76	0.73	0	0	0	0	0	0
TE	0.94	0.95	0.95	0.92	0.93	0.92	0.86	0.88	0.87	0.85	0.86	0.85
	0.83	0.95	0.89	0.80	0.92	0.85	0.75	0.86	0.80	0.72	0.83	0.77
TI	0.78	0.85	0.81	0.64	0.70	0.67	0.64	0.71	0.67	0.63	0.69	0.66
	0.69	0.86	0.76	0.62	0.77	0.69	0.64	0.79	0.71	0.61	0.76	0.68
US	0.73	0.61	0.67	0.69	0.58	0.63	0	0	0	0	0	0
	0.65	0.62	0.64	0.61	0.58	0.60	0	0	0	0	0	0
UW	0.90	0.53	0.67	0.51	0.30	0.38	0.55	0.33	0.41	0.51	0.30	0.38
	0.86	0.58	0.69	0.48	0.32	0.38	0.51	0.34	0.41	0.48	0.32	0.38

Temporal Tagger

HeidelTime (HE)
SUTime (SU)
GUTime (GU)
CAEVO (CA)
ClearTK (CL)
SynTime (SY)
TERNIP (TE)
TIPSem (TI)
USFD2 (US)
UWTime (UW)

Enhancements: Population

Legal Temporal Entity Extraction

Table 8 Evaluation results for the ECJ corpus for each temporal tagger, both for identification (two first columns, *lenient* and *strict*) and normalization (two last columns, *lenient* and *strict*). The first row (in white) corresponds to results against the *StandardTimeML* gold standard, while the second (in gray) corresponds to the *LegalTimeML* gold standard.

A	lenient			strict			lenient+ value			strict+ value		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
HE	0.48	0.95	0.64	0.47	0.94	0.63	0.47	0.94	0.62	0.47	0.93	0.62
	0.27	0.97	0.42	0.26	0.96	0.41	0.26	0.94	0.40	0.26	0.93	0.40
SU	0.81	0.97	0.88	0.79	0.95	0.86	0.78	0.93	0.85	0.77	0.92	0.84
	0.44	0.95	0.60	0.43	0.93	0.58	0.41	0.90	0.57	0.41	0.89	0.56
GU	0.97	0.87	0.91	0.97	0.86	0.91	0.94	0.84	0.89	0.94	0.84	0.88
	0.51	0.82	0.63	0.50	0.82	0.62	0.48	0.78	0.60	0.48	0.78	0.60
CA	0.89	0.74	0.81	0.85	0.70	0.77	0.86	0.71	0.77	0.85	0.70	0.77
	0.49	0.74	0.59	0.46	0.70	0.56	0.46	0.70	0.56	0.46	0.69	0.55
CL	0.77	0.88	0.82	0.32	0.36	0.34	-	-	-	-	-	-
	0.42	0.88	0.57	0.18	0.37	0.24	-	-	-	-	-	-
SY	0.89	0.99	0.93	0.81	0.90	0.85	0	0	0	0	0	0
	0.49	0.98	0.65	0.46	0.92	0.61	0	0	0	0	0	0
TE	0.97	0.88	0.92	0.96	0.88	0.91	0.96	0.87	0.91	0.95	0.87	0.91
	0.54	0.89	0.67	0.53	0.88	0.66	0.53	0.88	0.65	0.52	0.87	0.65
TI	0.72	0.81	0.76	0.64	0.72	0.68	0.62	0.70	0.65	0.61	0.69	0.65
	0.41	0.83	0.54	0.37	0.75	0.49	0.35	0.71	0.47	0.34	0.70	0.46
US	0.31	0.54	0.39	0.29	0.51	0.37	0.02	0.04	0.03	0.02	0.03	0.02
	0.20	0.65	0.31	0.19	0.61	0.29	0.02	0.06	0.03	0.02	0.05	0.02
UW	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-

Temporal Tagger

HeidelTime (HE)
SUTime (SU)
GUTime (GU)
CAEVO (CA)
ClearTK (CL)
SynTime (SY)
TERNIP (TE)
TIPSem (TI)
USFD2 (US)
UWTime (UW)

Enhancements: Population

Legal Temporal Entity Extraction

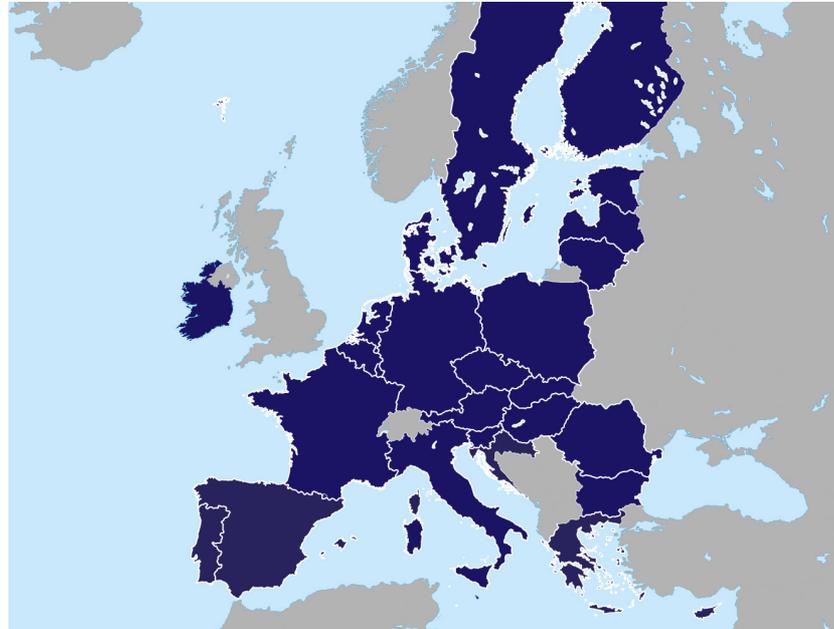
Table 9 Evaluation results for the USC corpus for each temporal tagger, both for identification (two first columns, *lenient* and *strict*) and normalization (two last columns, *lenient* and *strict*). The first row (in white) corresponds to results against the *StandardTimeML* gold standard, while the second (in gray) corresponds to the *LegalTimeML* gold standard.

A	lenient			strict			lenient+ value			strict+ value		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
HE	0.83	0.94	0.88	0.81	0.92	0.86	0.79	0.90	0.84	0.79	0.89	0.83
	0.29	0.97	0.44	0.26	0.88	0.40	0.20	0.67	0.31	0.19	0.64	0.29
SU	0.75	0.99	0.85	0.72	0.95	0.82	0.67	0.88	0.76	0.66	0.86	0.75
	0.25	0.98	0.40	0.23	0.90	0.36	0.18	0.72	0.29	0.17	0.69	0.28
GU	0.84	0.78	0.81	0.71	0.66	0.69	0.67	0.62	0.65	0.65	0.60	0.62
	0.25	0.69	0.36	0.16	0.45	0.23	0.12	0.34	0.18	0.10	0.27	0.14
CA	0.77	0.90	0.82	0.72	0.84	0.77	0.73	0.85	0.78	0.71	0.83	0.76
	0.23	0.82	0.36	0.21	0.72	0.32	0.21	0.73	0.33	0.20	0.69	0.30
CL	0.85	0.84	0.84	0.81	0.79	0.80	-	-	-	-	-	-
	0.30	0.89	0.45	0.26	0.78	0.39	-	-	-	-	-	-
SY	0.85	0.98	0.91	0.78	0.91	0.84	0	0	0	0	0	0
	0.28	0.98	0.44	0.24	0.84	0.37	0	0	0	0	0	0
TE	0.93	0.86	0.90	0.90	0.83	0.86	0.86	0.79	0.83	0.85	0.78	0.81
	0.32	0.90	0.48	0.29	0.81	0.43	0.25	0.69	0.37	0.23	0.64	0.34
TI	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-
US	0.50	0.30	0.37	0.39	0.23	0.29	0.08	0.02	0.08	0.02	0.01	0.02
	0.16	0.28	0.21	0.07	0.13	0.09	0.08	0.14	0.10	0.03	0.05	0.04
UW	-	-	-	-	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-	-	-	-	-

Temporal Tagger

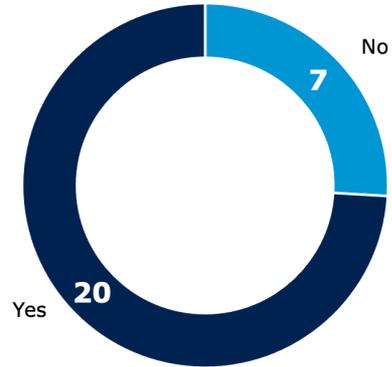
HeidelTime (HE)
SUTime (SU)
GUTime (GU)
CAEVO (CA)
ClearTK (CL)
SynTime (SY)
TERNIP (TE)
TIPSem (TI)
USFD2 (US)
UWTime (UW)

EU member state contribution to The European Legal Data Landscape

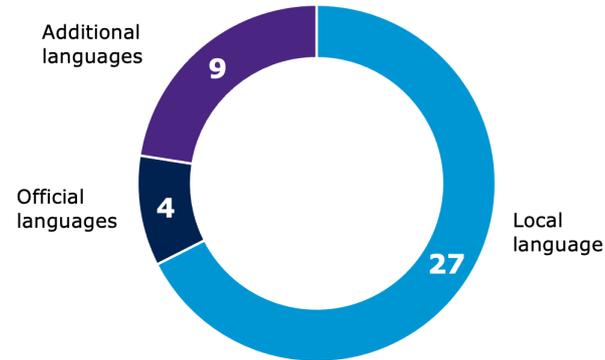


Legal data in Europe

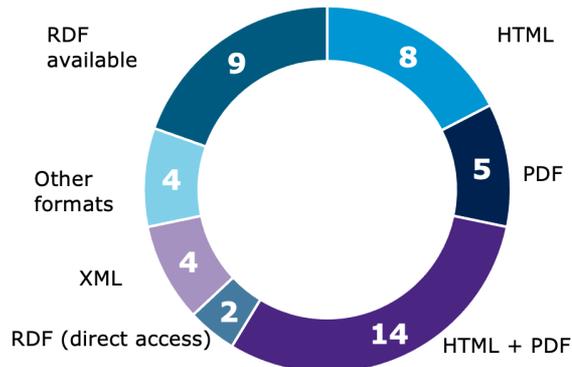
Central Search Interface



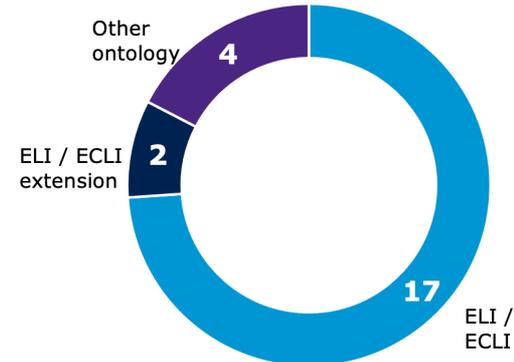
Documents available in language



Document formats



Used ontologies



Legal data in Europe

Country	Central Interface	ECLI Search	Search Interface	Document Format	Languages
Austria	✓	-	Keyword	HTML, PDF, RTF, XML	DE, EN*
Belgium	-	✓	Keyword	HTML	FR, NL, DE
Bulgaria	-	✓	Keyword	HTML, PDF	BG
Croatia	-	✓	Keyword	HTML	HR
Cyprus	✓	-	Keyword	PDF	EL
Czech Republic	-	✓	Keyword	PDF	CZ
Denmark	-	-	Faceted	HTML, PDF	DK
Estonia	✓	✓	Keyword	HTML, PDF, TXT, XML	EE, EN*
Finland	✓	✓	Keyword, SPARQL	HTML	FI, SE
France	✓	✓	Keyword	HTML, PDF	FR, EN*, DE*, IT*, ES*
Germany	-	✓	Keyword	HTML	DE, EN*
Greece	-	✓	Keyword	PDF	EL
Hungary	-	-	Keyword	HTML	HU, EN*
Ireland	-	-	Keyword	HTML, PDF	EN
Italy	-	✓	Keyword	HTML	IT
Latvia	-	✓	Keyword	HTML, PDF	LV, EN*, RU*
Lithuania	-	-	Faceted	HTML, PDF	LT
Luxembourg	-	-	Faceted, SPARQL	HTML, PDF, XML, RDF	FR
Malta	-	-	Keyword	PDF	MT, EN
Netherlands	-	✓	Both	HTML, PDF, RDF	NL, FR, EN*
Poland	-	-	Keyword	PDF	PL
Portugal	✓	✓	Faceted	HTML, PDF	PT, EN*
Romania	-	-	Keyword	HTML	RO
Slovakia	✓	-	Keyword	HTML, PDF	SK
Slovenia	-	✓	Keyword	HTML, PDF, DOCX	SI, EN*
Spain	-	✓	Both	HTML, PDF, XML, EPUB	ES
Sweden	-	-	Keyword	HTML	SE

Legal data in Europe

Country	Implementation ELI	Implementation ECLI	Data Availability	Information ELI / ECLI / NAL	Thesaurus
Austria	Identifier	Identifier	-	✓/ - / -	✓
Belgium	Identifier	Identifier	-	✓/ ✓/ -	-
Bulgaria	-	Identifier	-	- / ✓/ -	-
Croatia	Identifier	Identifier	-	✓/ - / -	-
Cyprus	-	-	-	- / ✓/ -	-
Czech Republic	-	Identifier	-	- / ✓/ -	-
Denmark	Identifier/Metadata	-	RDF	✓/ ✓/ ✓	-
Estonia	-	Identifier	-	- / ✓/ -	✓
Finland	Identifier/Metadata (+)	Identifier/Metadata (+)	RDF	✓/ ✓/ ✓	✓
France	Identifier/Metadata	Identifier	RDFa	✓/ ✓/ -	-
Germany	-	Identifier/Metadata	-	- / ✓/ -	-
Greece	-	Identifier	-	-	-
Hungary	-	-	-	✓/ - / -	-
Ireland	Identifier/Metadata	-	RDFa, RDF	✓/ ✓/ -	-
Italy	Identifier/Metadata	Identifier	RDFa, RDF	✓/ ✓/ ✓	-
Latvia	- (+)	Identifier	-	- / ✓/ -	-
Lithuania	-	-	-	- / ✓/ -	✓
Luxembourg	Identifier/Metadata (+)	-	RDFa	✓/ - / ✓	-
Malta	-	Identifier	-	✓/ - / -	-
Netherlands	- (+)	Identifier/Metadata	RDFa, RDF	- / ✓/ -	-
Poland	-	-	-	-	-
Portugal	Identifier/Metadata	Identifier (+)	RDFa	✓/ - / -	-
Romania	-	Identifier	-	- / ✓/ -	-
Slovakia	-	Identifier	-	- / ✓/ -	✓
Slovenia	- (+)	Identifier	-	- / ✓/ -	-
Spain	Identifier/Metadata	Identifier	RDFa	✓/ - / ✓	-
Sweden	-	-	-	-	-

Legal data in Europe

Project	Type	Using ELI / ECLI	Extension ELI / ECLI	Data Availability	Thesaurus	Open Data Linking	SPARQL
Legal Knowledge Graph	Linking	✓/✓	LKG / LKG	RDF	EuroVoc, Other	✓	✓
Semantic Finlex	Linking	✓/✓	SFL / SFCL	RDF	EuroVoc, Other	✓	✓
Nomothesia	Linking	✓/ -	Nomothesia / -	RDF	-	✓	✓
EUCases	Linking	✓/✓	-	-	EuroVoc, Other	-	-
Lynx	Linking	✓/ -	Lynx-LKG	RDF	EuroVoc, Other	✓	✓
GDPRtEXT	Linking	✓/ -	GDPRtEXT	RDF	-	-	✓
Linkoln	Extraction	✓/ -	-	-	-	-	-
BO-ECLI	Extraction	-/✓	-	-	-	-	-

Putting it all together!

How has the European Directive 2014/92/EU been implemented in EU member states?

```
PREFIX eli: <http://data.europa.eu/eli/ontology#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
select ?title ?document where {
  VALUES ?format {
    <http://www.iana.org/assignments/media-types/text/html>
    <http://www.iana.org/assignments/media-types/application/html> }
  ?n ?p <http://data.europa.eu/eli/dir/2014/92/oj> ;
    eli:is_realized_by ?r .
  ?r eli:is_embodied_by ?document ;
    eli:title ?title .
  ?document eli:format ?format .
}
```

Title	Document
Bundesgesetz, mit dem ein Bundesgesetz über die Vergleichbarkeit von Entgelten für Verbraucherzahlungskonten, ...	https://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2016_I_35/BGBLA_2016_I_35.html
European Union (Payment Accounts) Regulations 2016.	http://www.irishstatutebook.ie/eli/2016/si/482/made/en/html
Lov om betalingskonti	http://www.retsinformation.dk/eli/lta/2016/375/dan/html
Attuazione della direttiva 2014/92/UE, sulla comparabilità delle spese relative al conto di pagamento, ...	http://www.gazzettaufficiale.it/eli/id/2017/03/30/17G00051/sg/ita/html

A Roadmap for European Lined Legal Data

Challenges

- Slow uptake of ELI and ECLI
- Availability of legal documents in databases
- ELI / ECLI implementation information
- Deviation from ELI and ECLI
- Non-machine readable / proprietary data formats
- Search interfaces / data accessibility

Opportunities

- Involve governments
- Funding possibilities
- Licensing and access policies
- Provide implementation information
- Involve stakeholders for changes
- Provide data also using ELI and ECLI
- Use open / structured data formats
- Provide application programming interfaces (API)

Thank you / contact details



Department of Information Systems & Operations

Institute for Information Systems & New Media
Welthandelsplatz 1, 1020 Vienna, Austria

Dr. Sabrina Kirrane

T +43-1-313 36-4494
F +43-1-313 36-90 4494
sabrina.kirrane@wu.ac.at
www.wu.ac.at
www.sabrinakirrane.com
[@SabrinaKirrane](https://twitter.com/SabrinaKirrane)



Sabrina Kirrane is funded by the FWF Austrian Science Fund and the Internet Foundation Austria under the FWF Elise Richter and netidee SCIENCE programmes as project number V 759-N.