



Processing, Identification and Representation of Temporal Expressions and Events in Legal Documents

María Navas-Loro

Supervisors: Víctor Rodríguez-Doncel
Asunción Gómez-Pérez



mnavas@fi.upm.es



<https://mariannavas.oeg-upm.net>



17/01/2022



Online

Outline

- Introduction
- State of the Art
- Research Objectives
- Temporal Expressions
 - Corpora
 - Temporal Tagging
- Events
 - Corpus
 - Event Extraction
 - Event-related Resources
- Conclusions and Future Work

Outline

- Introduction
- State of the Art
- Research Objectives
- Temporal Expressions
 - Corpora
 - Temporal Tagging
- Events
 - Corpus
 - Event Extraction
 - Event-related Resources
- Conclusions and Future Work

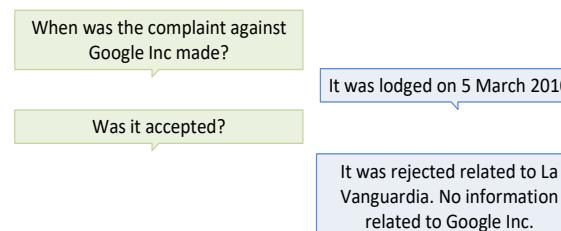
Temporal information is everywhere:
news, medical texts, legal documents...
... and we can use it for a lot of things!



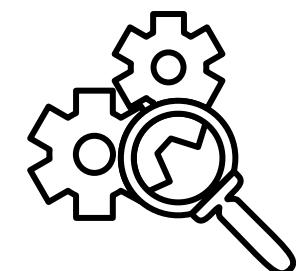
PATTERN
DETECTION



SUMMARIZATION



QA



SEARCH
ENGINE



TIMELINE GENERATION

Not a lot of tools in Spanish or in the legal domain.

Temporal taggers (TimeML, ISO standard):

1) Detect temporal expressions:

- DATE
- SET

- DURATION
- TIME

2) Normalize them.

The 3rd of April 2016 is OK, but the 20th May would be better.

```
<?xml version="1.0"?>
<!DOCTYPE TimeML SYSTEM "TimeML.dtd">
<TimeML>
<TIMEX3 tid="t4" type="DATE" value="2016-04-03">The 3rd of April 2016</TIMEX3>
is OK, but <TIMEX3 tid="t6" type="DATE" value="2016-05-20">the 20th May</TIMEX3>
would be better.
</TimeML>
```

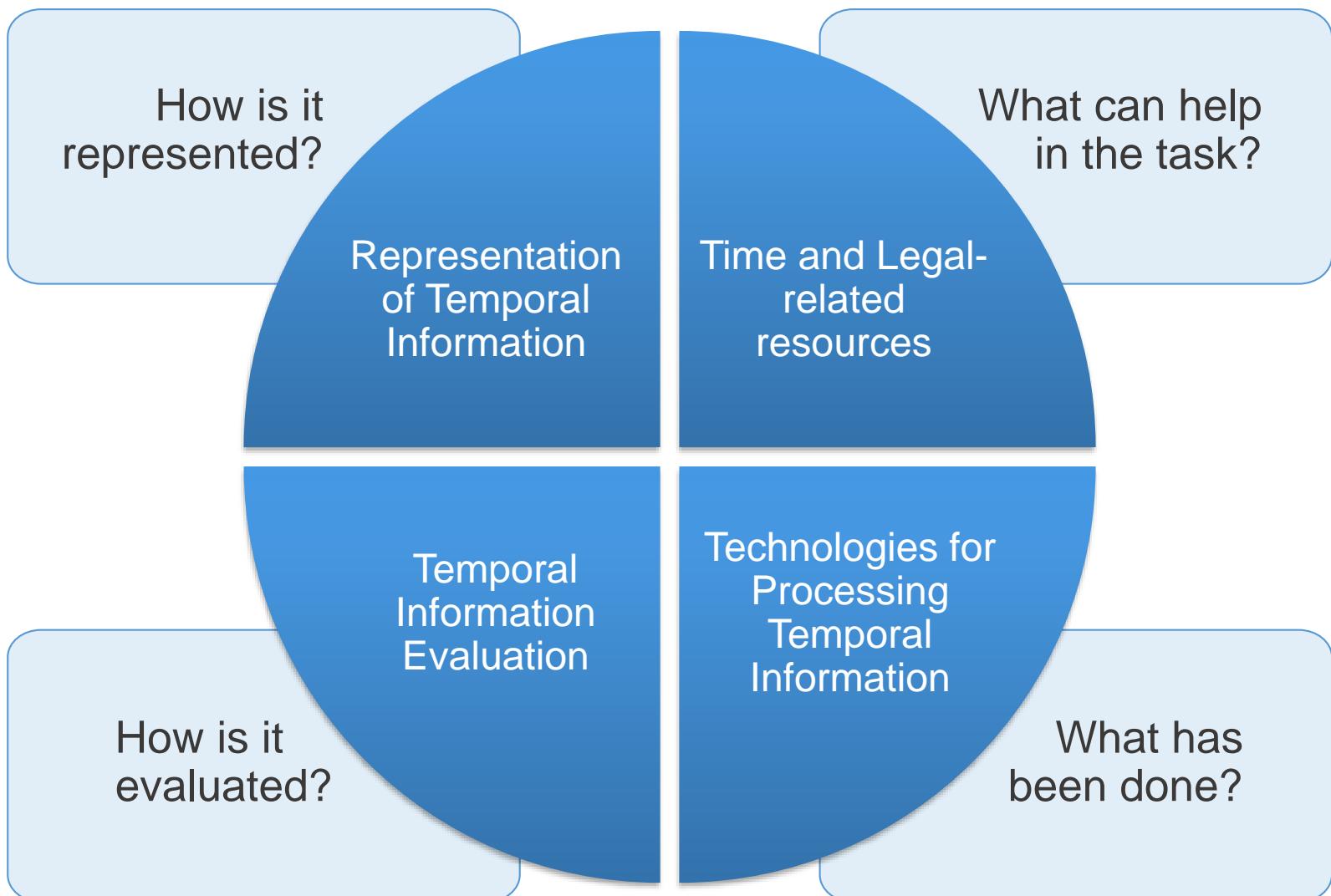
3) Additionally, event and relation extraction

State-of-the-art temporal taggers that implement TimeML annotations, also present shortcomings:

- Most **optional attributes are never used**, even if they are useful (e.g., modifiers), probably because they do not count in the evaluation of related challenges.
- TimeML **EVENT annotation is less common** than temporal expression annotation, and temporal taggers are often just used for temporal expression annotation, while the event extraction task is done otherwise.
- Most temporal taggers are not able to detect **compound durations**, since it is not usual to find them.
- Since it is the way they are usually evaluated, most temporal taggers focus on **covering expressions that can appear in news**.

Outline

- Introduction
- State of the Art
- Research Objectives
- Temporal Expressions
 - Corpora
 - Temporal Tagging
- Events
 - Corpus
 - Event Extraction
 - Event-related Resources
- Conclusions and Future Work





Temporal Expressions TimeML

TimeML precursors

TIDES TIMEX2
(Ferro et al., 2001)
Sheffield STAG
(Setzer, 2002)

TimeML

(Pustejovsky et al., 2010)

- Temporal Expressions
- Relations
- Events
- SIGNALS,MAKEINSTANCE

TimeML extensions

THYME Project
(Styler et al., 2014)
Probabilistic approach
(Angeli et al., 2012)

Events schemas and ontologies

Ann. Schemas

ACE, ERE, KBP, RED
Domain: GAF, CAMEO

Ontologies

W3C Time, SEM, TEO, ESO
Event Ont. + Timeline Ont.

Other Representations

Legal: LKIF, Akoma Ntoso
Generic: NIF, Web Annotations

➤ TimeML standard for Temporal Expressions, but no consensual way to represent events.

Need to integrate/ease the transition between representations.

Corpora



Timebank corpus (Pustejovsky et al., 2003b)

TempEval challenges

MEANTIME corpus (Minard et al., 2016)

English



Wikiwars corpus
(Mazur and Dale, 2010)



THYME corpus
(Styler et al., 2014)



Scientific abstracts
(Strötgen and Gertz, 2012)



Tweets
(Tabassum et al., 2016)



Colloquial texts
(Strötgen and Gertz, 2012)

Spanish



ModeS TimeBank
(17th and 18th centuries)



Lack of **Spanish corpora** annotated with temporal information.

Lack of **legal corpora** annotated with temporal information.



Temporal Taggers

T. Taggers	Approaches	Languages	Id.	Norm.	Ev.	Rel.
HeidelTime	Rule-based	ES,EN,DE,+200	✓	✓	-	-
SUTime	Rule-based	ES,EN	✓	✓	-	-
GUTime	Hybrid	EN	✓	✓	✓	✓
CAEVO	Hybrid	EN	✓	✓	✓	✓
ClearTK	Machine-Learning	EN	✓	-	✓	✓
SynTime	Rule-based	EN	✓	-	-	-
TERNIP	Rule-based	EN	✓	✓	-	-
TIPSem	Hybrid	ES,EN	✓	✓	✓	✓
USFD2	Hybrid	EN	*	*	-	*
UWTime	Hybrid	EN	✓	✓	-	-

Event Extraction

- Hagege and Tannier (2008) an event is any verb, any deverbal noun, any noun argument of the preposition during, or any time span noun.
- Capet et al. (2008) ad hoc templates (core + coordinates).
- Application: timeline generation from texts.
 - Generic: Linea (Etienne et al., 2015) and TimeLineCurator (Fulda et al., 2015)
 - Journalism (Tannier and Vernier, 2016), Medicine (Styler et al., 2014)



Approaches in the legal domain

- Schilder (2005): three types of documents:
 - Transactional documents
 - Constraints in statutes or regulations
 - Legal narratives in case law
- Iseemann et al. (2013) TI from regulations. Confusion between:
 - Episodic and generic statements
 - Legal references and dates

Event Extraction in the legal domain Legal Information Retrieval

- Lagos et al. (2010) several types of events (who, what, when and where).
- Maxwell et al. (2009) eventuality (event, state, or attribute)
- Spanish: Sierra et al. (2018), for instance, aims to extract events from Mexican legal texts (who, what, to whom and where)
- Portuguese: Bertoldi et al. (2014) manual semantic legal frames

Related tools



Temporal Expressions

NER

NLTK [dates, timex extension]
(Loper and Bird, 2002)

Spacy, OpenNLP [dates,times]
(Honnibal et al., 2020)
(Apache Soft. Fund., 2014)

CoreNLP [SUTime]
(Manning et al., 2014)

Event Extraction related tasks

Semantic Role Labelling

AllenNLP (Gardner et al., 2017)
IxaPipes (Agerri et al., 2014)

Open IE

AllenNLP (Gardner et al., 2017)
CoreNLP (Manning et al., 2014)

Frame Identification

OpenSESAME (Swayamdipta et al., 2017)
Framat (Roth and Lapata, 2015)
FRED (Gangemi et al., 2017), TakeFive (Alam et al., 2021)

- Lack of temporal tagger for Spanish handling frequent TEs
- No temporal tagger covering specific legal considerations.
- Not automatic relevant event extraction in the legal domain.



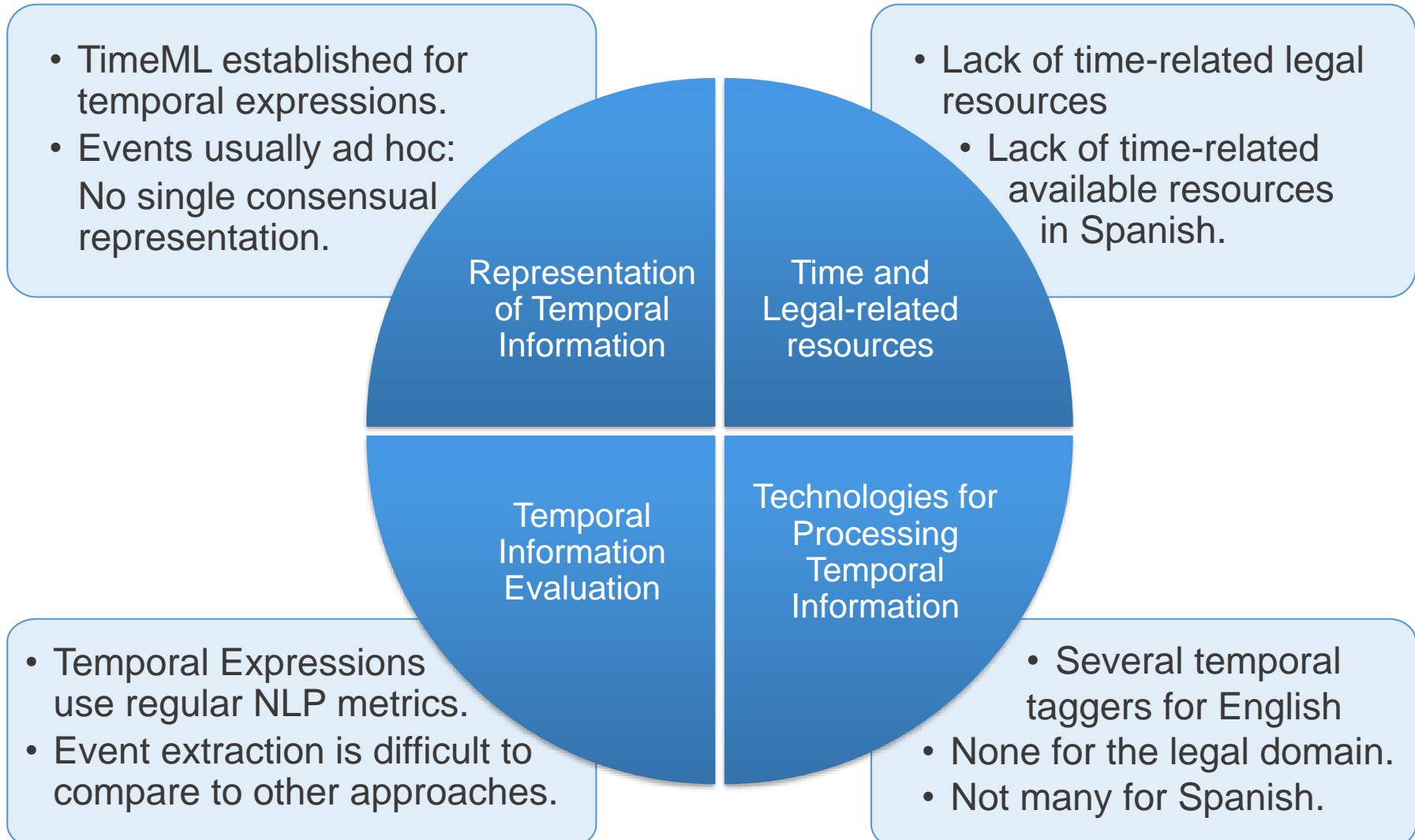
Temporal Expressions

- The TimeML standard is evaluated using NLP measures (Precision, Recall, F1-measure)
- Aspects usually evaluated are:
 - The extent of the annotation fits the reference annotation
 - The type of the expression is correctly classified
 - The normalized value equals the reference one

```
<TIME3 tid="t1" type="DATE" value="1990-07-06">6 July 1990</TIME3>
```

Events

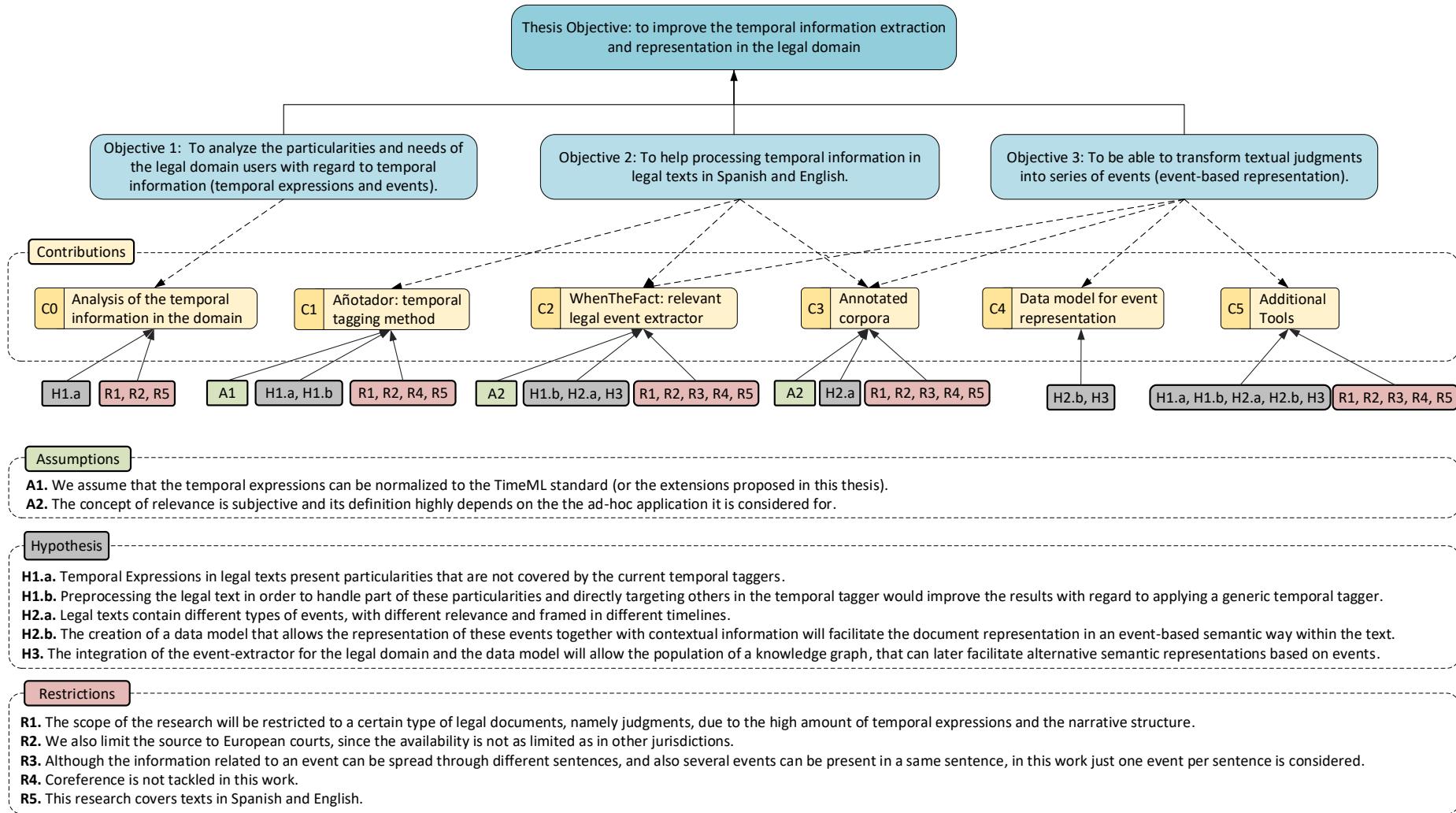
- TempEval: P,R,F1 to assess extent of the annotation and the assignment the event type, but in 2015 shifted to temporal QA.
- ACE VDR value, a metric taking into account:
 - The extent of the event annotation
 - The event arguments and their attributes value and modality.
- BioNLP'09 shared task on event extraction
 - Different levels: Core (trigger,type), arguments, factuality
- Other ways: Task oriented (e.g., summarization).



Outline

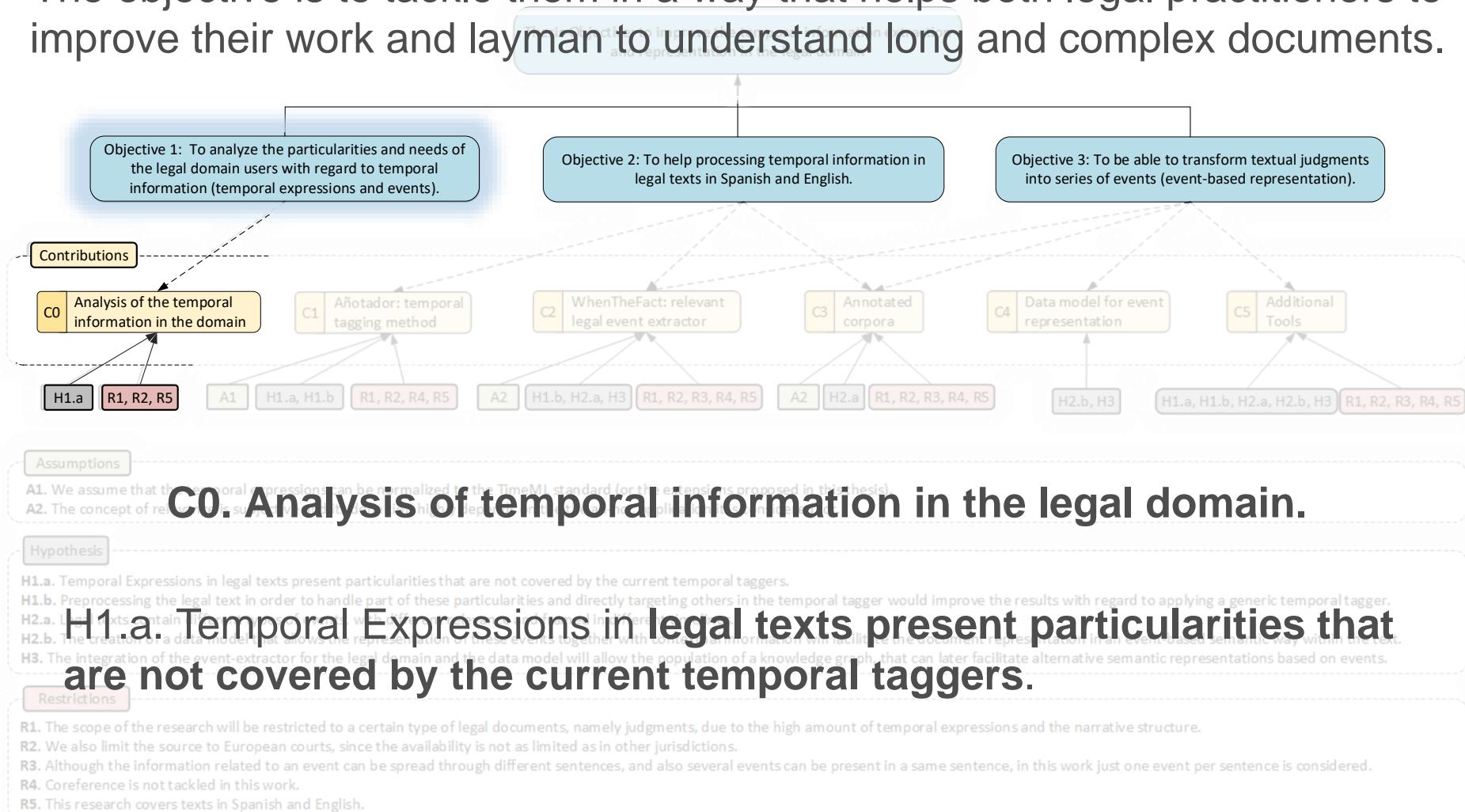
- Introduction
- State of the Art
- **Research Objectives**
- Temporal Expressions
 - Corpora
 - Temporal Tagging
- Events
 - Corpus
 - Event Extraction
 - Event-related Resources
- Conclusions and Future Work

Thesis objective: To improve the temporal information extraction and representation in the legal domain.

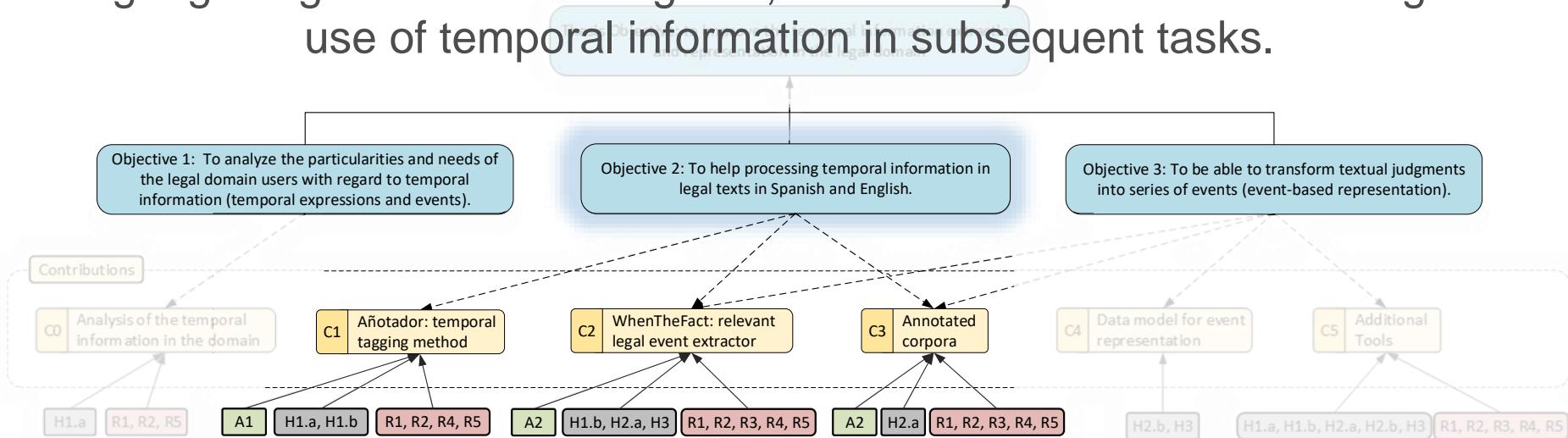


- H1.a.** Temporal Expressions in **legal texts present particularities that are not covered by the current temporal taggers.**
- H1.b.** Preprocessing the legal text in order to handle part of these **particularities** and directly targeting others in the temporal tagger would improve the results with regard to applying a generic temporal tagger.
- H2.a.** Legal texts contain **different types of events**, with different relevance and framed in different timelines.
- H2.b.** The creation of a **data model** that allows the representation of these events together with contextual information and their annotation details will allow **facilitating the document representation in an event-based semantic way** within the text.
- H3.** The **integration of the event-extractor for the legal domain and the data model will allow the population of a knowledge graph**, that can later facilitate alternative semantic representations based on events such as timelines, semantic searches or summarization generation.

O1. To analyze the particularities and needs of the legal domain users with regard to temporal information (temporal expressions and events).
 The objective is to tackle them in a way that helps both legal practitioners to improve their work and layman to understand long and complex documents.



O2. To help processing temporal information in legal texts in Spanish and English. Also, the lack of resources and technologies in the Spanish language in general will be targeted, and the objective is to encourage the use of temporal information in subsequent tasks.



C1. Añotador: temporal tagger, C2. WhenTheFact: relevant legal event extractor, C3. Annotated Corpora

A1. We assume that the temporal expressions can be normalized to the TimeML standard (or the extensions proposed in this thesis).

A2. The concept of relevance is subjective and its definition highly depends on the ad-hoc application it is considered for.

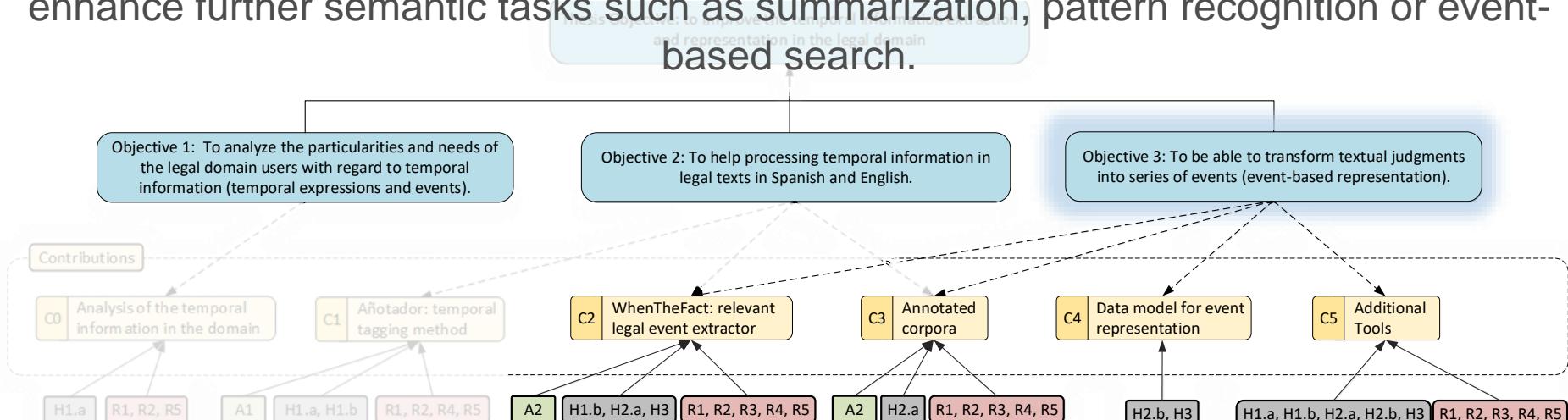
H1.a. Temporal Expressions in legal texts present particularities that are not covered by the current temporal taggers.

H1.a. Temporal Expressions in legal texts present particularities that are not covered by the current temporal taggers.
 H1.b. Preprocessing the legal text in order to handle part of these particularities and directly targeting others in the temporal tagger would improve the results with regard to applying a generic temporal tagger.
 H2.b. The creation of a data model that allows the representation of these events together with contextual information will facilitate the document representation in an event-based semantic way within the text.
 H3. The integration of the event-extractor for the legal domain and the data model will allow the population of a knowledge graph, that can later facilitate alternative semantic representations based on events such as timelines, semantic searches or summarization generation.

H2.a. Legal texts contain different types of events, with different relevance and framed in different timelines.

R1. The scope of the research will be restricted to certain type of legal documents, namely judgments, due to the high amount of temporal expressions and the narrative structure.
 R2. We also consider that the legal domain is very specific and that the particularities of the legal domain must be taken into account when dealing with legal texts.
 R3. Although the temporal expressions in legal texts are usually framed in different timelines, this is not always the case, so this aspect must be considered.
 R4. Coreference is not tackled in this work.
 R5. This is a first step towards the final goal.

O3. To be able to transform textual judgments into series of events. This event-based representation would allow expressing judgments in intuitive and easy to understand ways, such as timelines or event-based knowledge graphs, or to enhance further semantic tasks such as summarization, pattern recognition or event-based search.



C2. WhenTheFact: relevant legal event extractor, C3. Annotated Corpora, C4. FT3: Data Model, C5. Additional Tools

- A1. We assume that the temporal expressions can be normalized to the TimeML standard (the extraction approach is not specified).
- A2. The concept of relevance is subjective and its definition highly depends on the the ad-hoc applications considered for.

Hypothesis

- H2.b. The creation of a **data model** that allows the representation of these events together with contextual information and their annotation details will allow **facilitating the document representation in an event-based semantic way within the text**.
- H3. The integration of the event-extractor for the legal domain and the data model will allow the population of a knowledge graph, that can later facilitate alternative semantic representations based on events.

- R1. The scope of the work is limited to the European Union, since the availability of legal documents in other jurisdictions is not as extensive as in the EU.
- R2. We also limit the source to European courts, since the availability is not as limited as in other jurisdictions.
- R3. Although the implementation of the system is done through the use of the TimeML standard, the language of the input and output is not limited to English, as the context of the sentence is considered.
- R4. Coreference is not tackled in this work.
- R5. This research covers texts in Spanish and English.

- Corpora
- Temporal Tagging

TEMPORAL EXPRESSIONS

- Corpora
- Temporal Tagging

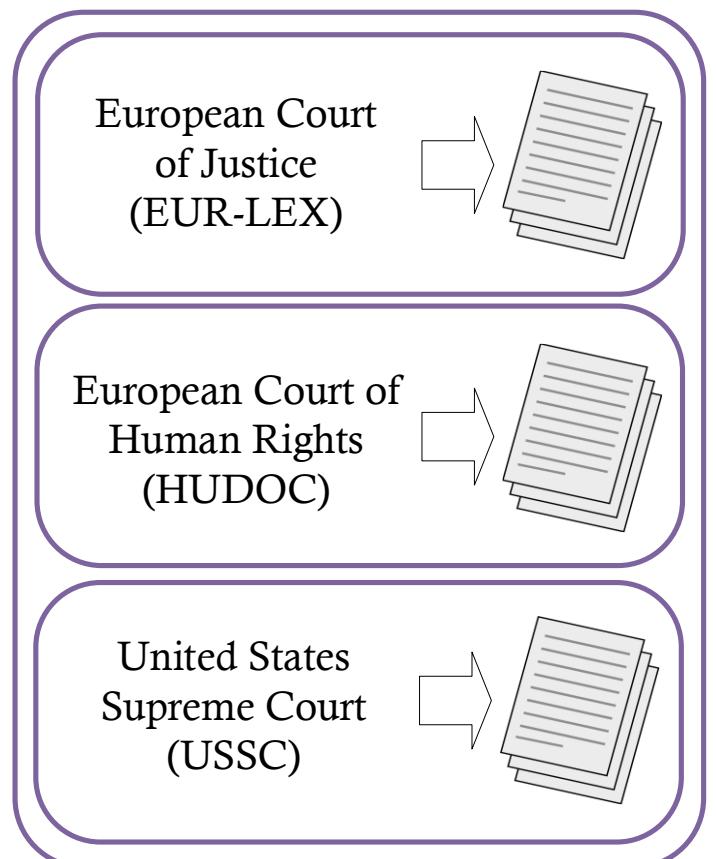
TEMPORAL EXPRESSIONS

Legal Corpus

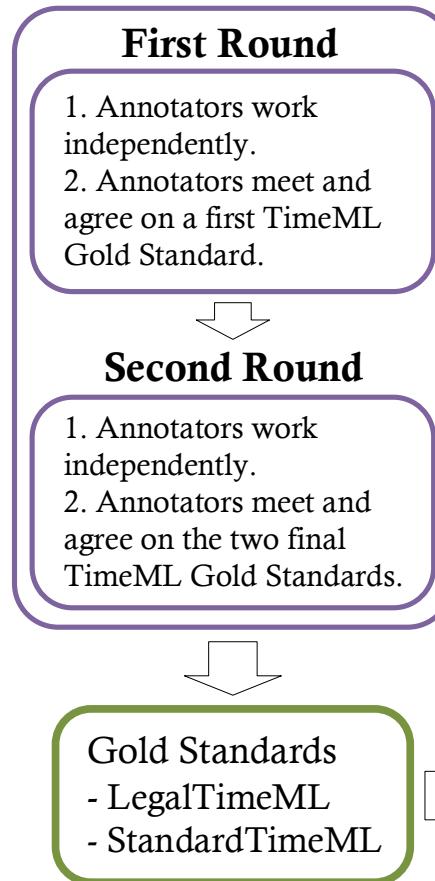
TEMPCOURT

First corpus of legal documents annotated with TEs

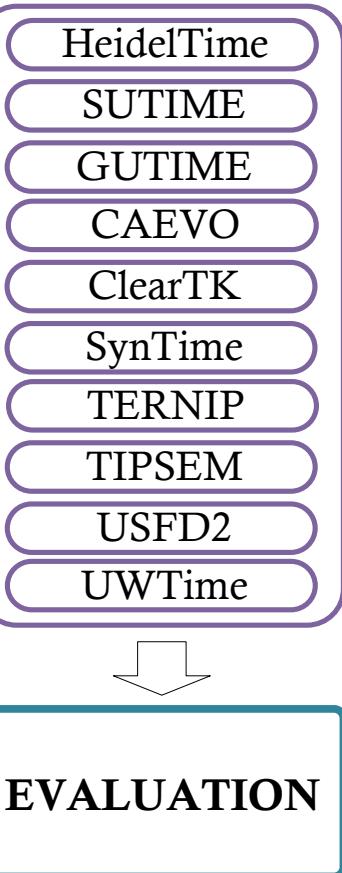
DOCUMENT COLLECTION



ANNOTATION



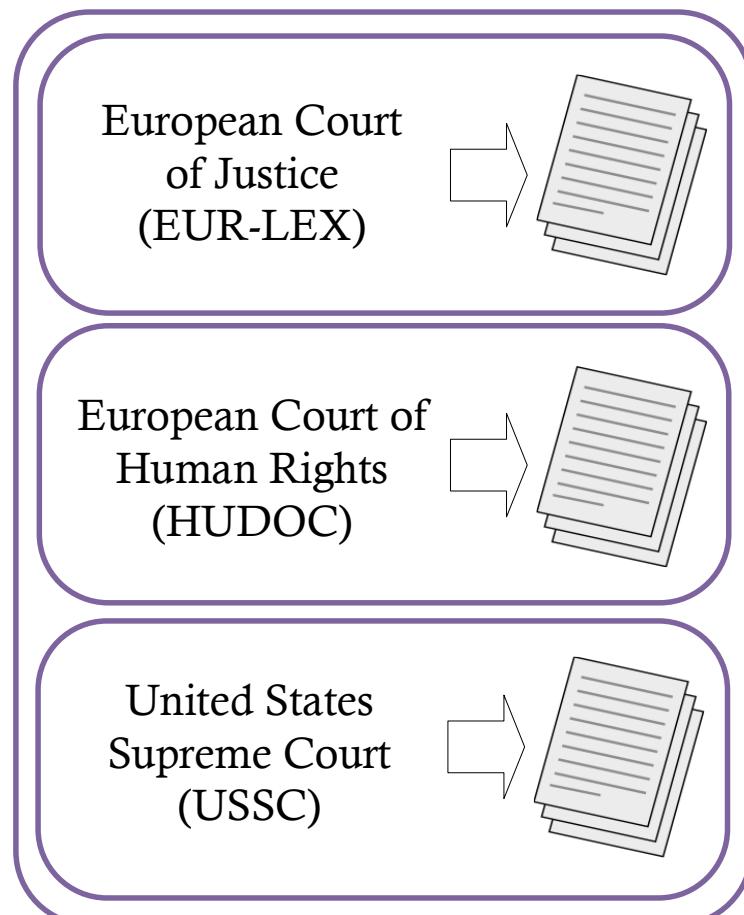
TAGGING



Gold Standards
- LegalTimeML
- StandardTimeML

EVALUATION

DOCUMENT COLLECTION



30 legal decisions:

- 10 from ECJ
 - 10 from ECHR
 - 10 from USSC

Normalization, standardization: from doc/pdf to text file

gold/pdt
Gold Standards

DOCUMENT COLLECTION

Two annotation sets:

- LegalTimeML
- StandardTimeML

European Court
of Human Rights
(HUDOC)

StandardTimeML

United States
Supreme Court
(USSC)

LegalTimeML

ANNOTATION

First Round

1. Annotators work independently.
2. Annotators meet and agree on a first TimeML Gold Standard.

Second Round

1. Annotators work independently.
2. Annotators meet and agree on the two final TimeML Gold Standards.

Gold Standards

- LegalTimeML
- StandardTimeML

TAGGING

HeidelTime

SUTIME

GUTIME

CAEVO

ClearTK

SynTime

TERNIP

TIPSEM

USFD2

UWTime

EVALUATION

DOCUMENT COLLECTION

Also 10 state-of-the-art temporal tagger annotations added to the corpus.

Result:

- Two gold standards
- Benchmark of results of 10 temporal taggers for comparison

ANNOTATION

1. Annotators work independently.
2. Annotators meet and agree on a first TimeML Gold Standard.

Second Round

1. Annotators work independently.
2. Annotators meet and agree on the two final TimeML Gold Standards.

- Gold Standards**
- LegalTimeML
- StandardTimeML

TAGGING

- HeideITime
- SUTIME
- GUTIME
- CAEVO
- ClearTK
- SynTime
- TERNIP
- TIPSEM
- USFD2
- UWTime

EVALUATION

The annotations by the temporal taggers were analyzed, and the main lacks in legal texts were detected:

- Date formats (eg, “DD/MM/YYYY”)
- Currency identified as a year (“EUR 2000”)
- Polysemous words (“fall”, “may”)
- SETs considered DURATIONs (“Once a week”)
- Compund durations are separated (“One year and one day”).
- Series of dates (“15 and 16 December”)
- MODs not used
- No in-text intervals
- Year-like expressions tagged (“No 1612/68”, “§1408”)

Navas-Loro, M., Filtz, E., Rodríguez-Doncel, V., Polleres, A., and Kirrane, S. (2019). “TempCourt: Evaluation of temporal taggers on a new corpus of court decisions”. *The Knowledge Engineering Review*, 34, E24.

Spanish Corpus

HOURGLASS

It is difficult to systematically test a temporal tagger.
We created a dataset named Hourglass, with two parts:

50-2-1991	false
En el 999 AC.	year
El 1, el 2 y el 3 de enero.	consecutive dates
Tres siglos y medio.	fractions
Este siglo.	relative
Él tiene 25 primaveras	coloquial special
Cinco para las 11	LatinAmerica yes
Se hizo illa tempora.	latin Standard
Bueno pues a y 45.	chat yes
...	

Synthetic part, developed for testing purposes.

- Includes Temporal Expressions that a temporal tagger should cover.
- Includes **tags** in order to facilitate the evaluation of different expressions.

People part, contributors foreign to the task provide expressions.

- They had different backgrounds.
- They came from different Spanish-speaking countries and regions.
- Each expression tagged with its **register**.

The following examples were difficult to handle to the taggers:

Example	Añotador	SUTime	HeidelTime
<u>“1 año, 6 meses y un día”</u> ("1 year, 6 months and one day")	<u>1 año, 6 meses y un día</u>	1 año, 6 meses y un día	<u>1 año, 6 meses y un día</u>
<u>“Cinco para las 11.”</u> ("Five to eleven.")	<u>Cinco para las 11.</u>	Cinco para las 11.	Cinco para las 11.
“lo vuestro dura <u>1h</u> , no?” ("your stuff lasts <u>1h</u> , right?")	lo vuestro dura <u>1h</u> , no?	lo vuestro dura 1h, no?	lo vuestro dura 1h, no?
<u>“en cero coma”</u> (in a short amount of time)	en cero coma	en cero coma	en cero coma

Navas-Loro, M. and Rodríguez-Doncel, V. “Annotador: a Temporal Tagger for Spanish”. Journal of Intelligent & Fuzzy Systems, vol. 39, no. 2, pp. 1979-1991, 2020.

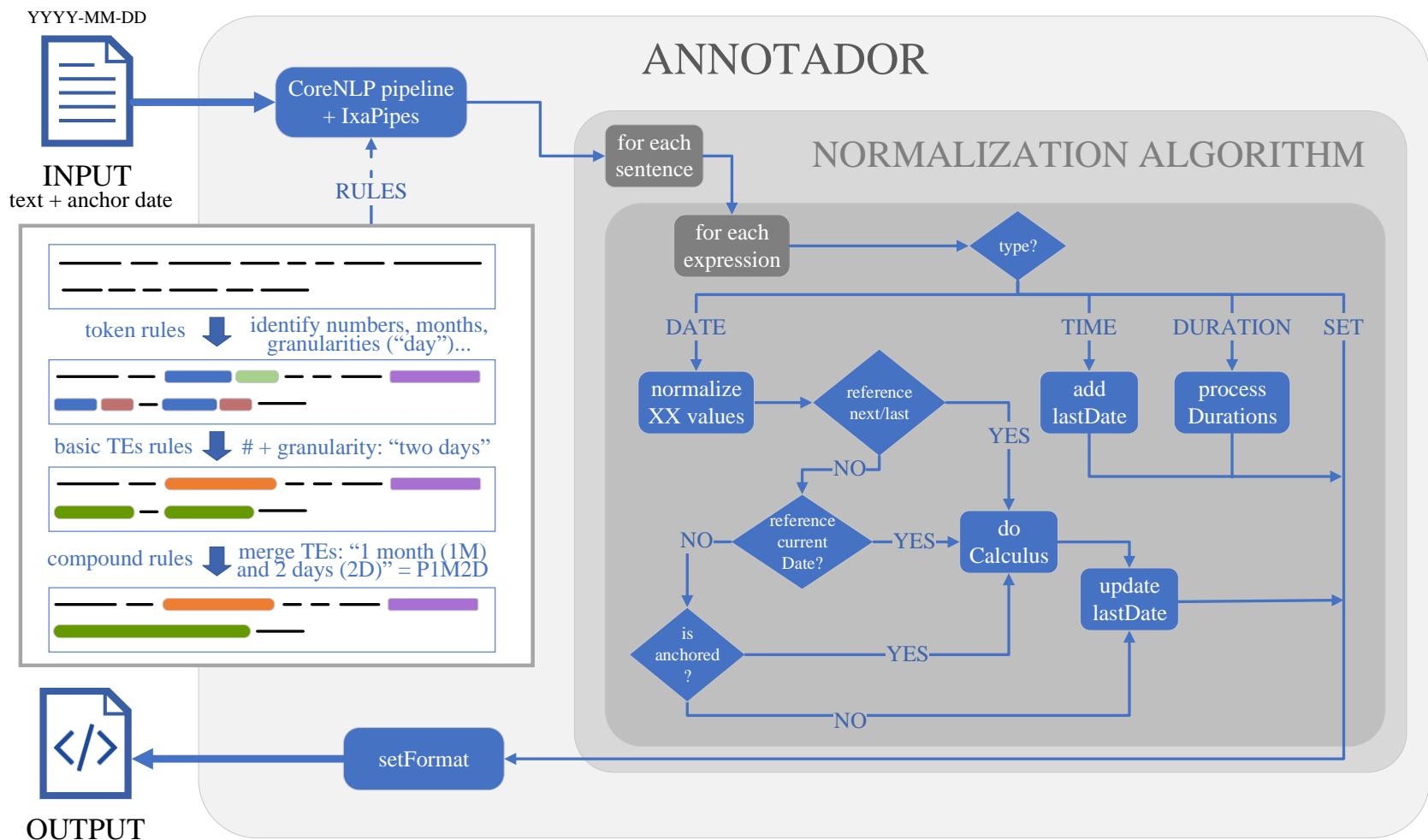
- Corpora
- Temporal Tagging

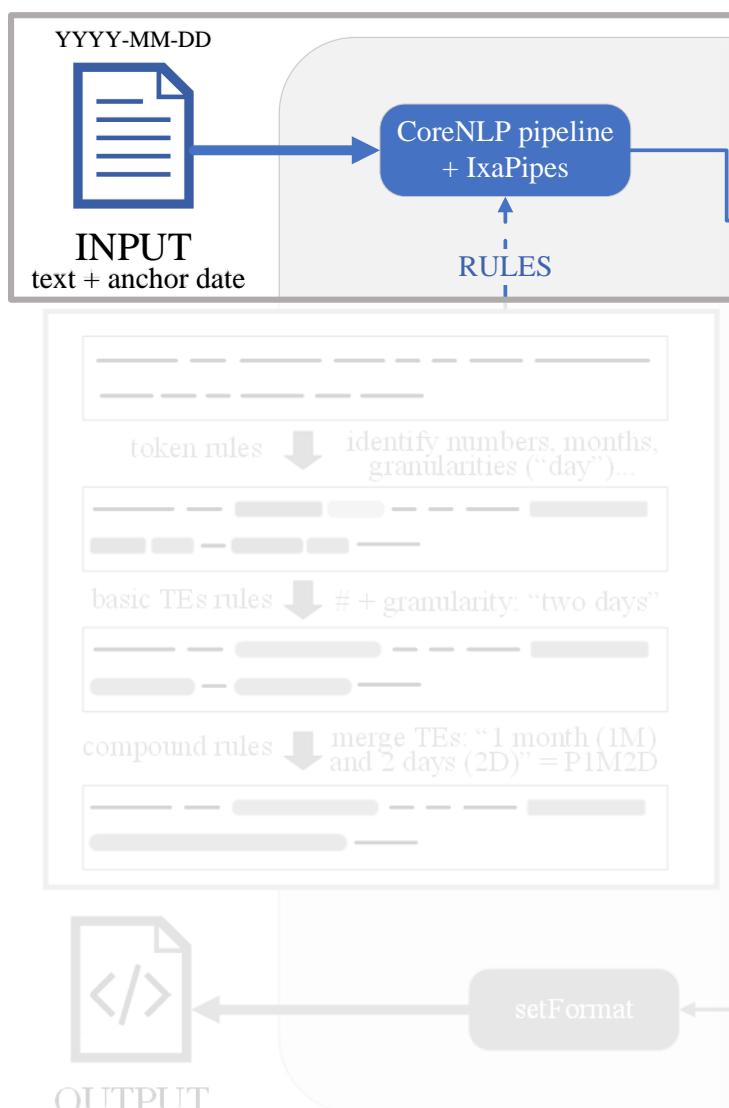
TEMPORAL EXPRESSIONS

Temporal tagging

AÑOTADOR

Añotador is a temporal tagger for Spanish and English that targets both generic texts or legal texts.

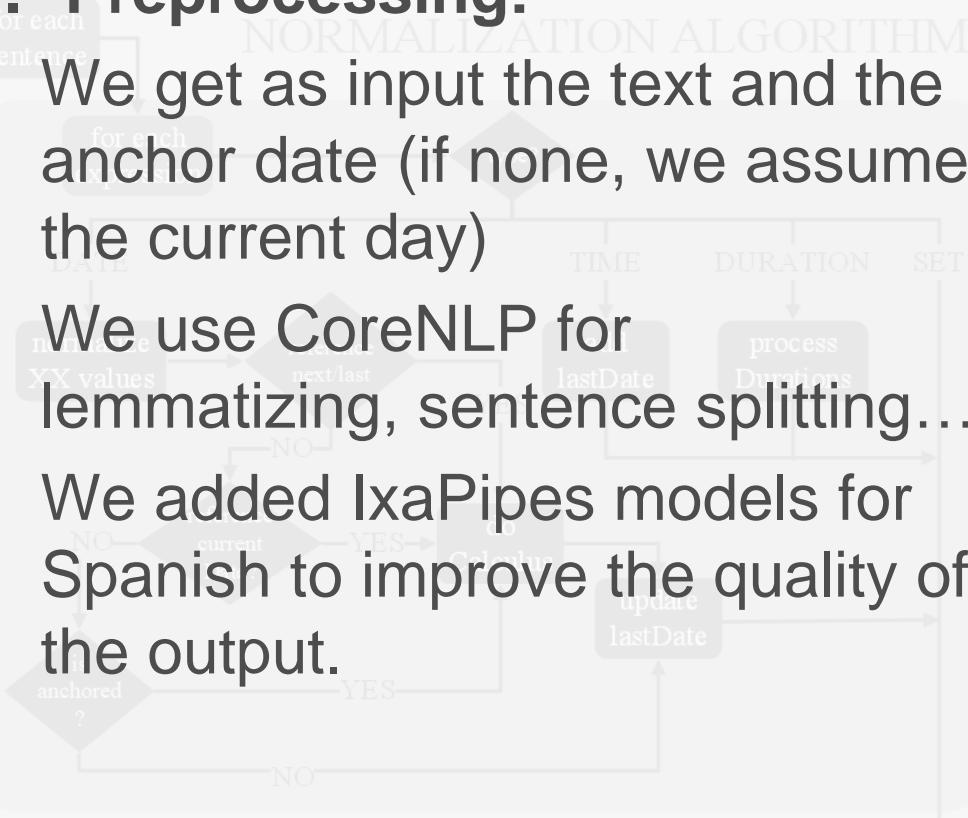


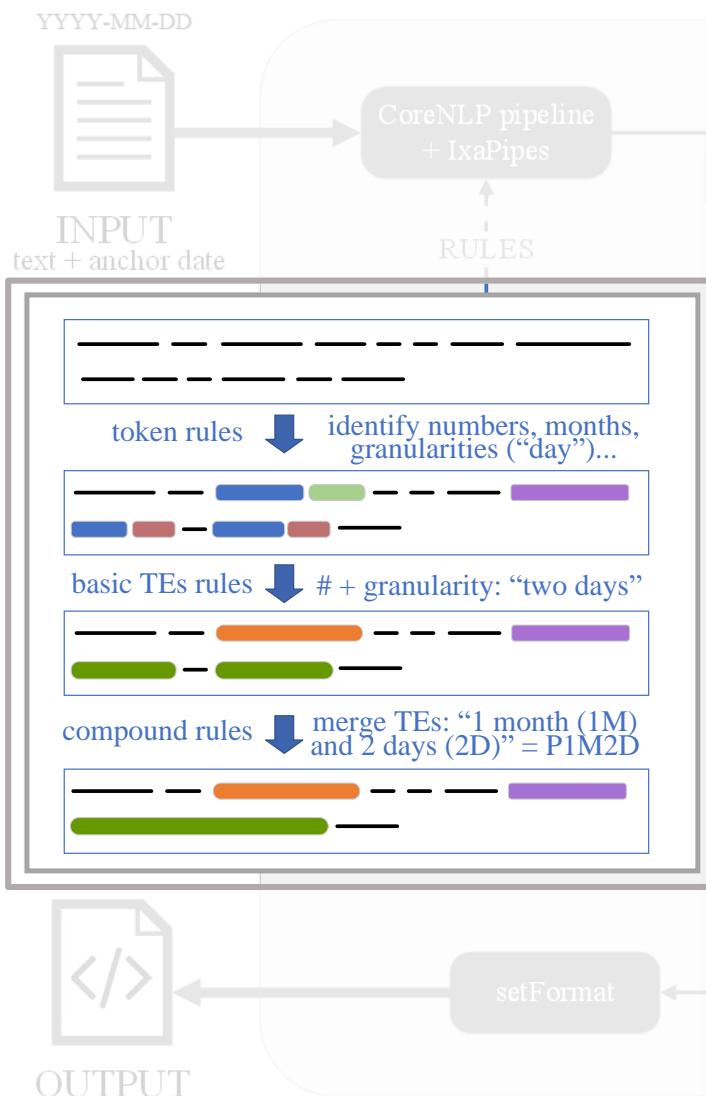


ANNOTADOR

1. Preprocessing:

- We get as input the text and the anchor date (if none, we assume the current day)
- We use CoreNLP for lemmatizing, sentence splitting...
- We added IxaPipes models for Spanish to improve the quality of the output.





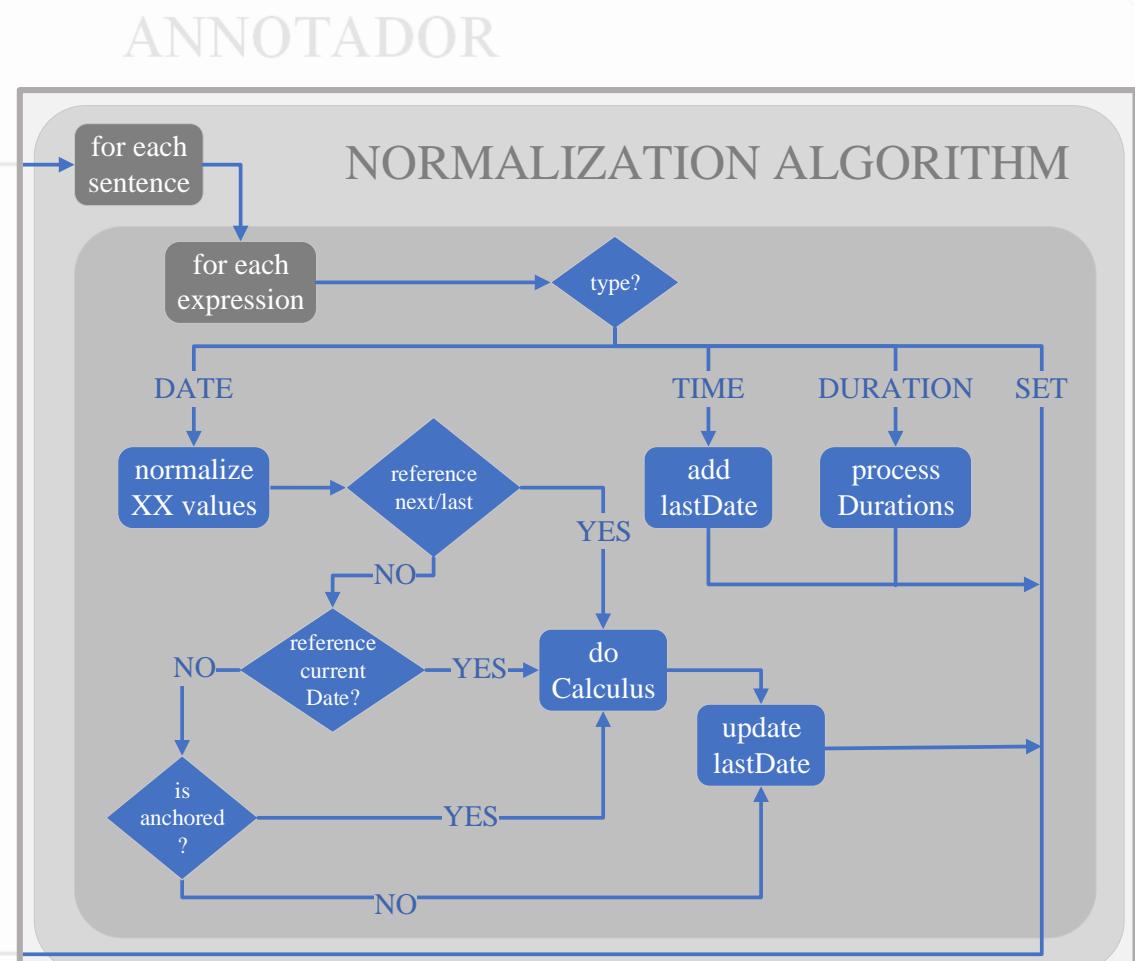
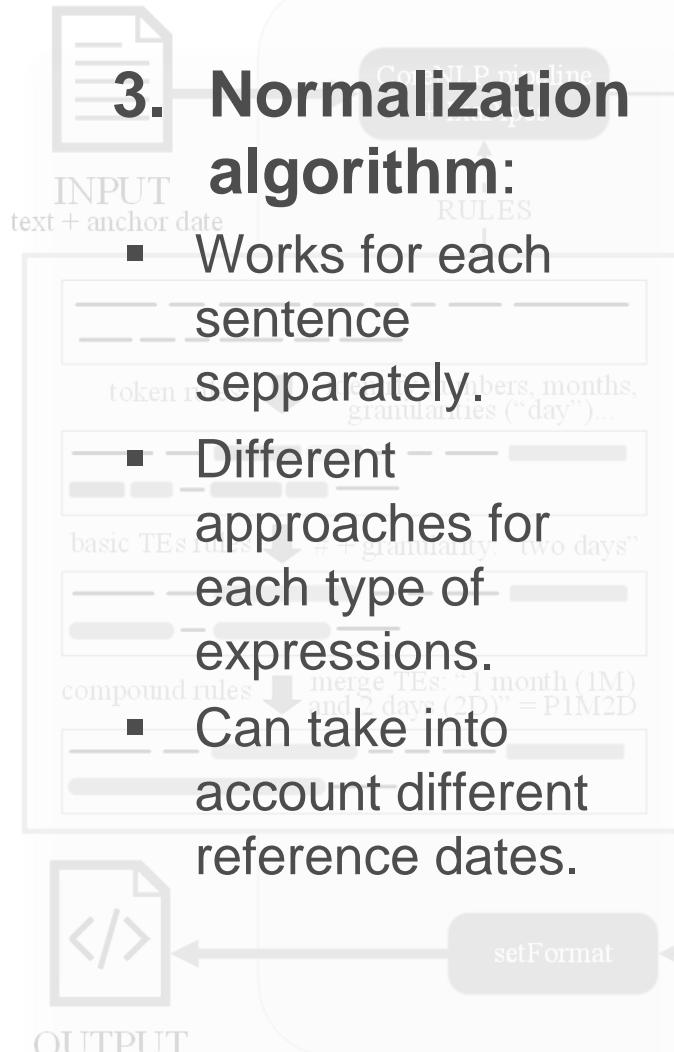
ANNOTADOR

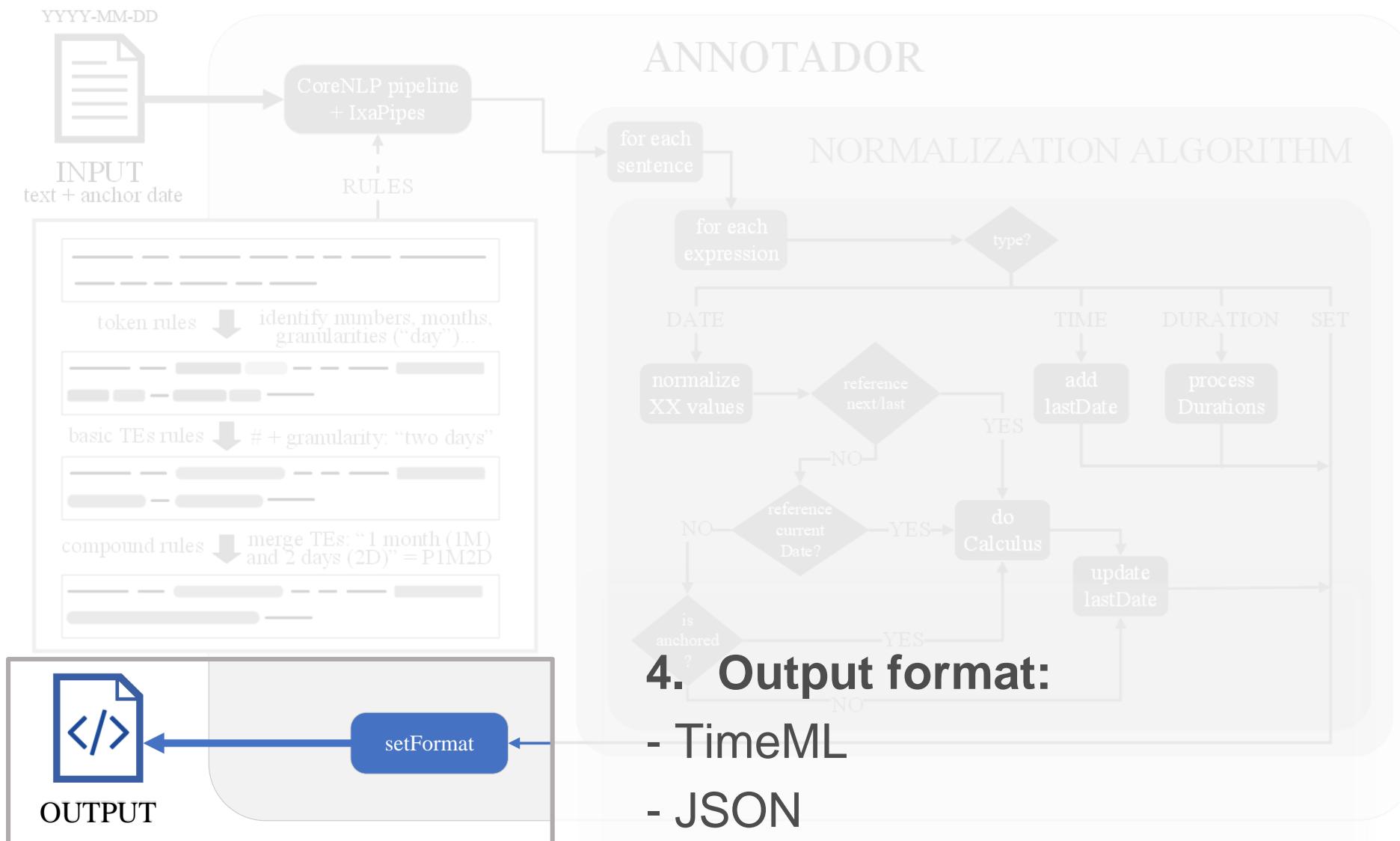
2. Rules:

More than 100 rules written in CoreNLP TokensRegex format.

- Token-based rules** for expressions such as numerals, granularities...
- Basic temporal expression rules**, working on previously found basic expressions
- Compound expression rules**, for inheritance values or composition.
- Literal expression rules**, for specific expressions.

YYYY-MM-DD





- **Format**-processing is covered (e.g. to avoid links).
- New tags for **intervals** and a new option to normalize the **granularity business days** (BD).
- Additionally, some frequent temporal expressions that are not used in this sense, but are frequent **taglines in the legal domain** (such as “now”), were omitted.
- Most temporal taggers were not able to identify **years when written with letters**, in particular for Spanish.
- **Composed DURATIONs** (e.g. “one day and three hours”) are often annotated separately by previous temporal taggers.

Regarding Spanish, Añotador covers cases that state-of-the-art temporal taggers do not meet:

- The tricky case of the word “mañana”, frequent in Spanish:
 - “mañana” (*noun*) means “morning”.
 - “mañana” (*adv*) means “tomorrow”.
 - “pasado mañana” (*adv*) means “the day after tomorrow”.
 - “pasado” (*adv*) equals to “pasado mañana”.
 - “pasado” (*noun or adjective*) means “past” (*noun or adj*).
 - Also syntactically: “por la mañana” vs “en la mañana” (“in the morning”)
- Different registers (*antaño*, a cultured way to say “in the past”), colloquial expressions.
- Latin American expressions, that have never been considered in previous temporal taggers.

Use Case 1 Lynx Project

- Añotador processed different types of legal texts in Spanish and English.

Use Case 2 Collaboration with CENDOJ (Judicial Documentation Centre)



Some of the main changes resulting from the feedback provided by CENDOJ:

- Misleading times/duration mentions (e.g. “A las 23 horas”, meaning “At 11pm”).
- Document Structure Extractor and Intervals added.
- References to the present time without temporal intention not annotated.
- More robustness against misleading capital letters and money symbols.

Use Case 3 As an occasional service within other tasks

- Terminology Extraction: to avoid dates to be included as relevant words (TermitUp).
- Translation Memory Matching: to detect dates (Ranasinghe et al., 2020).
- Anonymization: in a national project in order to delete dates from texts.

User validation

- Used in the Lynx project for both legal English and Spanish.
- Refined by CENDOJ for legal Spanish.
- Used as an API for several NLP tasks.
- Demo of Añotador freely accessible to any user; general users were asked to test the tool and report the main problems they found.

Corpora Evaluation

- Different aspects of temporal expressions covered:
 - Extension, normalization, type of temporal expression.
- These metrics will be considered:
 - Lenient: a partially tagged expression is considered a hit, even if not all its extent is marked by the tagger.
 - Strict: just expressions tagged exactly as in the test are considered correct.
 - Average: average of lenient and strict.

Hourglass

TempEval-2

Tagger	Attribute	strict			lenient			average		
		P	R	F1	P	R	F1	P	R	F1
Añotador	value	0.72	0.71	0.72	0.80	0.78	0.79	0.76	0.74	0.75
	type	0.79	0.77	0.78	0.89	0.87	0.88	0.84	0.82	0.83
	extent	0.83	0.82	0.82	0.95	0.92	0.94	0.89	0.87	0.88
Heidel	value	0.57	0.48	0.52	0.64	0.53	0.58	0.60	0.51	0.55
	type	0.61	0.51	0.55	0.82	0.69	0.75	0.72	0.60	0.65
	extent	0.62	0.52	0.57	0.87	0.73	0.80	0.75	0.63	0.68
SUTime	value	0.30	0.08	0.13	0.45	0.12	0.19	0.38	0.10	0.16
	type	0.47	0.13	0.20	0.80	0.21	0.34	0.64	0.17	0.27
	extent	0.47	0.13	0.20	0.89	0.24	0.37	0.68	0.18	0.29
Añotador	value	0.80	0.78	0.79	0.83	0.80	0.82	0.82	0.79	0.80
	type	0.84	0.82	0.83	0.91	0.88	0.89	0.88	0.85	0.86
	extent	0.87	0.84	0.85	0.93	0.90	0.92	0.90	0.87	0.89
Heidel	value	0.84	0.75	0.80	0.86	0.77	0.82	0.85	0.76	0.81
	type	0.85	0.76	0.81	0.89	0.79	0.84	0.87	0.78	0.82
	extent	0.90	0.81	0.85	0.94	0.84	0.89	0.92	0.83	0.87
SUTime	value	0.64	0.22	0.33	0.83	0.29	0.43	0.73	0.26	0.38
	type	0.65	0.23	0.34	0.93	0.32	0.48	0.79	0.28	0.41
	extent	0.67	0.23	0.35	0.96	0.33	0.49	0.81	0.28	0.42

TempCourt - ECHR

StandardTimeML
 LegalTimeML

	A	lenient			strict			lenient + value			strict + value		
		P	R	F1	P	R	F1	P	R	F1	P	R	F1
AN	0.98	0.96	0.97	0.94	0.93	0.93	0.91	0.89	0.90	0.88	0.87	0.87	0.87
	0.87	0.97	0.92	0.83	0.93	0.88	0.81	0.90	0.85	0.77	0.86	0.81	
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	0.78
	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	0.75
	0.76	0.85	0.80	0.71	0.80	0.76	0.66	0.74	0.79	0.64	0.72	0.68	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	0.85
	0.84	0.92	0.88	0.83	0.92	0.87	0.74	0.82	0.78	0.74	0.82	0.78	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
	0.75	0.85	0.80	0.70	0.79	0.74	0.65	0.74	0.69	0.64	0.72	0.67	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
	0.86	0.93	0.90	0.70	0.76	0.73	0	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	0.85
	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
	0.65	0.62	0.64	0.61	0.58	0.60	0	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	

TempCourt - ECJ

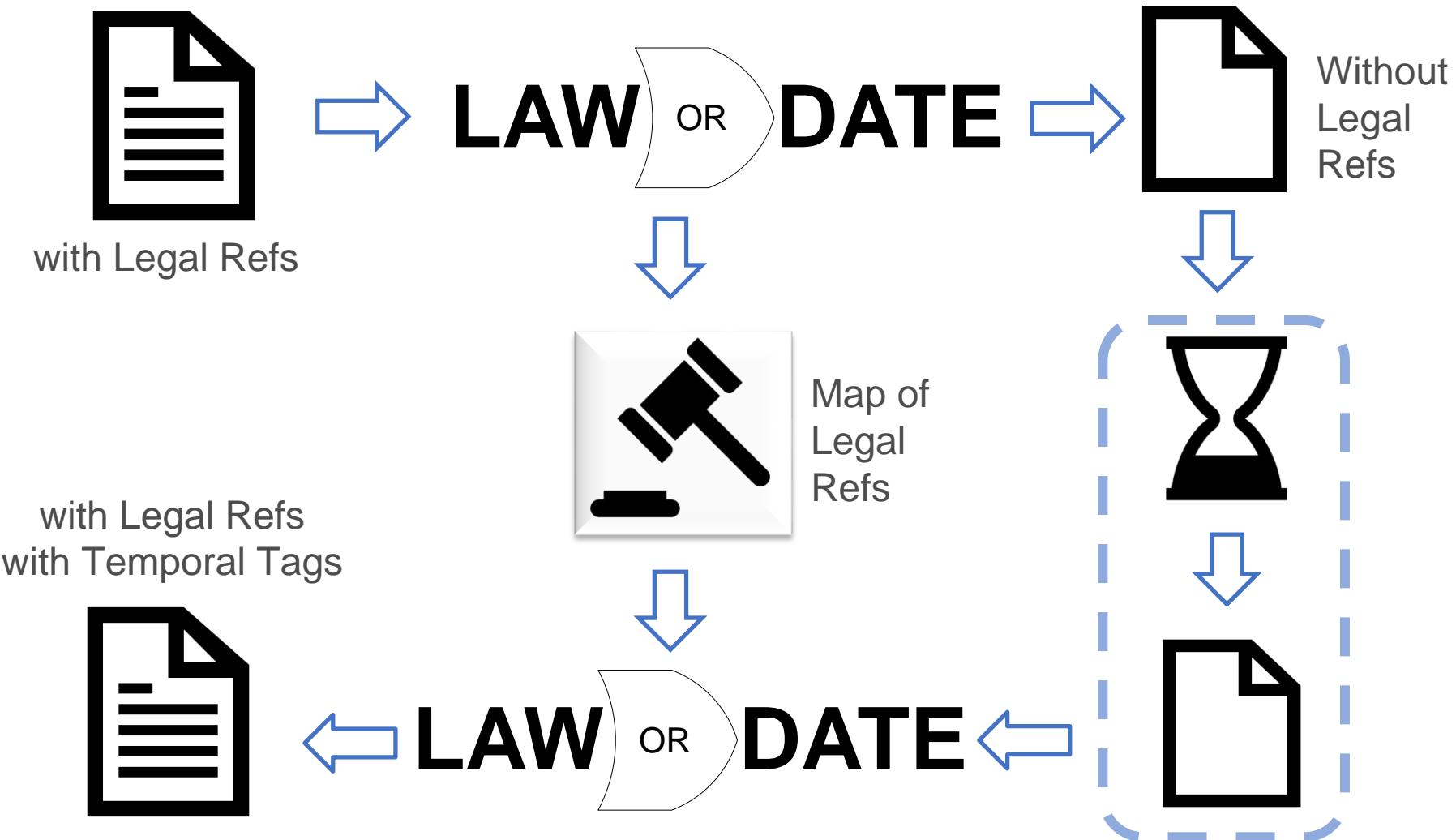
StandardTimeML
 LegalTimeML

A	lenient			strict			lenient + value			strict + value		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
AN	0.98	0.94	0.96	0.96	0.92	0.94	0.96	0.92	0.94	0.94	0.90	0.92
	0.54	0.94	0.69	0.52	0.91	0.67	0.53	0.92	0.67	0.51	0.89	0.65
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 tagging

LAWORDATE

Idea Patters to “clean” misleading expressions for temporal taggers



FINAL TEXT

Your original text with real temporal annotations provided by state-of-the-art temporal tagger as [HeidelTime](#) after our preprocessing

```
<?xml version="1.0"?>
<!DOCTYPE TimeML SYSTEM "TimeML.dtd">
<TimeML>
An example: "En la presente base de datos se recogen los elementos
inscritos en el registro creado via el Real Decreto 2093/2008 de 19 de
diciembre. Ha sido actualizado por ultima vez <TIMEX3 tid="t3"
type="DATE" value="2017-08-13">el 13 de agosto de 2017</TIMEX3>."
</TimeML>
```

MAP OF REPLACEMENTS

The replacements done by LawORDate before applying [HeidelTime](#) are the following:

Replacement	Original
RefRealDecreto	Real Decreto 2093/2008, de 19 de diciembre

ALTERNATIVE FINAL TEXT

Without our LawORDate preprocessing, the result by [HeidelTime](#) would have been:

```
<?xml version="1.0"?>
<!DOCTYPE TimeML SYSTEM "TimeML.dtd">
<TimeML>
An example: "En la presente base de datos se recogen los elementos
inscritos en el registro creado via el Real Decreto <TIMEX3 tid="t1"
type="DATE" value="2093">2093</TIMEX3><TIMEX3 tid="t2"
type="DATE" value="2008">2008</TIMEX3>, <TIMEX3 tid="t4"
type="DATE" value="2008-12-19">de 19 de diciembre</TIMEX3>. Ha
```

Result of Heideltime using LawORDate:

- It does not annotate Real Decreto as a date...
- ... because it is in the map of LawORDate

Result of Heideltime without using LawORDate:

- Annotates the different parts of Real Decreto as a date...
- ... so the annotations are not correct.

Temporal Expression-related contributions

H1.a

An analysis of temporal tagging of the legal domain.

H1.a

H2.a

Corpus of legal decisions in English.

Corpus of short texts in Spanish to systematically test temporal taggers.

H1.b

A tool that allows the user to preprocess citations that can be misleading to temporal taggers.

H1.a

H1.b

A temporal tagger for Spanish and English that

- 1) covers untackled particularities of the Spanish language,
- 2) has a special implementation for the legal domain

- Corpus
- Extraction
- Resources

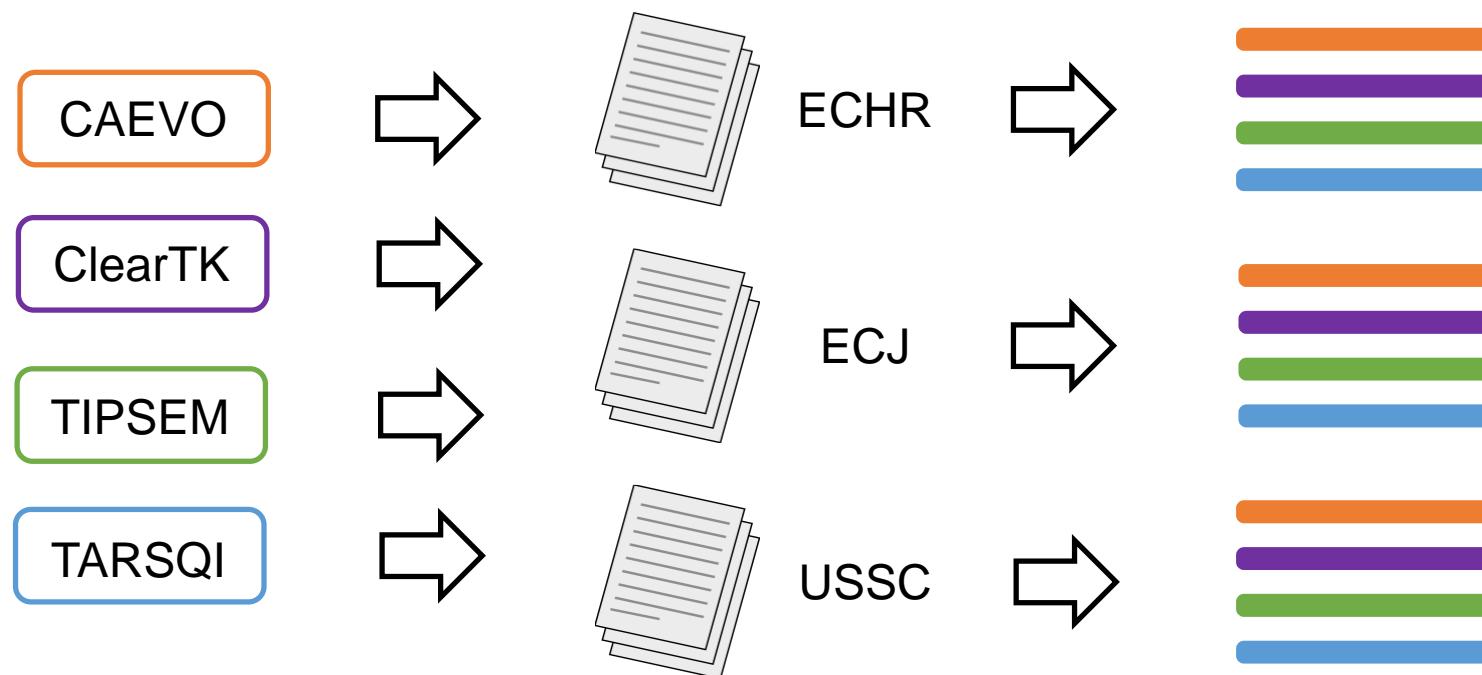
EVENTS



WHO HOW WHAT
WHEN WHERE
WHY

- Is it clear what is an event?
 - How many events are there?
 - What is annotated as event?
 - Is there agreement?

- Test in TempCourt



How many events are there?

Between the 7 and the 9% of the tokens in a document (12 and 13% in the case of TARSQI) are considered events!

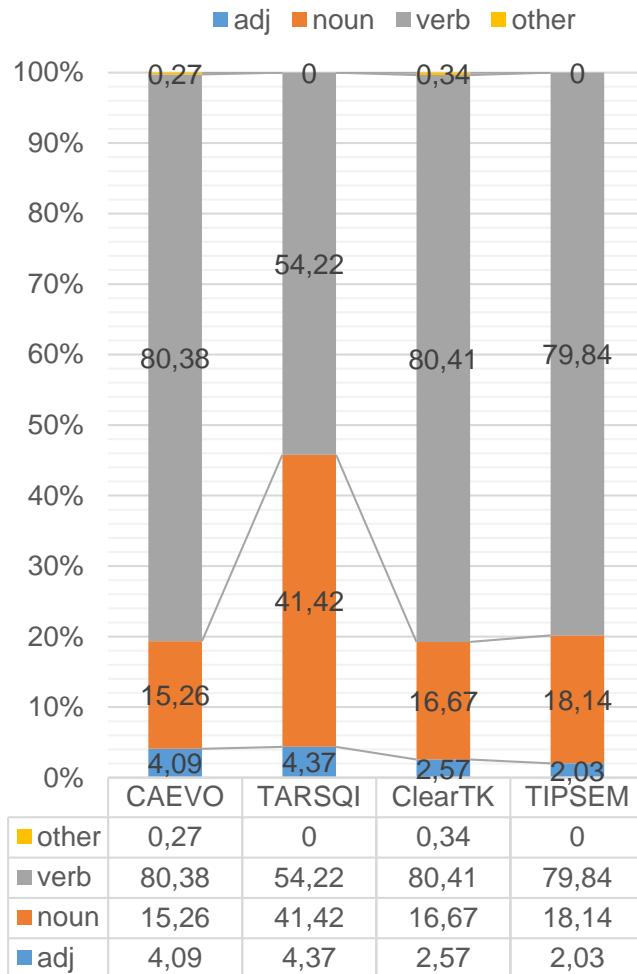
Calculus per sentence:

- 0,99~1,59 events per sentence for the ECHR part
- 2,26~4 events per sentence for ECJ
- 1,43~2,31 events per sentence for USSC

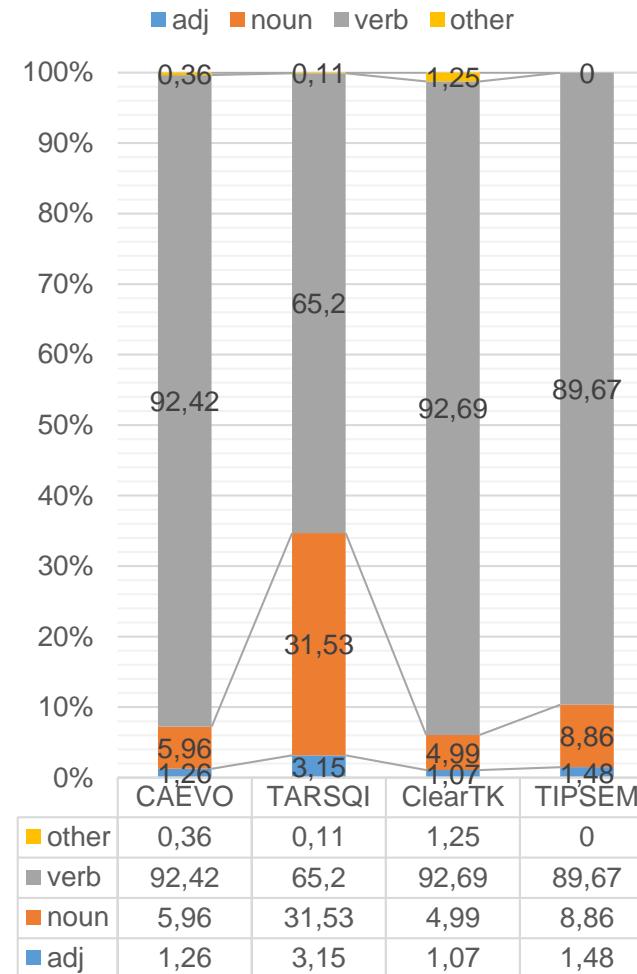
On average every sentence has at least one event, and even more in the case of documents with longer sentences.

What is annotated as event?

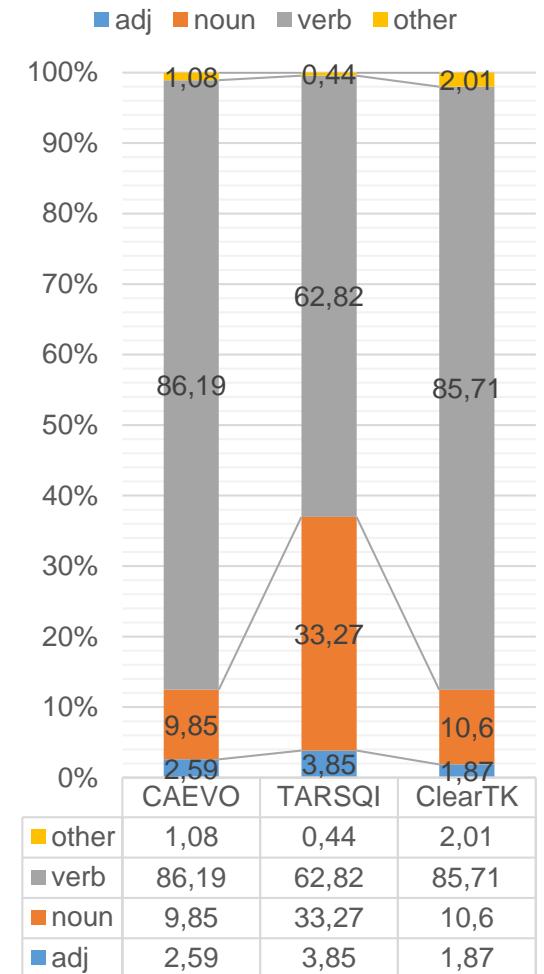
ECJ



ECHR



USSC



Is there agreement?

ECHR

	CAEVO	CLEARTK	TARSQI	TIPSEM
CAEVO		0,76	0,56	0,72
CLEARTK	0,76		0,58	0,79
TARSQI	0,56	0,58		0,55
TIPSEM	0,72	0,79	0,55	

ECJ

	CAEVO	CLEARTK	TARSQI	TIPSEM
CAEVO		0,69	0,48	0,69
CLEARTK	0,69		0,50	0,73
TARSQI	0,48	0,50		0,52
TIPSEM	0,69	0,73	0,52	

USSC

	CAEVO	CLEARTK	TARSQI
CAEVO		0,25	0,52
CLEARTK	0,25		0,17
TARSQI	0,52	0,17	

Temporal Dimensions

- ▶ ■ Temporal dimension of the case
- ▶ ▶ ■ Temporal dimension of the legal process (national or not)
- Temporal dimension of the legal context and other events

▶ Mr Buivids made a video recording in a station of the Latvian national police

▶ Mr Buivids published the recorded video which showed police officers, on the internet site www.youtube.com

The National Data Protection Agency requested Mr Buivids to remove that video from the internet site www.youtube.com and from other websites

▶ Mr Buivids brought an action before the administratīvā rajona tiesa (District Administrative Court, Latvia) claiming compensation

▶ The administratīvā rajona tiesa (District Administrative Court, Latvia) dismissed the action

▶ The Administratīvā apgabaltiesa (Regional Administrative Court, Latvia) dismissed the appeal brought by Mr Buivids against the decision of the administratīvā rajona tiesa (District Administrative Court)

▶ Mr Buivids filed an appeal in cassation before the referring court, the Augstākā tiesa (Supreme Court, Latvia), against the judgment of the Administratīvā apgabaltiesa (Regional Administrative Court)

▶ In those circumstances, the Augstākā tiesa (Supreme Court) decided to stay the proceedings and to refer the following questions to the Court of Justice for a preliminary ruling

▶ REQUEST for a preliminary ruling from the Augstākā tiesa (Supreme Court, Latvia) to the ECJ

▶ FINAL DECISION by the ECJ (Second Chamber)

2013

2014

2015

2016

2017

2018

2019

2020

Conclusions from analysis:

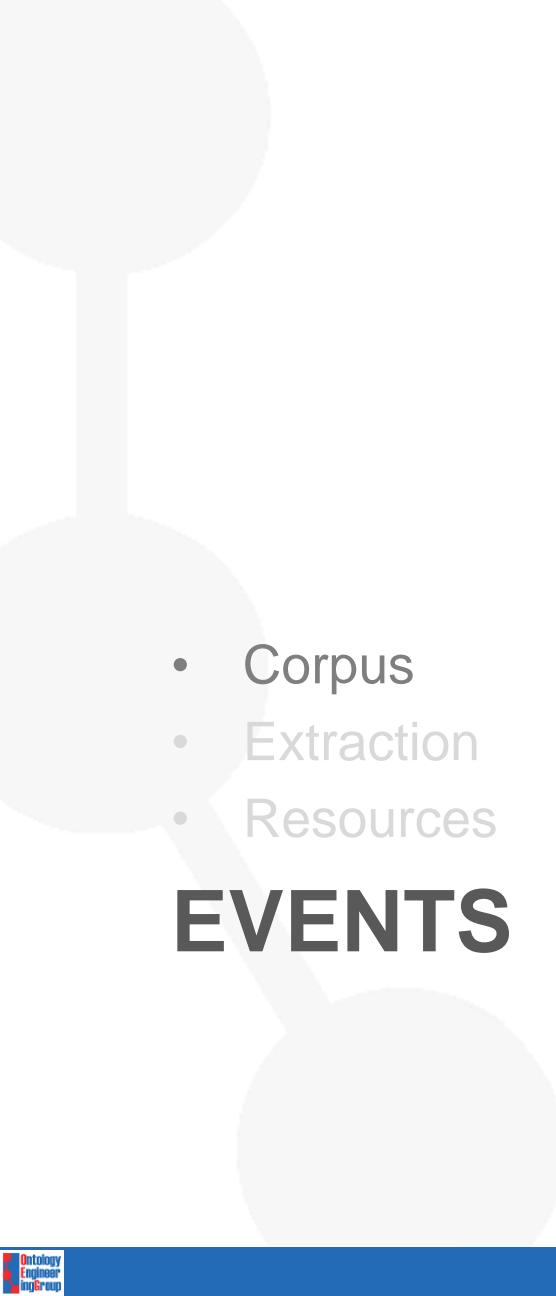
- High variability of what is an event.
- Too many events! Preferably only **relevant** ones.
- Not all the events are recognized, importance of light verbs (e.g. request a preliminary **ruling**)
- TimeML just covers time-related information.

Creation of the EventsMatter corpus.

Filtz, E., **Navas-Loro, M.**, Santos, C., Polleres, A., and Kirrane, S. (2020). "EventsMatter: Extraction of Events from Court Decisions". *JURIX2020, Frontiers in Artificial Intelligence and Applications. Volume 334: Legal Knowledge and Information Systems, p.33-42, 2020*

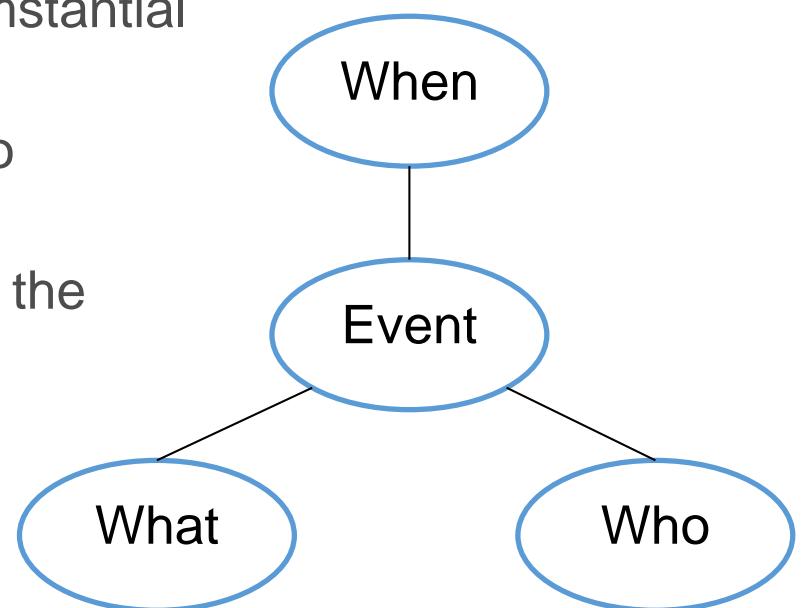
- Corpus
- Extraction
- Resources

EVENTS



WHO HOW WHAT
WHEN WHERE
WHY

- 30 documents from the European Court of Human Rights.
- Annotation of event-what-when-who by two experts.
- Two types of events: procedural and circumstances.
- Several annotation rounds and discussion on tricky cases and relevance; main difficulties:
 - Events can be procedural or circumstantial depending on the case.
 - “Legal verbs” do not always refer to procedural events.
 - Nested events, where one event is the “when” of another.
 - Factuality of the event.
 - Relevance of the event.
 - Annotation of repeated events.



- Corpus
- Extraction
- Resources

EVENTS

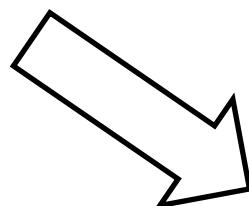


WHO HOW WHAT
WHEN WHERE
WHY

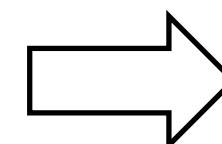
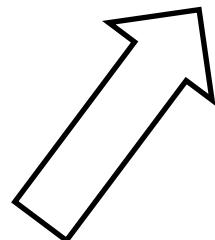
Events Extraction

CONTRACTFRAMES

Legal cases



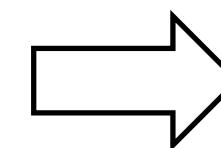
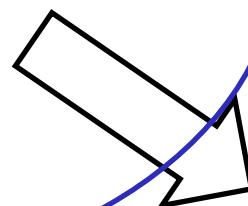
Laws



Knowledge



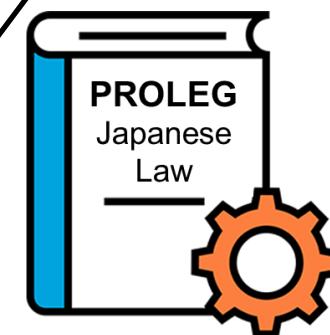
Legal cases



Knowledge

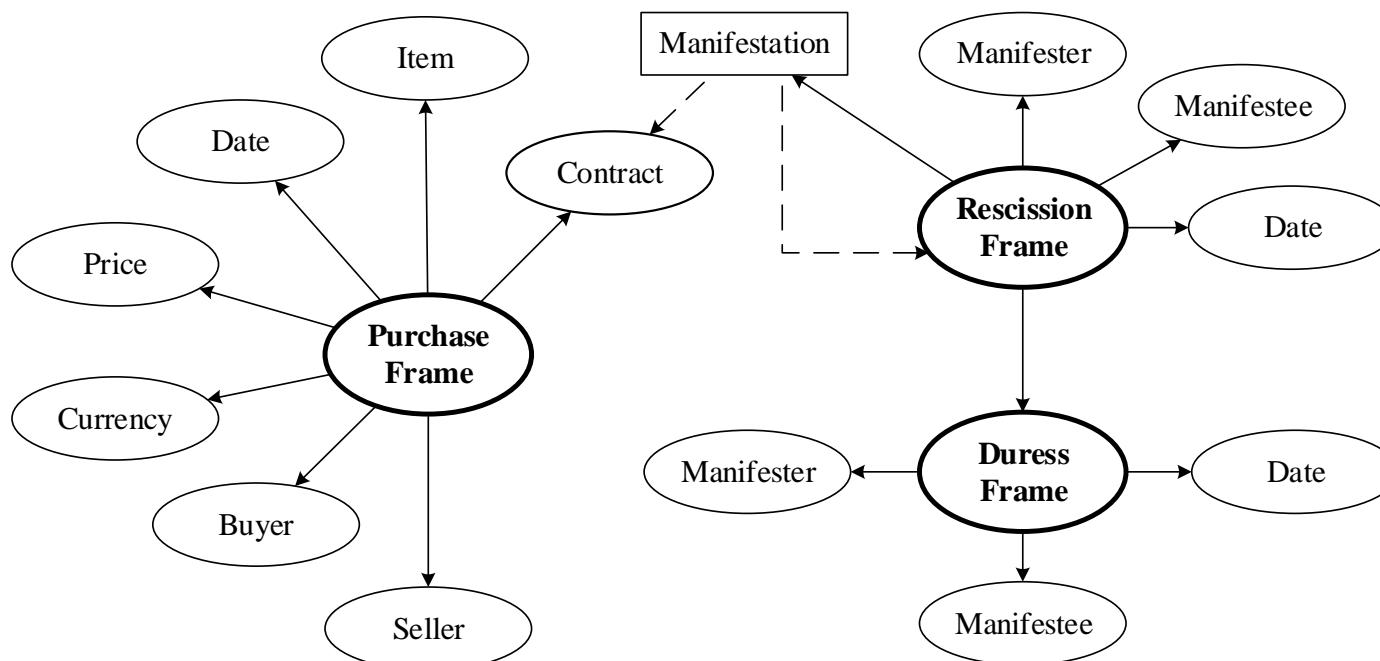


Laws



ContractFrames - Modelization of the events

- We build **frames** for representing the different events that can affect to the status of a contract.
- Output them as PROLEG facts.



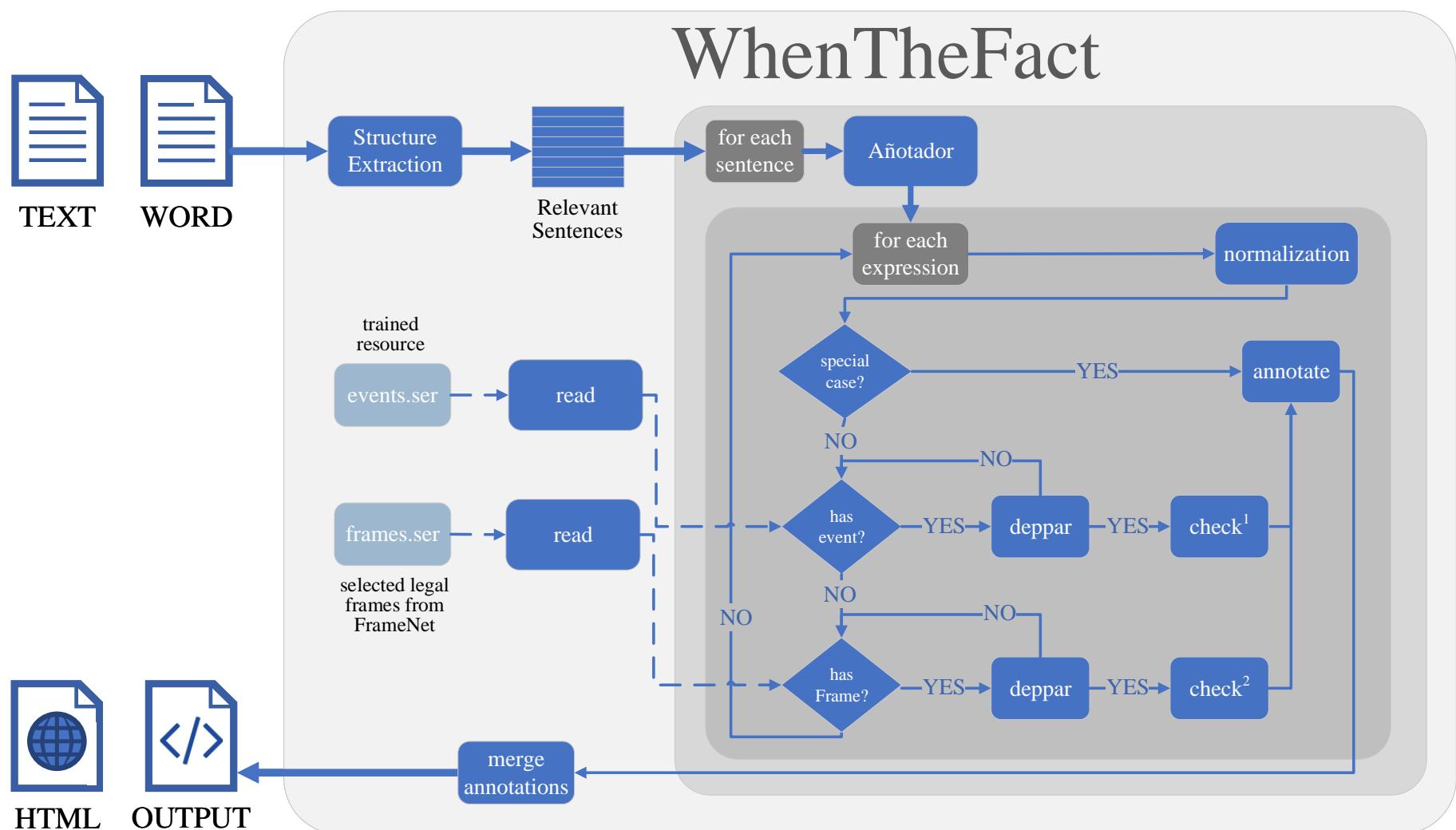
Navas-Loro, M., Satoh, K., Rodríguez-Doncel, V. (2018). "Contract-Frames: Bridging the gap between natural language and logics in contract law". *JSAI International Symposium on Artificial Intelligence*. Springer, Cham, 2018.

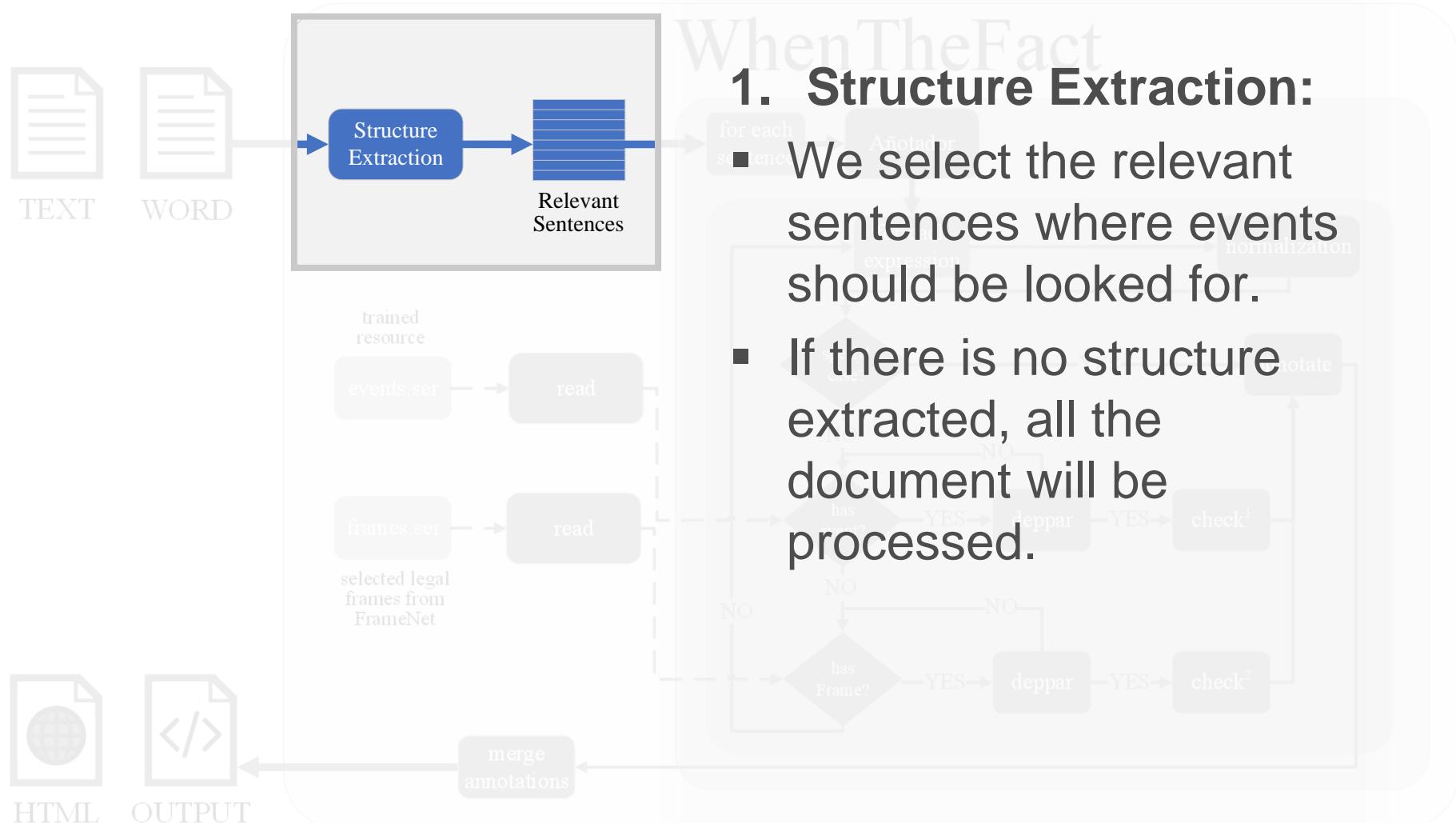
Events Extraction

WHEN THE FACT

Objective: Find relevant events in European legal texts.

- Finds who-when-what and the type of event (procedure/circumstance).
- Uses Structure Extraction and semantic similarity for relevance.
- Different training strategies:
 - Automatically from EventsMatter.
 - Manually hierarchy-based selection of legal-related frames from FrameNet.
- Event extraction based on Añotador structure.
- Checked auxiliary verbs, complements for light verbs...





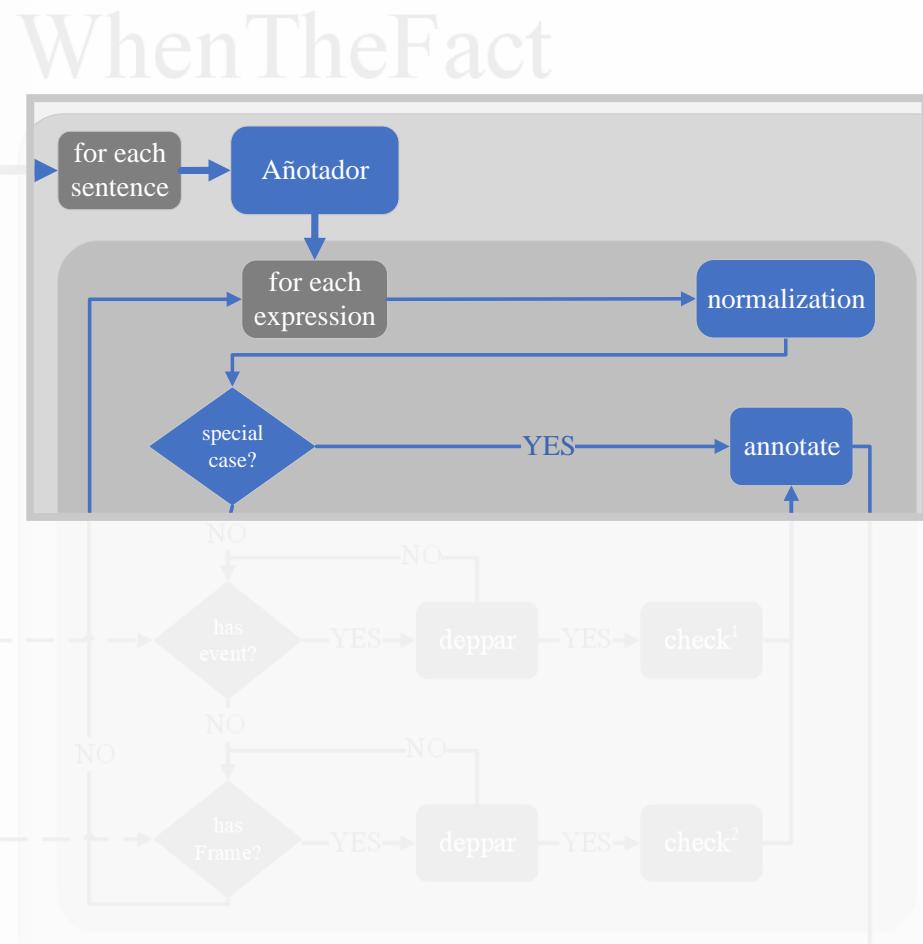
1. Structure Extraction:

We select the relevant sentences where events should be looked for.

- If there is no structure extracted, all the document will be processed.

2. Añotador looks for dates.

- In the same iteration we look for events.
- We detect special cases (such as sentence structures that always appear in the same way in specific documents).

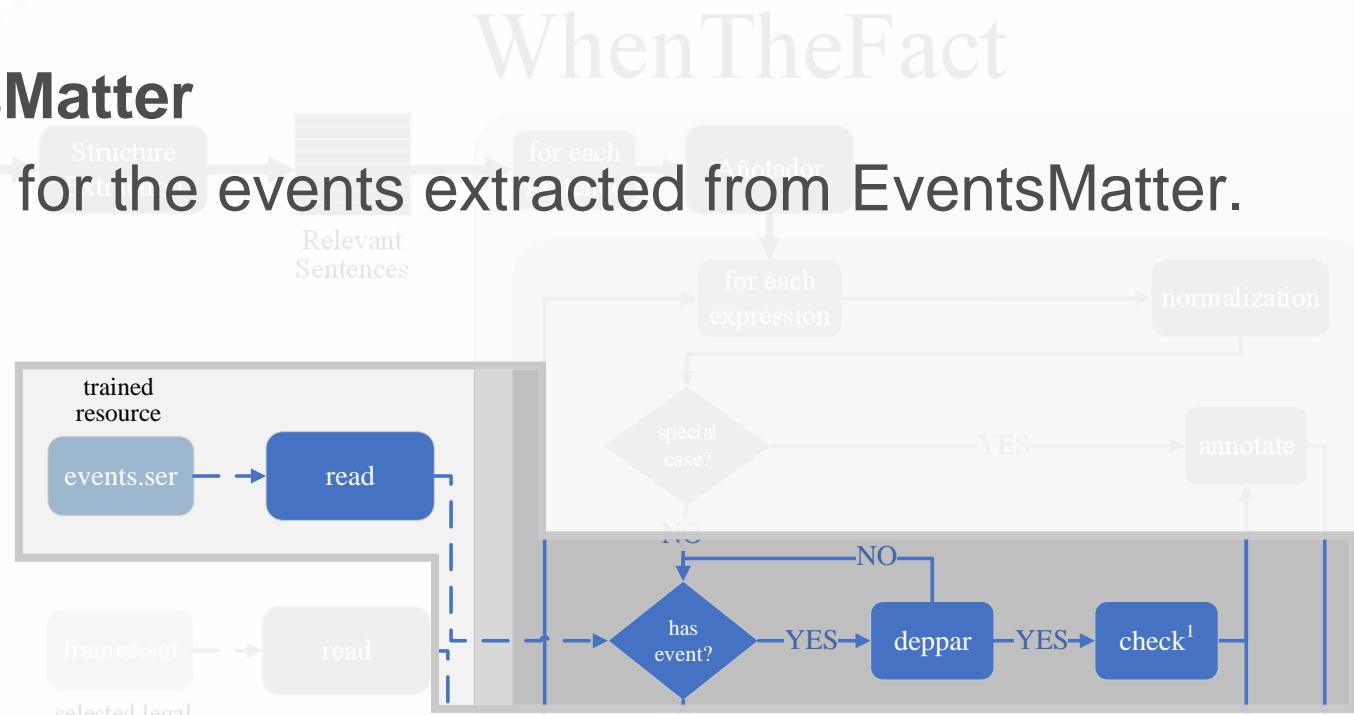


3. EventsMatter

- We look for the events extracted from EventsMatter.

TEXT

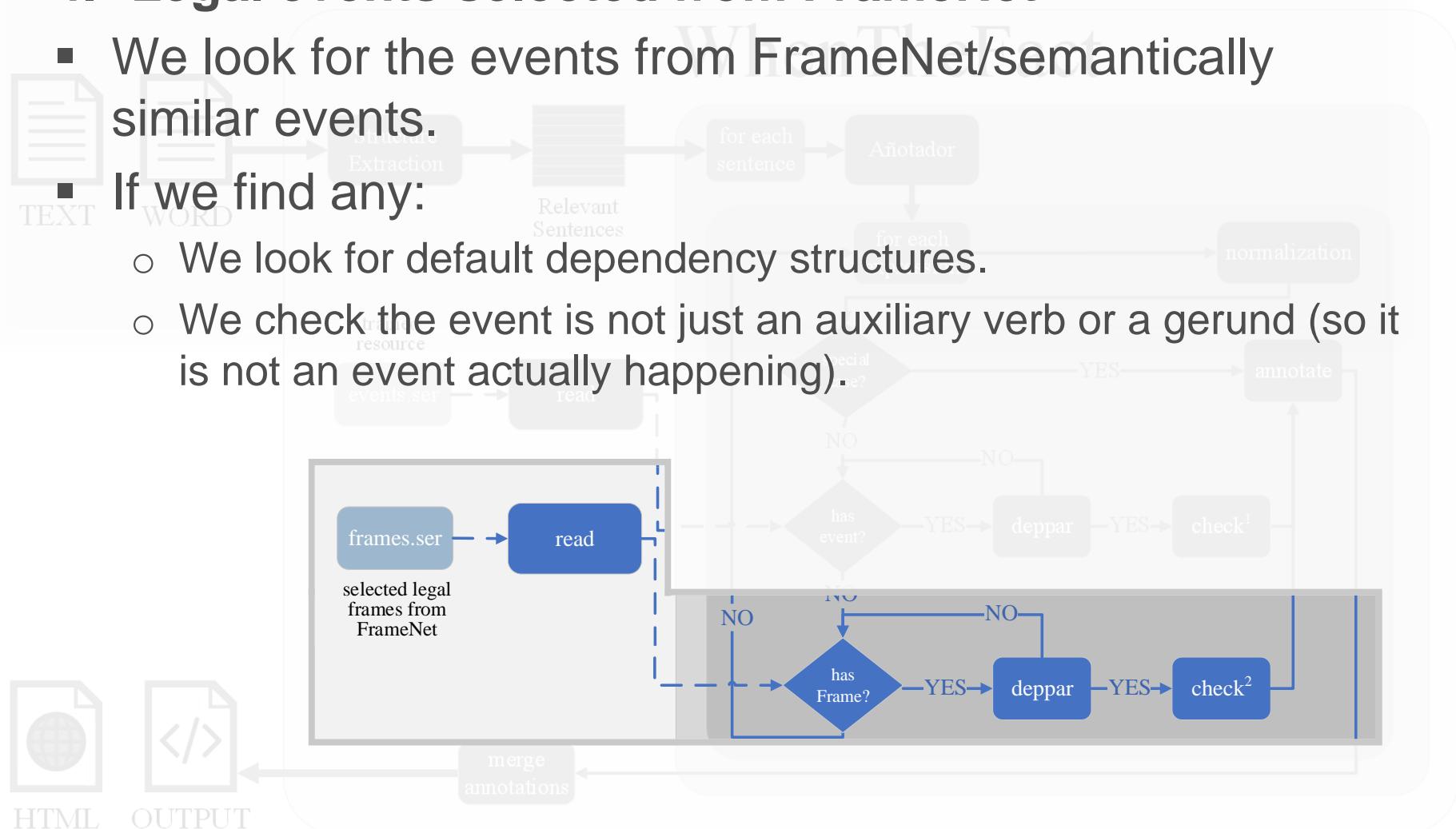
WORD



- If we find any:
 - We look for dependency structures previously marked as relevant in order to find who and what.
 - We check the event is not just an auxiliary verb or a gerund (so it is not an event actually happening).

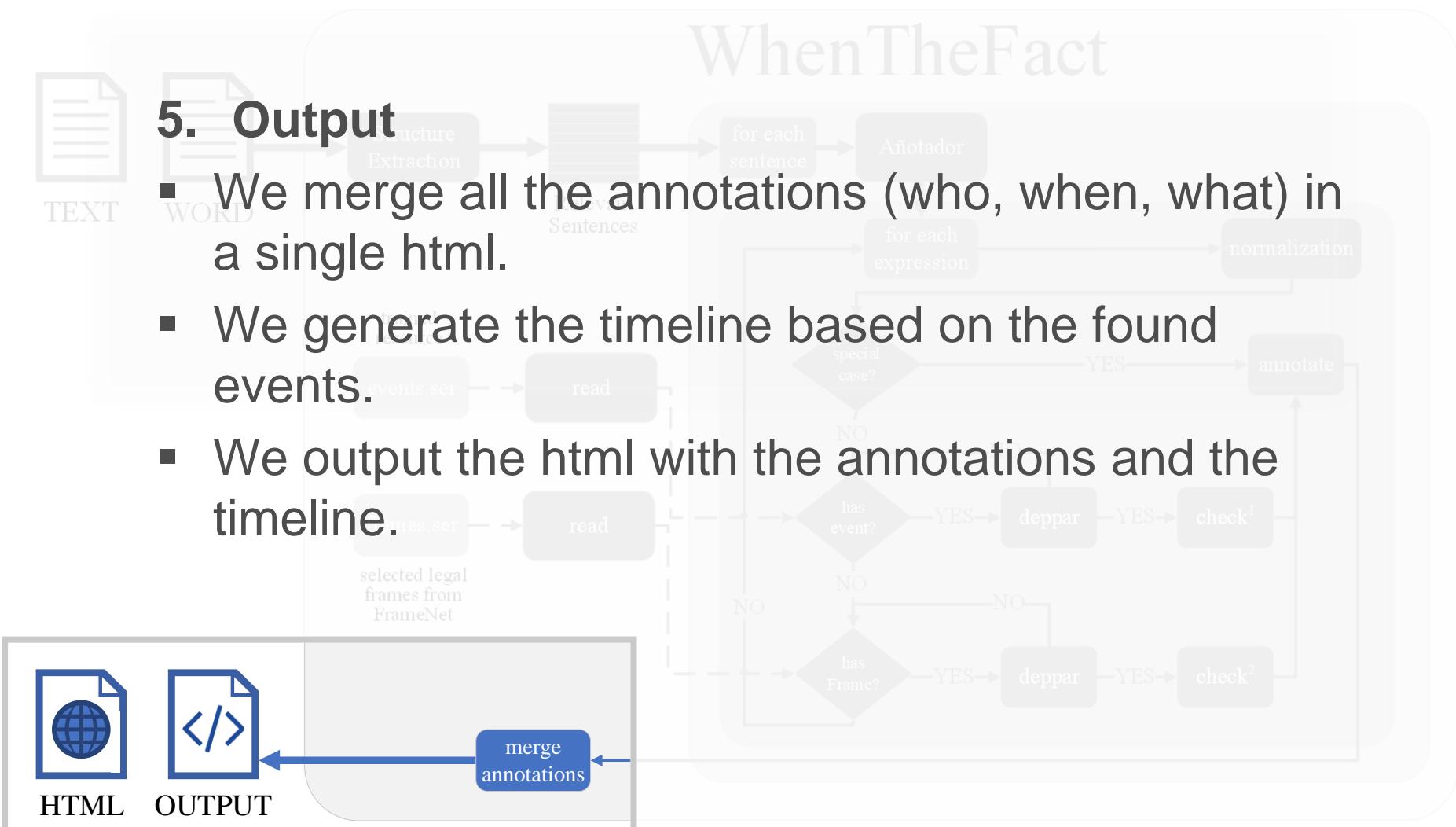
4. Legal events selected from FrameNet

- We look for the events from FrameNet/semantically similar events.
- If we find any:
 - We look for default dependency structures.
 - We check the event is not just an auxiliary verb or a gerund (so it is not an event actually happening).



5. Output

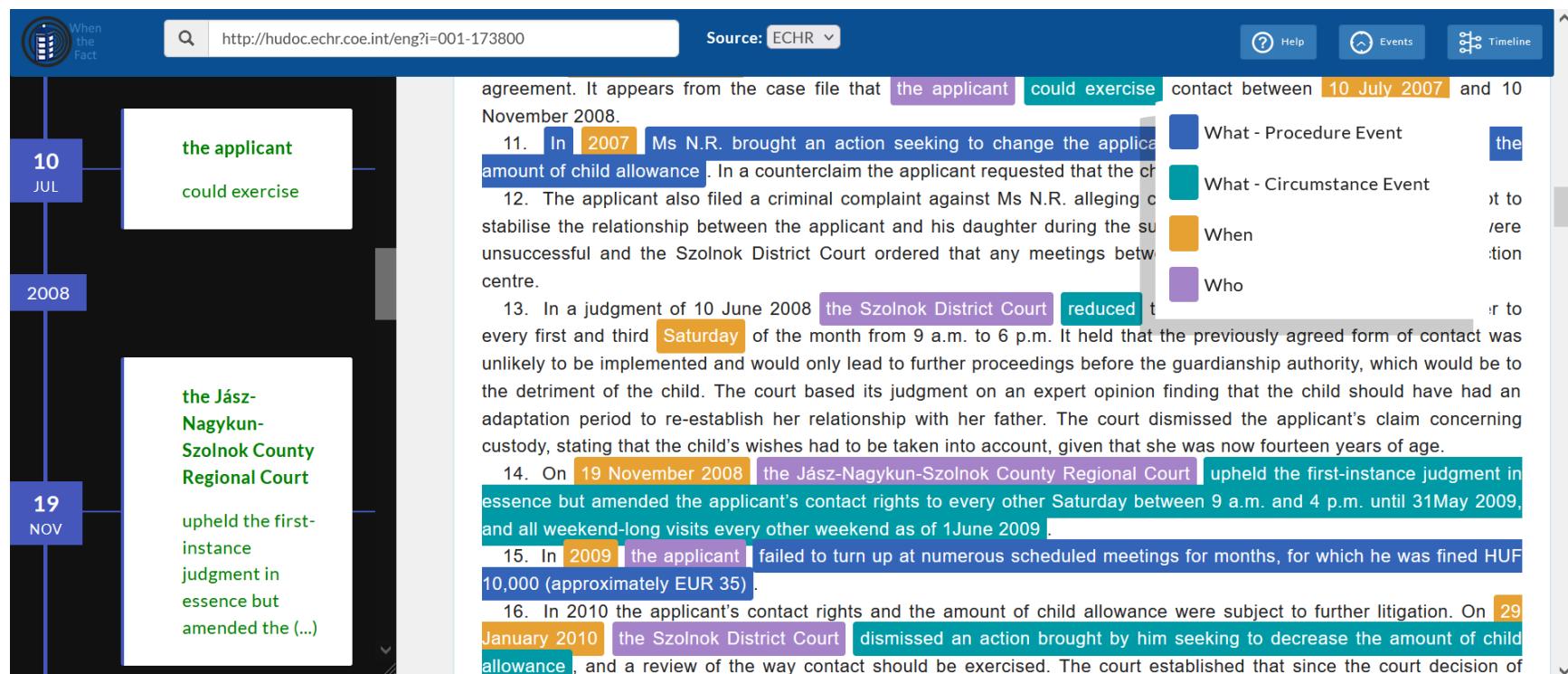
- We merge all the annotations (who, when, what) in a single html.
- We generate the timeline based on the found events.
- We output the html with the annotations and the timeline.



Demo available for ECHR and ECJ documents.

- A timeline is created based on the relevant events found.
- It allows to navigate through the document.

 WhenTheFact Source: ECHR



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 application amount of child allowance. In a counterclaim the applicant requested that the court stabilise the relationship between the applicant and his daughter during the unsuccessful and the Szolnok District Court ordered that any meetings between the parties should take place at the centre.

12. The applicant also filed a criminal complaint against Ms N.R. alleging that she had been physically abusing their daughter. The Szolnok District 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.

13. In a judgment of 10 June 2008 the Szolnok District Court reduced the amount of child allowance 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).

16. In 2010 the applicant's contact rights and the amount of child allowance were subject to further litigation. On 29 January 2010 the Szolnok District Court dismissed an action brought by him seeking to decrease the amount of child allowance, and a review of the way contact should be exercised. The court established that since the court decision of

- Corpus
- Extraction
- Resources

EVENTS



Who How
When ? What
Where Why

Event-related resources

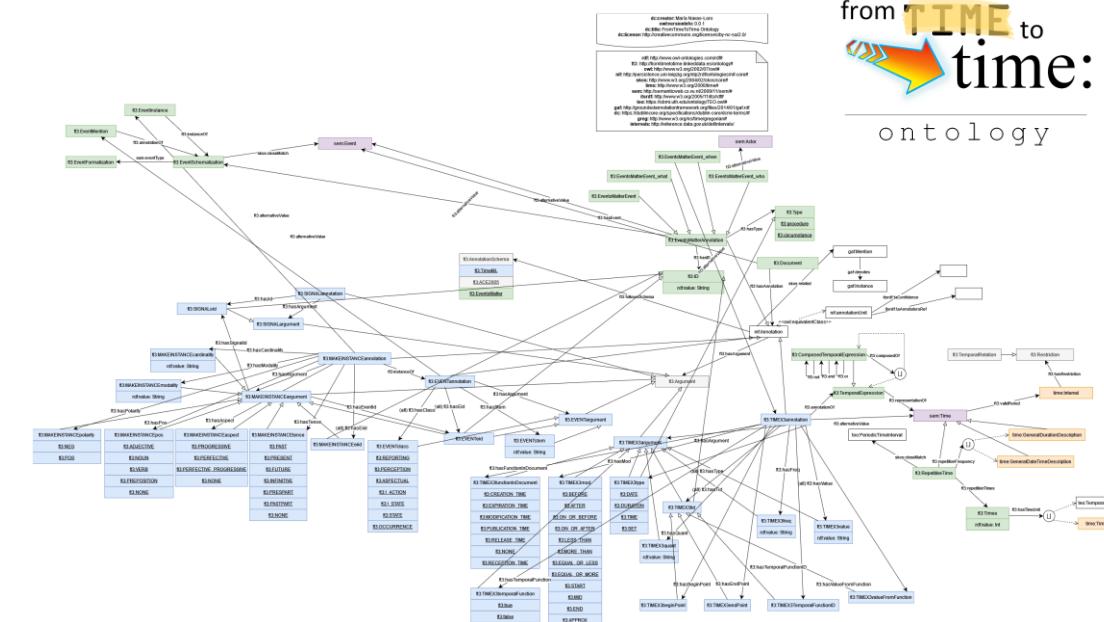
FT3 ONTOLOGY

Objectives

1. Event-based representation of information.
2. Representation of events and annotations for latter tasks.
3. Facilitate translation among time-related annotation formats and ontologies.

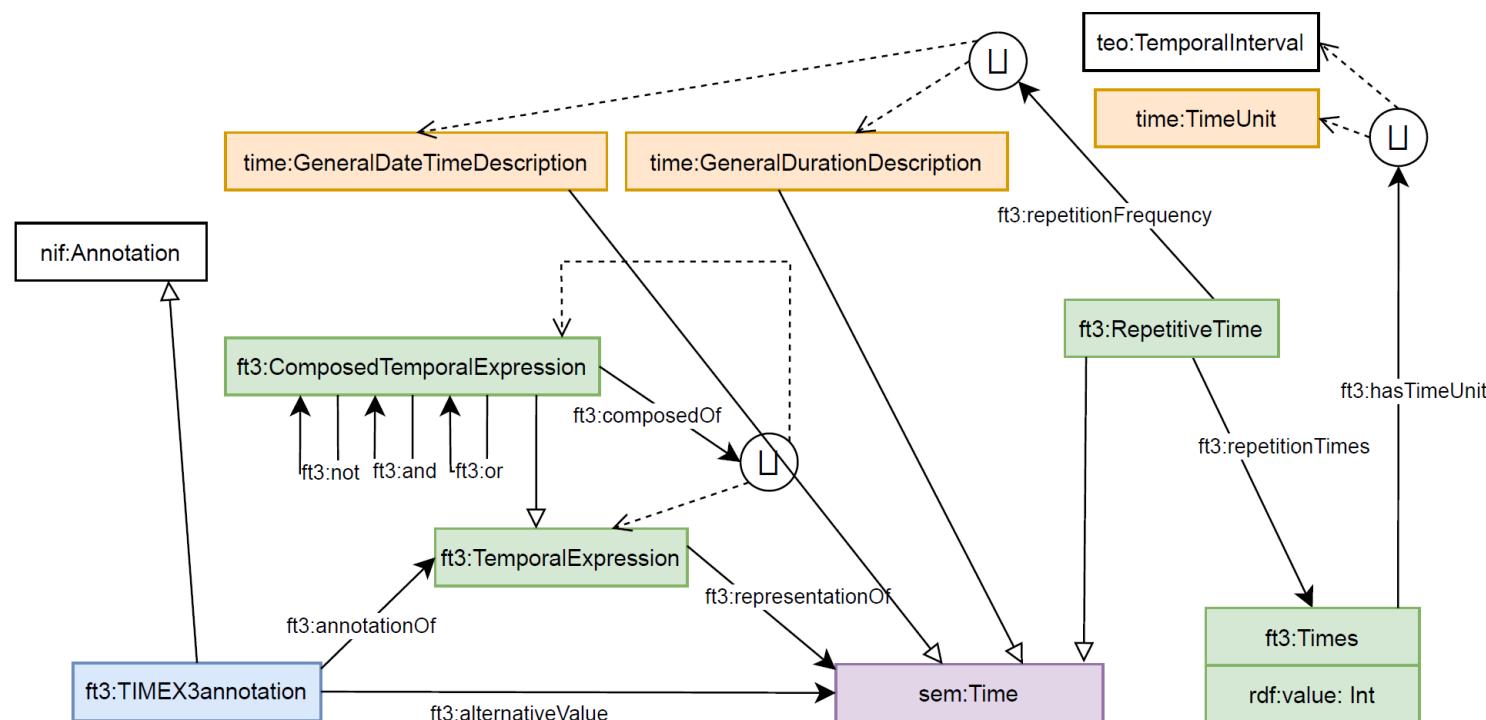
Main design decisions

- High level classes:
 - *ft3:Guidelines*
 - *ft3:Annotation*
 - *ft3:Argument*.
- Ontology reusage:
 - SEM
 - TEO
 - NIF
 - TIME

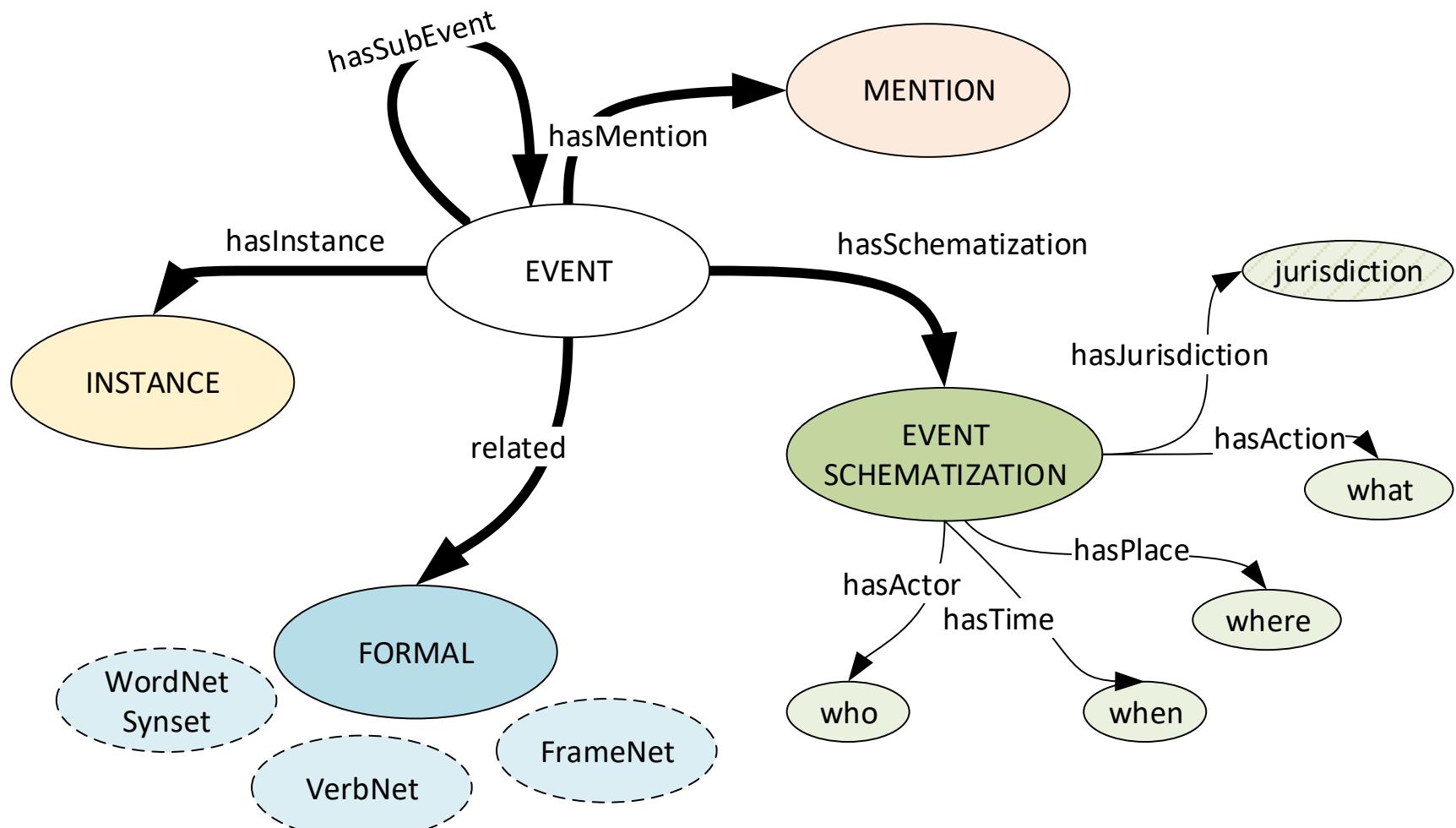


Temporal Expressions

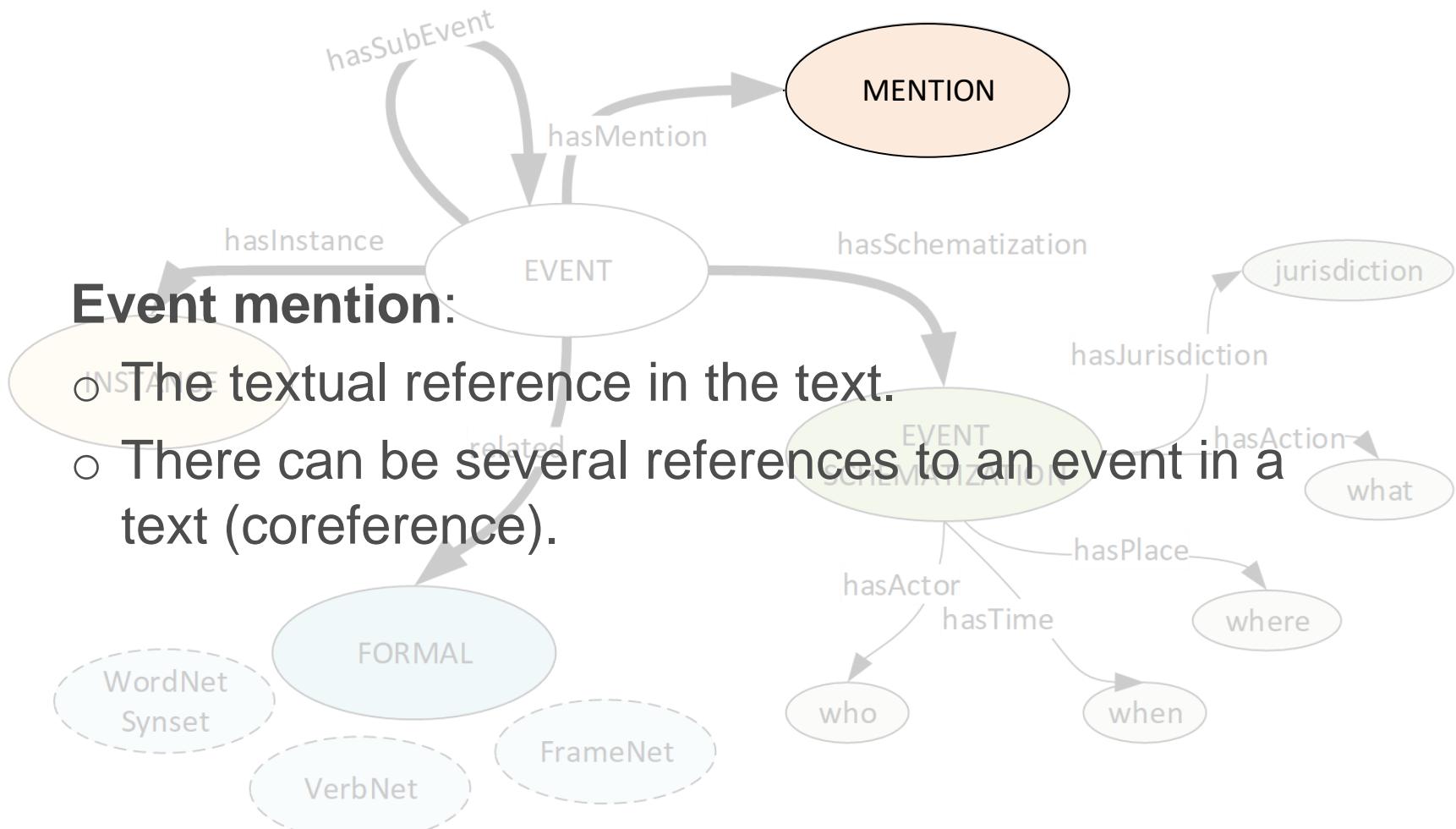
1. Expressivity: *ft3:ComposedTemporalExpressions*.
 2. Bridge between the annotation and the normalization.



Events distinction among different concepts



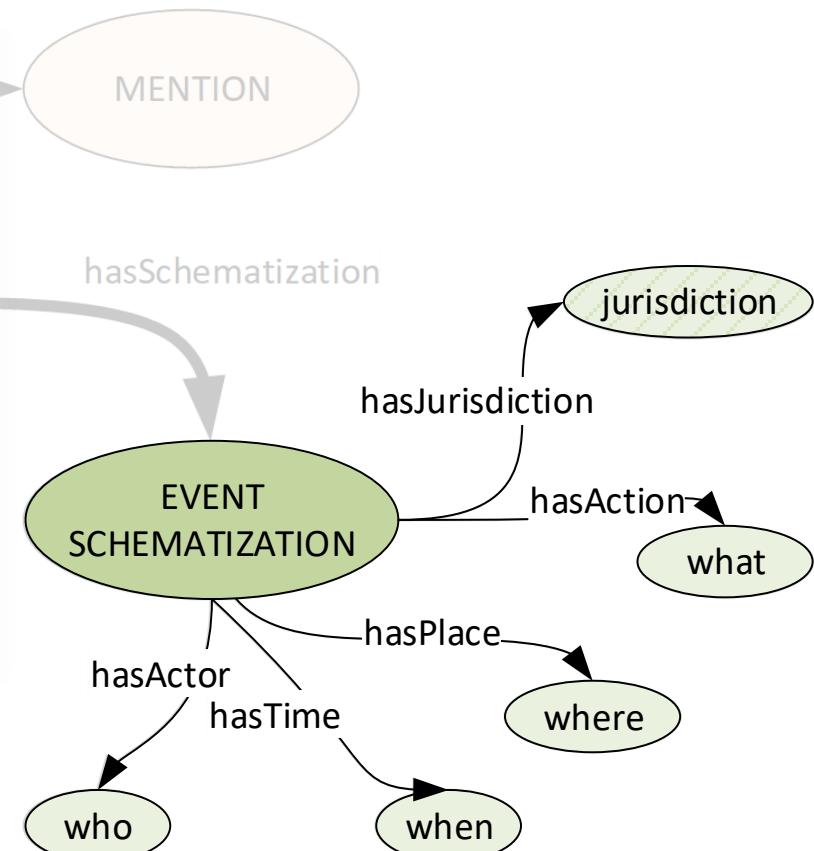
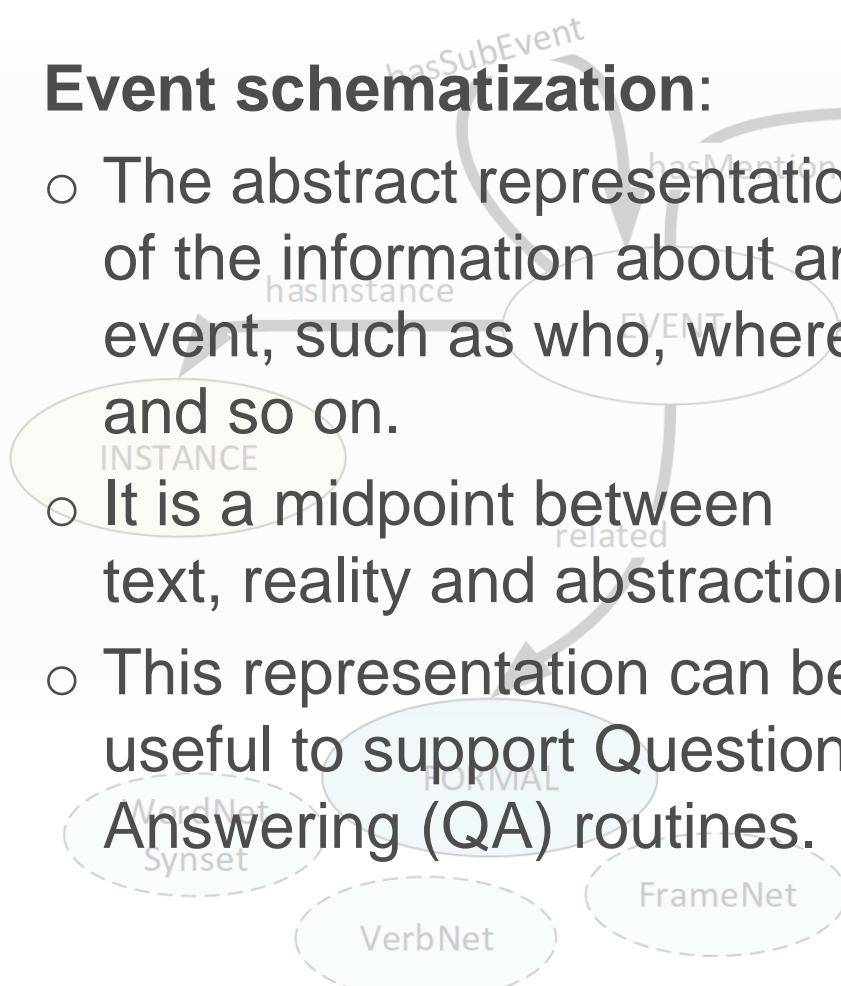
Events distinction among different concepts



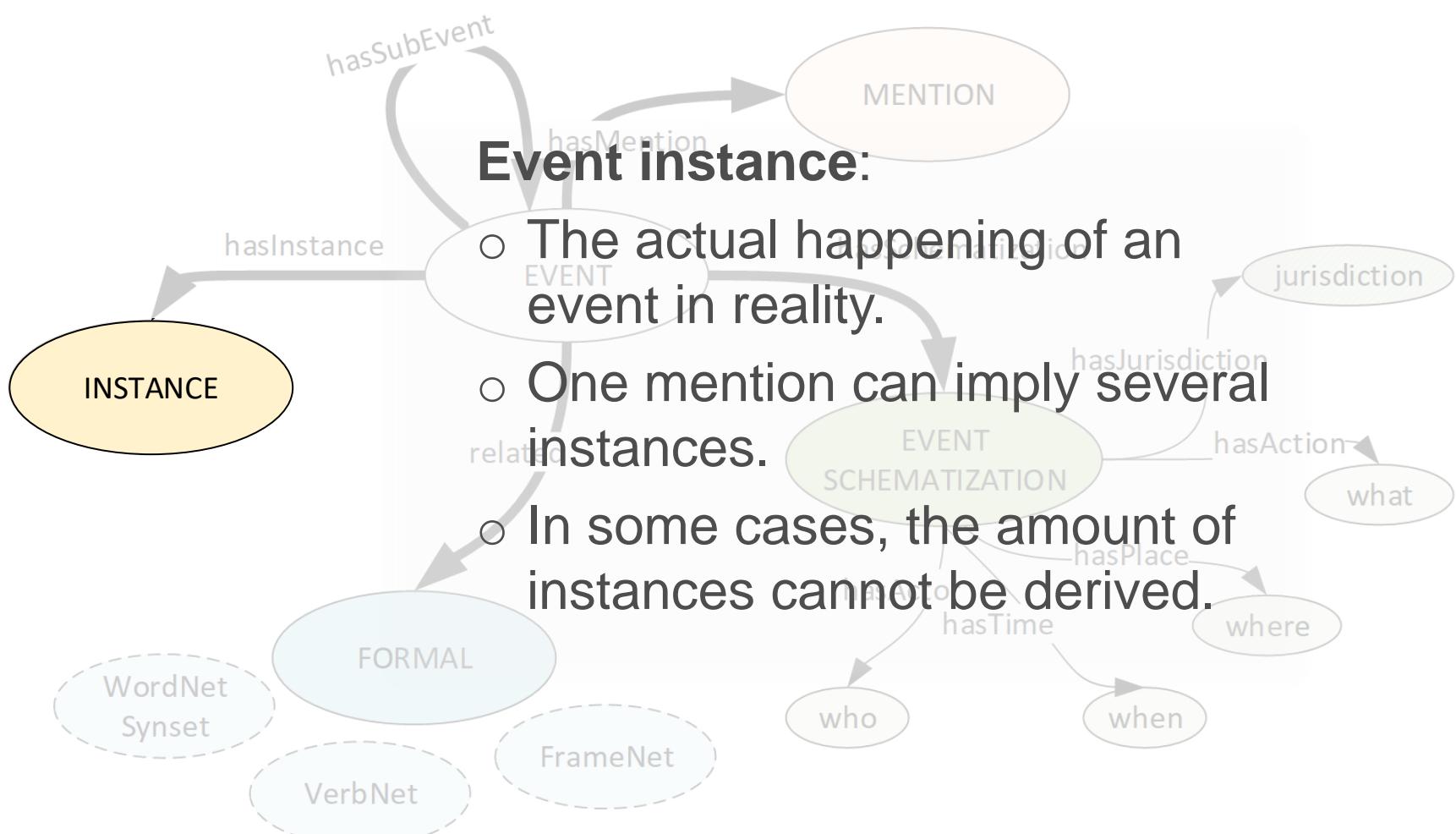
Events distinction among different concepts

Event schematization:

- The abstract representation of the information about an event, such as who, where, and so on.
- It is a midpoint between text, reality and abstraction.
- This representation can be useful to support Question Answering (QA) routines.



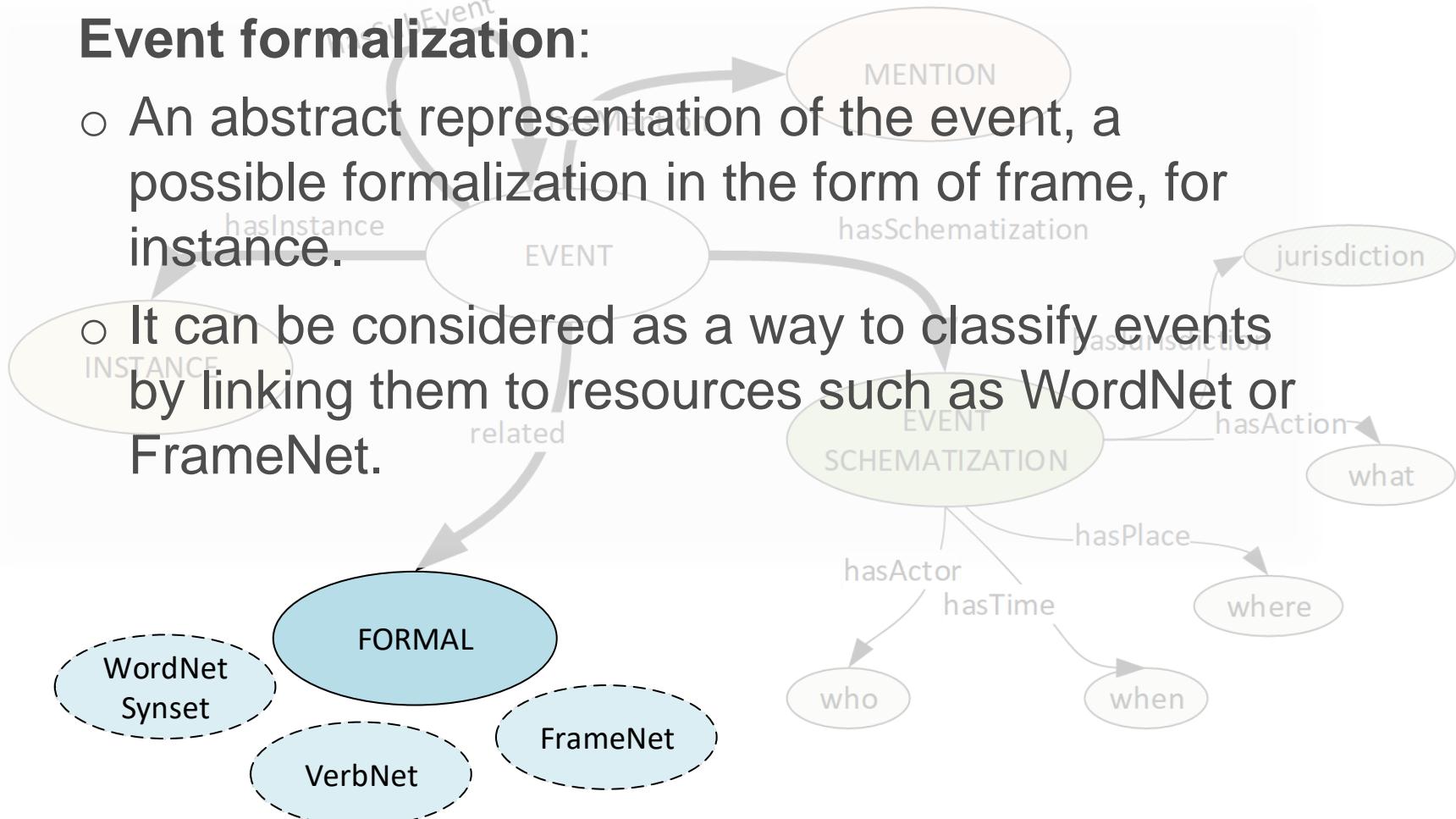
Events distinction among different concepts



Events distinction among different concepts

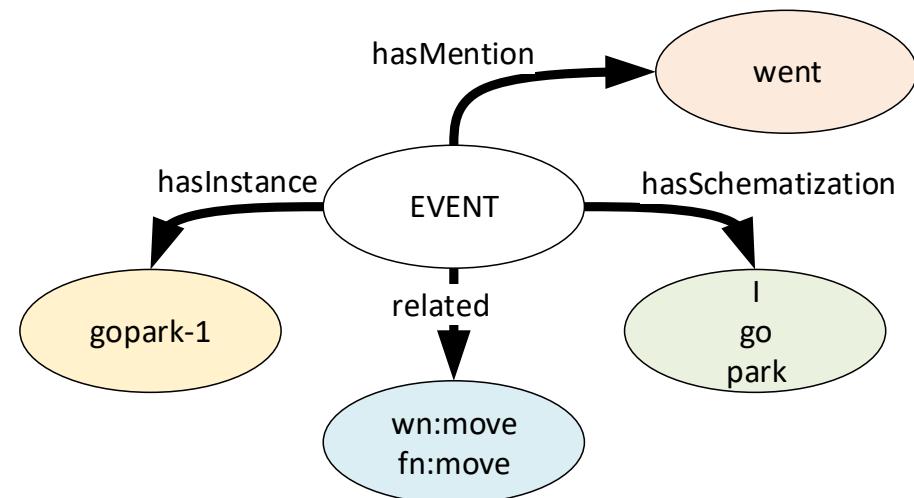
Event formalization:

- An abstract representation of the event, a possible formalization in the form of frame, for instance.
- It can be considered as a way to classify events by linking them to resources such as WordNet or FrameNet.

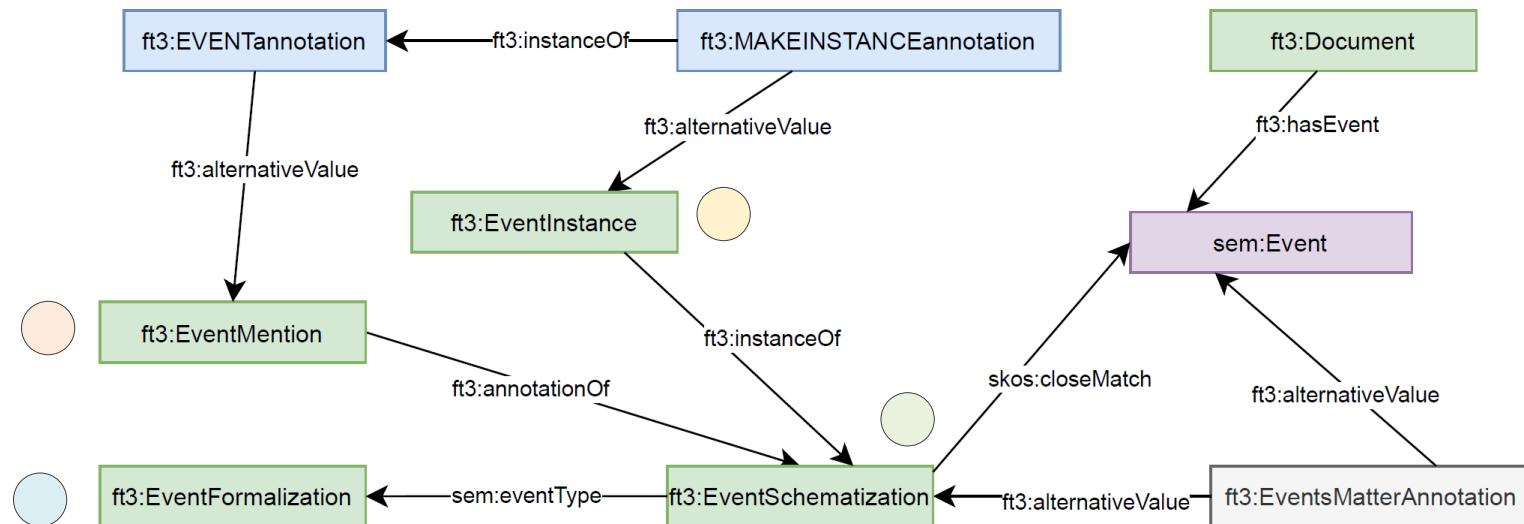


Example

I went to the park

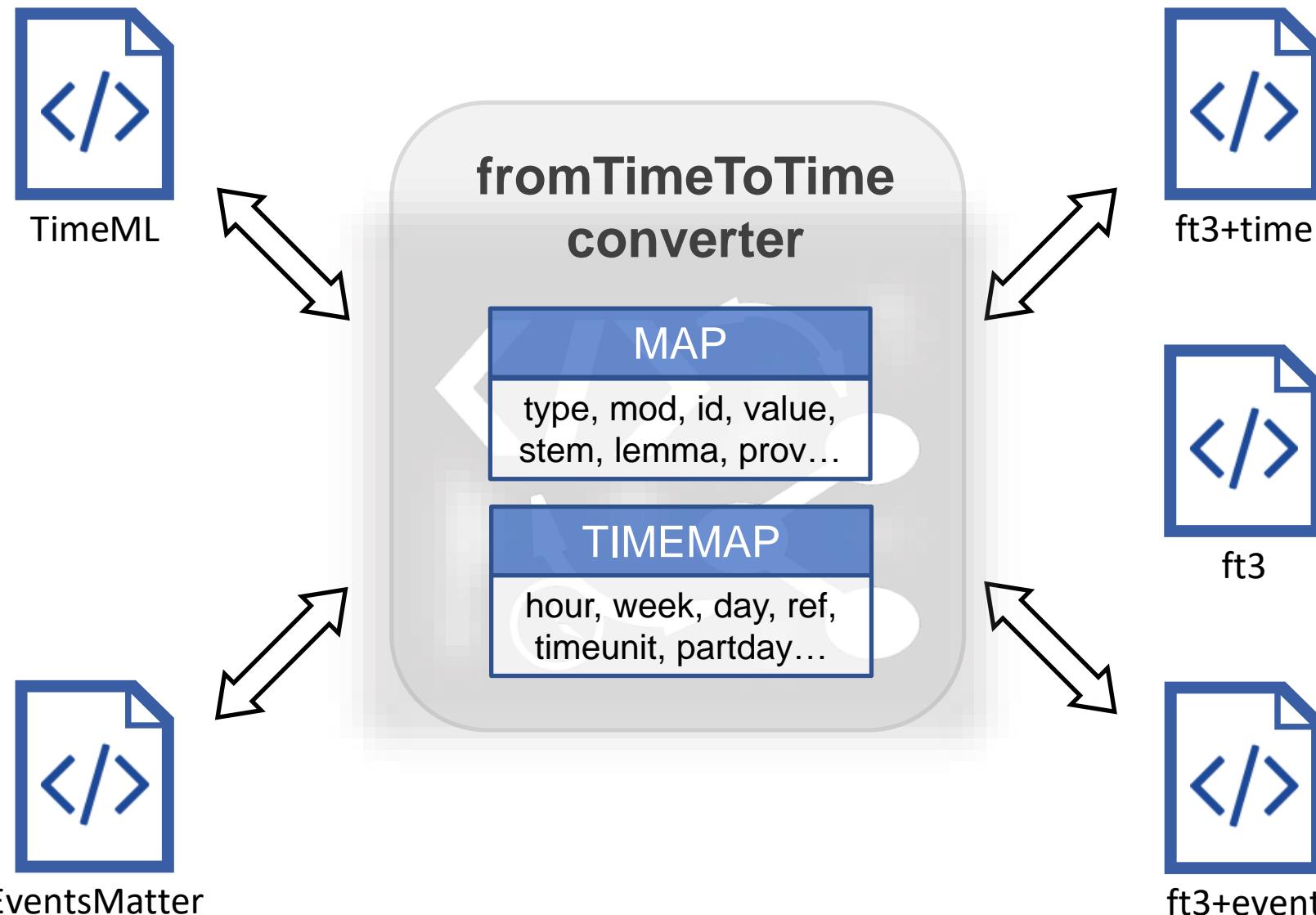


Events in ft3



Event-related resources

FT3 CONVERTER





ft3 Converter: On 6 October 1990 he married Ms N.R.

EventsMatter

On <Event_when tid="t4" type="DATE" value="1990-10-06">**6 October 1990**</Event_when> <Event_who argument="who" tid="t4">**he**</Event_who><Event_what argument="what" tid="t4" type="circumstance" prov="eventsmattertrain" lemma="marry">**married**</Event_what> **Ms N.R.**

TimeML

```
<?xml version="1.0" ?>
<TimeML xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="http://timeml.org/timeMLdocs/TimeML_1.2.1.xsd">
On <TIMEX3 tid="t4" type="DATE" value="1990-10-06">6 October 1990</TIMEX3> he
<EVENT eid="t4" class="circumstance">married</EVENT> Ms N.R.
</TimeML>
```

ft3

```
<https://fromtimetotime.linkeddata.es/doc/samples/doc002>
a nif:Context , ft3:Document ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "36"^^xsd:nonNegativeInteger ;
nif:title "X"^^xsd:String ;
nif:isString """On 6 October 1990 he married Ms N.R.""" ;
nif:AnnotationUnit [
...
]
```

```
<https://fromtimetotime.linkeddata.es/doc/samples/doc002/EventsMatter/Event_whenannotation_t4_5> [  
    a ft3:EventsMatterEvent_when ;  
    nif:beginIndex "3"^^xsd:nonNegativeInteger ;  
    nif:endIndex "17"^^xsd:nonNegativeInteger ;  
    ft3:hasID "t4"^^xsd:String ;  
    nif:isString """6 October 1990""";  
    ft3:hasTid "t4"^^xsd:String;  
    ft3:hasValue "1990-10-06"^^xsd:String;  
    ft3:hasType ft3:DATE ;  
];  
<https://fromtimetotime.linkeddata.es/doc/samples/doc002/EventsMatter/Event_whatannotation_t4_6> [  
    a ft3:EventsMatterEvent_what ;  
    nif:beginIndex "21"^^xsd:nonNegativeInteger ;  
    nif:endIndex "28"^^xsd:nonNegativeInteger ;  
    ft3:hasID "t4"^^xsd:String ;  
    nif:isString """married""";  
    ft3:hasType ft3:circumstance ;  
    ft3:hasProv "eventsmattertrain"^^xsd:String;  
    ft3:hasLemma "marry"^^xsd:String;  
];  
<https://fromtimetotime.linkeddata.es/doc/samples/doc002/EventsMatter/Event_whoannotation_t4_7> [  
    a ft3:EventsMatterEvent_who ;  
    nif:beginIndex "18"^^xsd:nonNegativeInteger ;  
    nif:endIndex "20"^^xsd:nonNegativeInteger ;  
    ft3:hasID "t4"^^xsd:String ;  
    nif:isString """he""";  
];
```

```
ft3:alternativeValue [  
<https://fromtimetotime.linkeddata.es/doc/samples/doc002/Time_t4> [  
    a sem:Time,  
    time:GeneralDateTimeDescription ;  
    time:year "1990"^^xsd:gYear ;  
    time:monthOfYear greg:October ;  
    time:month "--10"^^xsd:gMonth ;  
    time:day "---06"^^xsd:gDay ;];  
];
```

ft3+time

```
ft3:hasEvent [  
<https://fromtimetotime.linkeddata.es/doc/samples/doc002/EVENT_t4> [  
    a sem:Event ;  
    sem:EventType "marry" ;  
    ft3:hasType ft3:circumstance ;  
    ft3:hasID """t4""";  
    sem:hasTime [  
        <https://fromtimetotime.linkeddata.es/doc/samples/doc002/Time_t4> [  
            a sem:Time, time:GeneralDateTimeDescription ;  
            time:year "1990"^^xsd:gYear ;  
            time:monthOfYear greg:October ;  
            time:month "--10"^^xsd:gMonth ;  
            time:day "---06"^^xsd:gDay ;];  
        ] ;  
        sem:hasActor """he"""^^xsd:String ;  
    ]  
.].
```

ft3+events

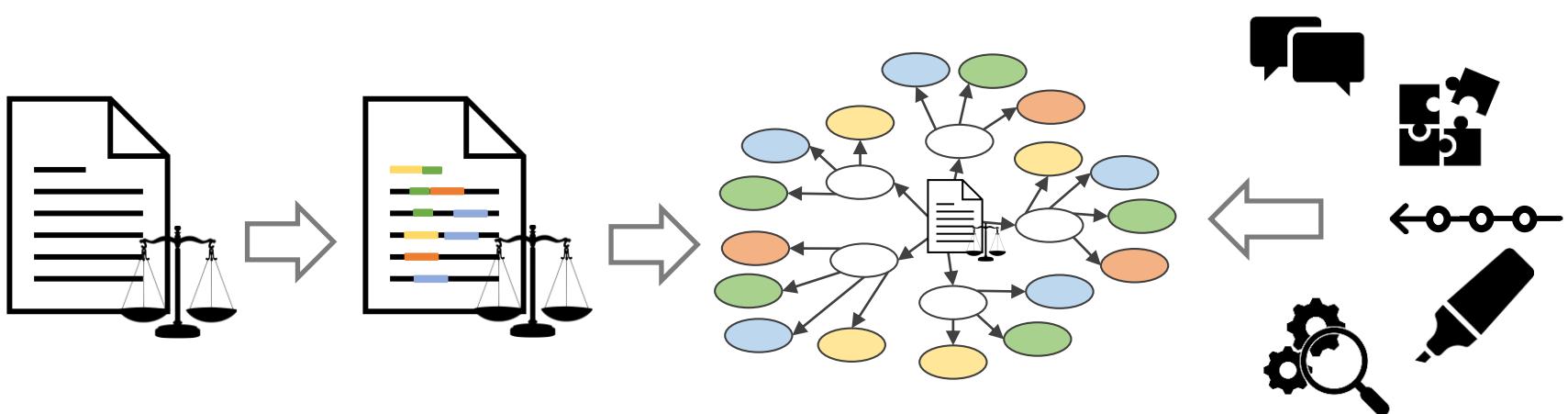
Event-related resources

EVENT-BASED KNOWLEDGE GRAPH

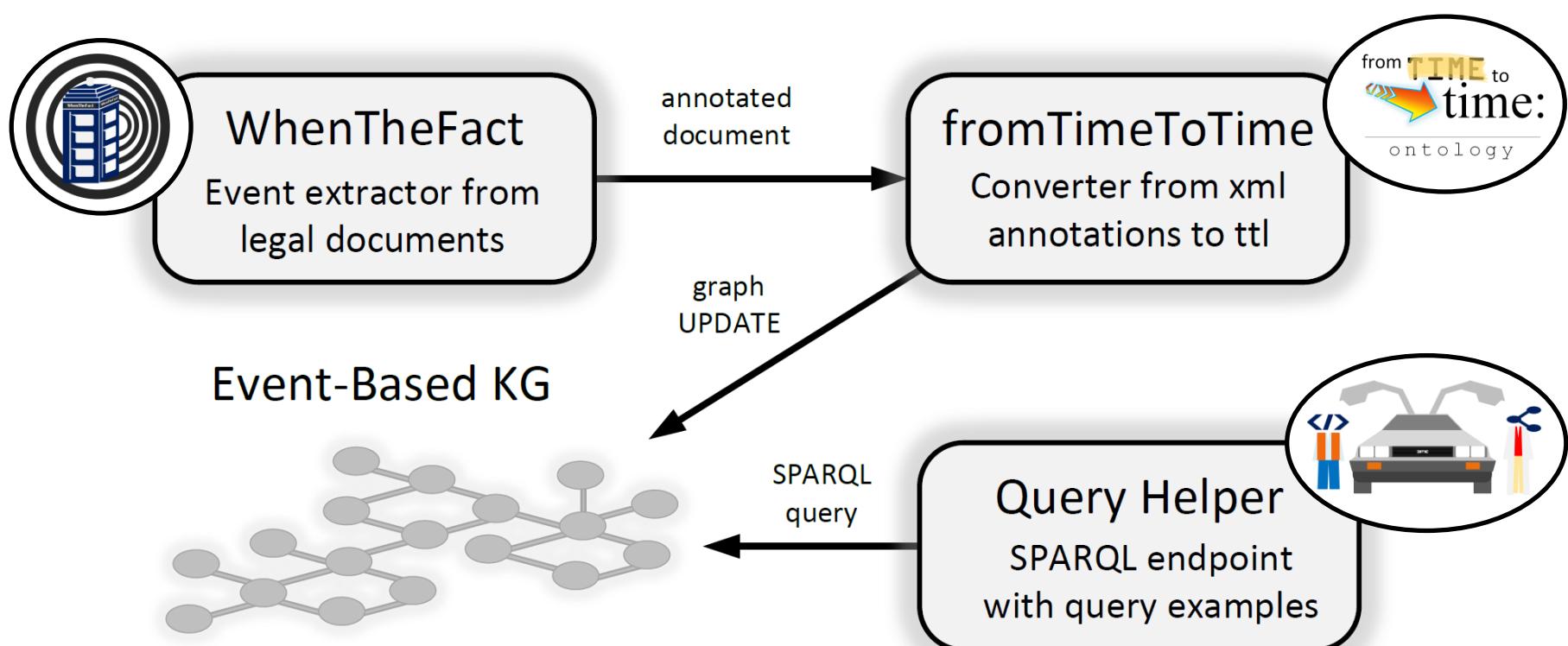
We propose Event-Based Knowledge Graphs:

“Knowledge Graphs where information is represented as a series of events.”

- We describe legal decisions using the events as the basis, being *blocks* that describe the legal judgment.
- A case is considered a narrative of events in different dimensions, namely procedural or relative to the case under judgment.
- Useful for various applications within the legal domain.



Pipelining all the tools in the thesis, we can go from a legal decision to a Legal Event-Based Knowledge Graph that can be queried for further application.



Event-related contributions

H2.a

EventsMatter, a legal corpus annotated with relevant events.

H1.b

H2.a

H3

ContractFrames for contract lifecycle events
Event Extractor **WhenTheFact** for legal texts

H2.b

H3

fromTimeToTime pipeline of resources:

- **ft3 Ontology** for temporal information and annotation representation.
- **ft3 Converter** for conversion among formats.
- **Legal EBKG+query helper** for further exploitation.

Outline

- Introduction
- State of the Art
- Research Objectives
- Temporal Expressions
 - Corpora
 - Temporal Tagging
- Events
 - Corpus
 - Event Extraction
 - Event-related Resources
- Conclusions and Future Work

C0. Analysis of the temporal information in the legal domain. An analysis of the particularities and challenges in the legal domain with regard to temporal information processing.

C1. Añotador. Design and implementation of a temporal tagger for Spanish and English.

C2. WhenTheFact. Design and implementation of an event extractor for European judgments.

C3. Corpora. Annotated with temporal information.

C3.1. TempCourt corpus. Corpus of judgments in English from different courts.

C3.2. HourGlass corpus. Corpus of short texts in Spanish of different provenance.

C3.3. EventsMatter corpus. Corpus of judgments in English annotated with events.

C4. ft3 Ontology. Representation of temporal information and data related to its annotation.

C5. Additional Tools. Standalone applications, integrated, or complementary to others.

C5.1. lawORdate. Web service that handles misleading legal citations in Spanish.

C5.2. ContractFrames. Software that extracts events about the lifecycle of a contract in English.

C5.3. Structure Extractor. Section detector from judgments, part of WhenTheFact.

C5.4. ft3 Converter. Online converter among different temporal annotation formats.

C5.5. Legal EBKG. Knowledge graph populated with events of legal decisions.

-  Spanish corpora for Sentiment Analysis: a survey. (2019) **M. Navas-Loro**, V. Rodríguez-Doncel. Language Resources and Evaluation, pp 1–38.
 -  TempCourt: evaluation of temporal taggers on a new corpus of court decisions. (2019) **M. Navas-Loro**, E. Filtz, V. Rodríguez-Doncel, A. Polleres, S. Kirrane. The Knowledge Engineering Review, Vol 34, E24.
 -  Annotador: a Temporal Tagger for Spanish. (2020) **M. Navas-Loro**, V. Rodríguez-Doncel. Journal of Intelligent & Fuzzy Systems 39 (2020), Vol 2, 1979–1991
 -  Lynx: A Knowledge-based AI Service Platform for Content Processing, Enrichment and Analysis for the Legal Domain. (2020) J. Moreno Schneider, G. Rehm, E. Montiel-Ponsoda, V. Rodríguez-Doncel, P. Martín-Chozas, **M. Navas-Loro**, M. Kaltenböck, A. Revenko, S. Karampatakis, C. Sageder, J. Gracia, F. Maganza, I. Kernerman, D. Lonke, A. Lagzdins, J. Bosque Gil, P. Verhoeven, E. Gomez Diaz, P. Boil Ballesteros. Special Issue of the Information Systems Journal.
 -  (TO BE SUBMITTED) Tools for building an event-based knowledge graph from legal decisions. (2021) **M. Navas-Loro**, V. Rodríguez-Doncel.
-

-  Spanish Corpus for Sentiment Analysis Towards Brands. **M. Navas-Loro**, V. Rodríguez-Doncel, I. Santana-Perez, A. Sánchez. In Speech and Computer: 19th International Conference, SPECOM 2017, Hatfield, UK, September 12-16, 2017, Proceedings. Cham: Springer, pp. 680–689.
-  ContractFrames: Bridging the gap between natural language and logics in contract law. **M. Navas-Loro**, K. Satoh, and V. Rodríguez-Doncel. JSAI International Symposium on AI. Springer, Cham, 2018.
-  Events Matter: Extraction of Events from Court Decisions. E. Filtz, **M. Navas-Loro**, C. Santos, A. Polleres, S. Kirrane. In Proceedings of the 33rd International Conference on Legal Knowledge and Information Systems (JURIX 2020), pp. 33-42.

Workshop Papers and other outcomes

-  Mining, Representation and Reasoning with Temporal Expressions in the Legal Domain (2017). **M. Navas-Loro**. In *Proceedings of the Doctoral Consortium, Challenge, Industry Track, Tutorials and Posters (RuleML+RR 2017)*.
-  OEG at TASS 2017: Spanish Sentiment Analysis of tweets at document level (2017). **M. Navas-Loro**, V. Rodríguez-Doncel. In *Proceedings of the Workshop TASS (SEPLN 2017)*. pp. 43–49.
-  MAS: A Corpus of Tweets for Marketing in Spanish (2018). **M. Navas-Loro**, V. Rodríguez-Doncel, I. Santana-Perez, A. Fernández-Izquierdo, A. Sánchez. In *The Semantic Web: ESWC 2018 Satellite Events*. pp. 363–375.
-  LawORDate: a Service for Distinguishing Legal References from Temporal Expressions (2017). **M. Navas-Loro**. In *Proceedings of the 1st Workshop TeReCom (JURIX 2011)*. pp. 25–31.
-  Events in the legal domain: first impressions (2018) **M. Navas-Loro**, C. Santos. In *Proceedings of the 2nd Workshop TeReCom (JURIX 2018)*.
-  Spanish Legislation as Linked Data (2018) V. Rodríguez-Doncel, **M. Navas-Loro**, E. Montiel-Ponsoda, P. Casanovas. In *Proceedings of the 2nd Workshop TeReCom (JURIX 2018)*.
-  TimeLex: a Suite of Tools for Processing Temporal Information in Legal Texts (2019) **M. Navas-Loro**, V. Rodríguez-Doncel. In *Proceedings of the 2nd Workshop Iberlegal (JURIX 2019)*.

- Software registered in the Registry of the Region of Madrid.
- Collaborations and projects:
 - Lynx
 - LPS-BIGGER
 - CENDOJ
- ActúaUPM competition finalist
- Usage of the software:
 - Anonymization
 - Date detection
 - Named Entity Recognition
 - Legal document annotation



(Jul-Oct 2017)

WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

Vienna, Austria

Prof. Sabrina Kirrane
and Prof. Axel Polleres
Tempcourt (KER)

(Jul-Oct 2018)



大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics

Tokyo, Japan

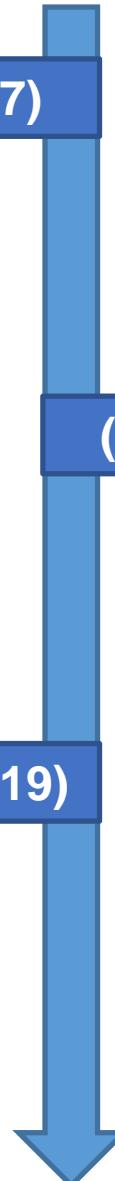
Prof. Ken Satoh
Contract Frames
(JURISIN)



(Sep-Dec 2019)

Bologna/Rome, Italy

Prof. Aldo Gangemi

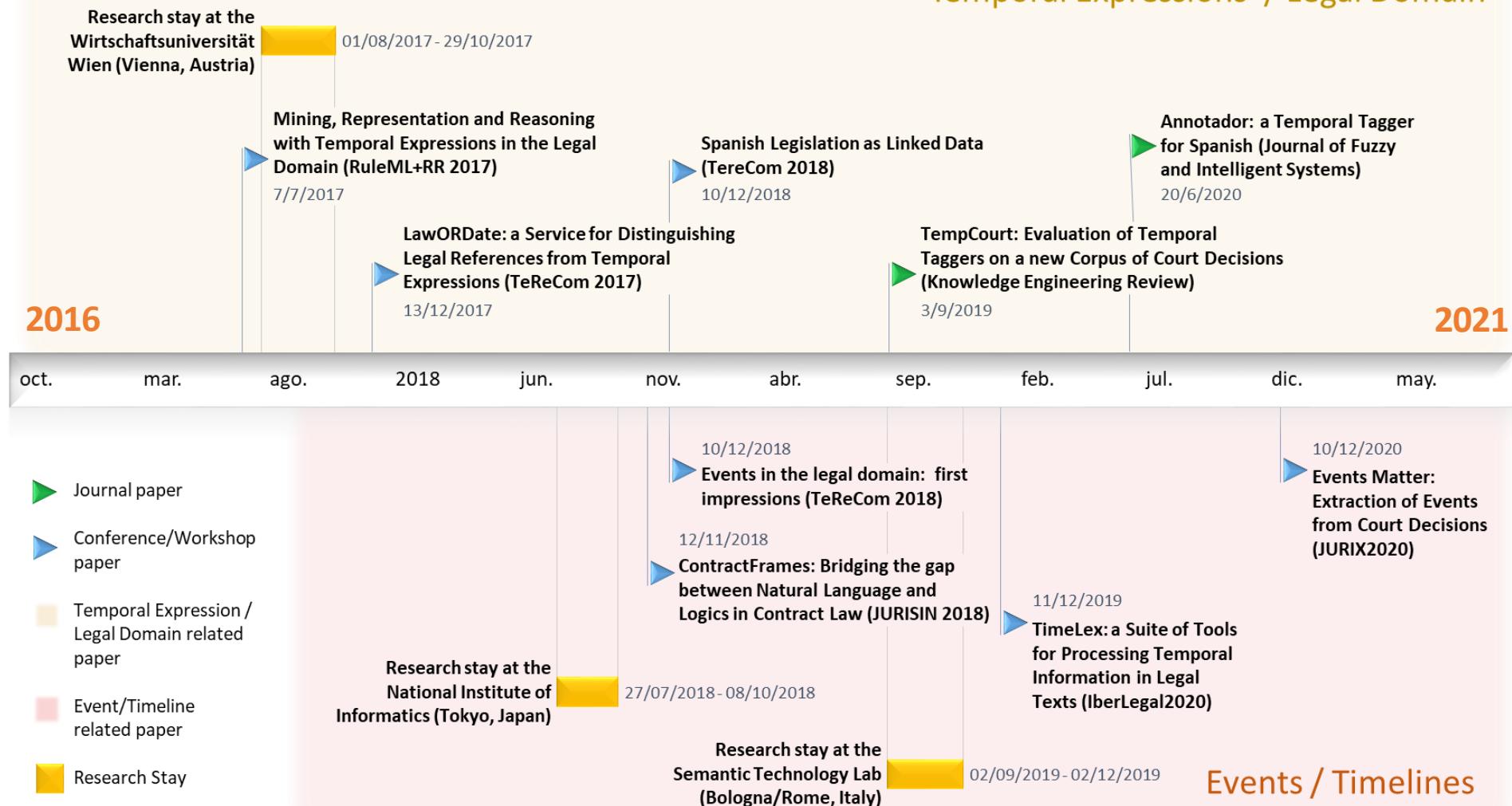


Additionally, visit to



Barcelona

Temporal Expressions / Legal Domain



- Extending the event extraction to **more languages**.
- Extending the corpora available: both **languages and domains**.
- Event **co-reference**: this would allow for instance to retrieve for instance the timeline of one actor's involvement in a case.
- Enriching the knowledge graph with **metadata**: helpful for co-reference.
- Processing **more types of documents**.
- **Deep Learning** for covering more events.
- Facilitate the **queries** to the EBKG for legal practitioners and further **exploit the EBKG**: QA, summarization...



Processing, Identification and Representation of Temporal Expressions and Events in Legal Documents

María Navas-Loro

Advisors: Víctor Rodríguez-Doncel
Asunción Gómez-Pérez

 mnavas@fi.upm.es

 <https://marinavas.oeg-upm.net>

 17/01/2022

 Online