

Annotation Scheme

Different parts of court decisions contain different information. As such, it may be desirable to treat different parts differently.

Following segments were annotated:

- Header: contains identification of participants, brief case description and ruling of court in the case
- History: contains procedural history including summary of previous court decisions
- Submission / Rejoinder: contains submission of petitioner, and rejoinder of opposing party
- Argumentation: contains court's own consideration and argumentation about the case
- Footer: contains legally required information on possibility of further appeal, or calculation of legal costs
- Footnotes: contains additional information or references
- Dissent: contains dissenting opinion of single judge or multiple judges.

Annotation Pipeline

At least one annotator have annotated every document in the dataset. 280 decisions were annotated by one trained annotator and then manually adjudicated by one trained editor/curator. 70 decisions were annotated by two trained annotators and then manually adjudicated by one trained editor/curator. Editor/curator was allowed to mark annotator's work as correct or incorrect, and to create editorial annotations to replace incorrect annotations.

Corpus contains only annotations marked as correct by editor/curator or created by editor/curator.

Designations of segments (EN to CZ)

Description above is in English, however name convention used in corpus is in Czech. To account for this, we provide comparison of English (used throughout this ReadMe) and Czech (used in corpus) designations.

Segments	
Header	01 Hlavička
History	02 Procesní historie
Submission/Rejoinder	03 Repliky
Argumentation	04 Argumentace
Footer	05 Patička
Footnotes	06 Poznámky pod čarou
Dissent	ÚS-odlišná stanoviska

Suggested Citation

HARAŠTA, Jakub, Jaromír ŠAVELKA, František KASL a Jakub MÍŠEK. Automatic Segmentation of Czech Court Decisions into Multi-Paragraph Parts. *Jusletter IT*, Weblaw AG, 2019, roč. 4, 23. Mai 2019, s. 1-10. ISSN 1664-848X.

Acknowledgment

Supported by the Czech Science Foundation under grant no. GA17-20645S.