

Human Language Understanding for Legal Informatics

Adam Wyner, University of Aberdeen, UK

CHIST-ERA Conference 2014

Istanbul, Turkey

Conceptual Problems

- Distinguish legal from non-legal language.
- Identify and formalise rules and norms.
- Legal ontology.
- Interpretation of law in contexts.
- Vagueness, ambiguity.
- Non-monotonic.



Technical Problems

- Volume.
- Interrelating documents and information.
- Linguistic complexity, e.g. length, terminology, complex clauses.
- Document complexity, e.g. list structures, references across documents, document dynamics.
- Formal representation of concepts.

Technologies to Address Problems

- Parsing.
- Machine-readable semantics.
- Semantic annotation.
- Extraction of and querying of named entities, relations, and properties.
- Ontologies.
- Linking.
- Semi v. fully automatic tools.
- Refinement from intermediate forms.
- Gold standard corpora development.
- Machine learning.

Corpora

- legislation, regulation, guidance, case law, policy consultations....

Scale

- Macro (Transnational, EU), Midi (National), Micro (local government).

Use Cases

- Compliance rulebooks - extracting rules for specific industries.
- Case search – finding relevant precedents.
- Decision tools – executing legislation.
- Jurisdictional influence networks – relating laws.
- Crowdsourcing law.
- Jury support.
- Support – managing legalities for emergencies, security, privacy, rights, policy making, business decision making, ...