Crowdsourcing Legal Text Annotation

Adam Wyner, University of Liverpool
Wim Peters, University of Sheffield

LawTech Camp
June 29, 2012
London
Needles and Haystacks

• Lots of open, unstructured legal data – cases, legislation, regulation.

• How to find the information you want?

• Want to make cases and legislation searchable, but by concepts, not just by strings.

  – IBM is the plaintiff, Seventh Circuit Court is the jurisdiction, program code is the trade secret, and the decision upholds a lower court decision in favour of IBM.

  – The holdings by judge X and with respect to cause of action A against defendant Z
Case Briefs – Focused Problem

To analyse a legal case, legal professionals annotate the case into its constituent parts. The analysis is summarised in a case brief. However, the current approach is very limited:

• Analysis is time-consuming and knowledge-intensive.
• Case briefs may miss relevant information.
• Case analyses and briefs are privately held.
• Case analyses are in paper form, so not searchable over the Internet.
• Current search tools are for text strings, not concepts.
Solution – Crowdsourcing Annotation

• Create a richly annotated, open source corpus for:
  – Machine learning.
  – Online search
  – Semantic web applications and information extraction

• Used by:
  – Law school students learning case analysis.
  – Legal professionals in identifying relevant cases.
  – Researchers in legal informatics.
Who Does This? Familiarity?

• The annotators would be law school students or legal professionals.
• We use the same annotations that are learned in law school and used in practice. No need to teach the annotations.
• Adjudicate conflicts.
• Make this a 'learning' tool.
Project Website

• [http://wyner.info/LanguageLogicLawSoftware/?p=1315](http://wyner.info/LanguageLogicLawSoftware/?p=1315)

• The website has further information and a video on how to use the tool.
In the following, we present the various annotations that are searchable in the corpus.
Argument for Party

Arguments for a particular party, using the most general notion:

• for Appellee, for Appellant, for Defendant, for Plaintiff.
Facts

Legal and procedural facts:

• Cause of Action – the specific legal theory upon which the plaintiff brings the suit.
• Defenses raised by Defendant – the defendant defenses against the cause of action.
• Legal Facts – the legally relevant facts of the case that are used in arguing the issues.
• Remedy requested by Plaintiff – what the plaintiff asks the court to grant.
Indexes

Various indicative information:

• Case Citation – the citation of the particular case being annotated.
• Court Address – the address of the court.
• Hearing Date – the date of the hearing.
• Judge Name – the names of the judge, annotated one at a time.
• Jurisdiction – the legal jurisdiction of the case.
Issues before the court:

• Procedural Issue – what the appellee claims that the lower court did wrong.
• Substantive Issue – the point of law that is in dispute (legal facts have their own annotation).
Legal Roles

The role of the parties in the case:

• Appellee, Appellee’s Lawyer, Appellant, Appellant’s Lawyer, Defendant, Defendant’s Lawyer, Plaintiff, Plaintiff’s Lawyer.

• General – buyer/seller, employer/employee, landlord/tenant, etc.
Other

Relevant information not covered by the other annotations:

- Appellee, Appellee’s Lawyer, Appellant, Appellant’s Lawyer, Defendant, Defendant’s Lawyer, Plaintiff, Plaintiff’s Lawyer.
- General – buyer/seller, employer/employee, landlord/tenant, etc.
Various parts of the legal decision:

- Concurring Opinion.
- Dicta – commentary about the judgement and holding, but not part of the rationale.
- Dissenting Opinion.
- Holding – the rule of law or legal principle that was applied in making the judgement.
Reasoning Outcomes II

Various parts of the legal decision:

• Judgement – the court’s final decision about the rights of the parties, the court’s response to a party’s request for relief, and bearing on prior decisions (e.g. affirmed, reversed, remanded, etc.).

• Rationale – the court’s analysis of the issues and the reasons for the holding.
728 F.2d 818

220 U.S.P.Q. 167

AMERICAN CAN COMPANY, Plaintiff-Appellee,
v.
Ishwar MANSUKHANI, et al., Defendants-Appellants.

No. 82-2004.

United States Court of Appeals,
Seventh Circuit.
Tool II

have discovered such formulations, this is not a patent action and "obviousness" is not a
benchmark. See Ferroline Corp. v. General Aniline & Film Corp., 207 F.2d 912 (7th Cir

The compositions of series 400 inks were not known to others in the industry. Steps
taken to guard their secrecy--notably confidentiality agreements employees were reque
t to sign with M & T and with Whittaker. There was no dispute as to the value of the for-
or with respect to the time and effort spent in developing them. Duplication was not se-
as to deprive the series 400 inks of trade secret status. A few sophisticated competi

Annotating Legal Texts, LawTechCamp
2012, London
Results of Annotation

- The annotators carry out their task and complete the project.
- Carry out inter-annotator agreement analysis.
- Curate the disagreements to create a Gold Standard corpus. Use this for machine learning.
- Search the annotations using an online tool, e.g. ANNIC.
Searching Annotations

{Legal Roles.Feature=="Plaintiff"}

Corpus: Entire datastore  Annotation set: All sets
Results:  Context size:

Search  Clear  Next page of 50 results

Page 1 (12 results)  Export

<table>
<thead>
<tr>
<th>Left context</th>
<th>Match</th>
<th>Right context</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>I. Jonas SPECNER, Trustee in Bankruptcy of ABC</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellant, v. REYNOLDS METALS</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellant, v. MARCAN PRODUCTS</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellant, v. PRESIDENT AND</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellee, v. MARCAN PRODUCTS</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellee, v. PRESIDENT AND</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellant, v. WILTSCHICK et al.</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellant, v. PRESIDENT AND</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellee, v. WILTSCHICK et al.</td>
<td></td>
</tr>
<tr>
<td>Metals Company 279 F.2d 337</td>
<td>Plaintiff Appellee, v. PRESIDENT AND</td>
<td></td>
</tr>
</tbody>
</table>

Annnotations

June 29, 2012
Searching Annotations

{Indexes.Feature=="Case Citation"}

Page 1 (21 results)

<table>
<thead>
<tr>
<th>Left context</th>
<th>Match</th>
<th>Right context</th>
<th>Features</th>
<th>Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jurist Home 510 F2d 894</td>
<td>Telex Corporation v...Machines Corporation</td>
<td>510 F2d 894 184 U.S.P.Q.</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Jurist Home 279 F2d 337</td>
<td>Speciner Abc v. Reynolds Metals Company</td>
<td>279 F2d 337 I. Jonas</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Jurist Search Open Jurist Home</td>
<td>510 F2d 894</td>
<td>Telex Corporation v. International Business</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>International Business Machines Corporation</td>
<td>510 F.2d 894</td>
<td>184 U.S.P.Q. 521, 1975-1 Trade</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>389 F2d 11 Emery Com...Products Corporation a R E</td>
<td>Open</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jurist Search Open Jurist Home</td>
<td>389 F2d 11 Emery Com...Products Corporation a R E</td>
<td>389 F.2d 2</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Jurist Search Open Jurist Home</td>
<td>389 F2d 11</td>
<td>Emery Company v. Marcan Products</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Products Corporation a R E</td>
<td>389 F.2d 11</td>
<td>158 U.S.P.Q. 529 The A.</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Jurist Search Open Jurist Home</td>
<td>279 F2d 337</td>
<td>Speciner Abc v. Reynolds Metals</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Abc v. Reynolds Metals Company</td>
<td>279 F.2d 337</td>
<td>I. Jonas SPECINER, Trustee in</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
<tr>
<td>Jurist Search Open Jurist Home</td>
<td>248 F2d 799</td>
<td>Irizarry Puente v. President and</td>
<td>{Indexes.Feature}</td>
<td></td>
</tr>
</tbody>
</table>
FAQ

• How easy is it to learn to use the tool?
• What if I don't agree with some of your annotations or features?
• Can I get the results of my annotations?
• How many cases must I do?
• How much time will it take?
• What if I have a problem with using the tool or if the tool is buggy?
• How thoroughly should I annotate the cases?
• Do we track good annotators and bad annotators?
Other Tasks

• Regulatory rules - the rules that define a regulation such as 'if...then' or 'must'.
• Legal case factors for case based reasoning - the mitigating or aggravating circumstances that contribute to the decision.
• Argumentation patterns.
To Participate

- Email me:

  adam@wyner.info

- I will send further information, a user name, and password to access the system.
Thanks for your attention!

• Contacts:
  – Adam Wyner, adam@wyner.info
  – Wim Peters, w.peters@dcs.shef.ac.uk

• Crowdsourcing Legal Cases Website:
  
  http://wyner.info/LanguageLogicLawSoftware/?p=1315