Digital humanities methods in comparative law

Quantitative analysis on a plain text corpus of books to trace the diffusion of legal concepts in public spheres
Introduction

Goal of this presentation

• Give an example of how digital humanities methods can be used in comparative law:
  – Tracing the diffusion of a legal tool: direct democracy by the ballot, from Switzerland to the United States, at the end of the 19th century.

• Think together in the potential of these methods for legal research:
  – comparative (space)
  – historical (time)
Introduction

Context

• Digital Humanities are the application of computing methods to humanities and social sciences disciplines.
  – Digital methods can process quantitative but also qualitative information

• Digital Humanities are a trans-disciplinary maturing field:
  – Digital Humanities researchers are tinkerers, opportunistic and willing to share and experiment with new methods
  – Methods are often new and their scientific potential and validity is still sometimes blurry

• Although legal studies handle a great number of text, they are not very present in Digital Humanities, yet.
Tracing legal concepts

Researching legal corpora

Looking for direct democracy
Legal research follows a stepping stones strategy

• Legal scholars reconstruct the logics of legal evolution:
  – from case to case
  – from statute to statute

• Focus on a selected path of cases.

• Is this path telling the whole story?
A quantitative approach is also possible

• Legal activity produces a lot of documents
  – These documents can be counted and organized.
• Legal studies rely heavily on text material.
  – Text material is easy to process programmatically in series.
• What can we learn from studying a large collection of legal texts, not only focusing on a selected path of cases, but on relative quantities of documents?
Concepts and practice go hand in hand

- Sociology of Law is a well established tradition of research
  - Max Weber: legal domination
  - Émile Durkheim: growth of civil law as society becomes more complex
- Law & Society journal
- The interaction of law and society is well researched, but relies heavily on case studies, few field researches and almost no account of extensive bodies of cases and legislation.
  - “highlights”, “illustrates”
- Quantitative research can help understand legal practices in society.
What do we need to track the diffusion of direct democracy?

- During the progressive era, at the beginning of the end of the 19th century, 20+ western US States implemented direct democracy by the ballot (i.e. unassembled)
- Two possible sources of inspiration:
  - New England Town Meetings
  - Switzerland modern constitution of 1848
- Scholars consistently cite Switzerland as the main source:
  - Based on similarities of procedure
  - But evidence of this influence is slim and anecdotal
Tracing legal concepts

Researching legal corpora

Looking for direct democracy
Indexing is useful to find stepping stones

- Legal scholars have developed clever and extensive index systems.
  - Stepan classification for comparative law (ISDC)
- Legal information retrieval relies on human resources and time.
- A clear index provides unambiguity and efficiency by reducing the amount of information that the researcher has to process.
- When you conduct an analysis on a selected number of cases, it’s a useful tool to guide you through these cases.
Indexing impairs quantitative analysis

- An index also impairs the researcher wanting to do quantitative analysis with a rigid formal system, in which an hypothesis must be formulated.
- Indexing imposes a strict taxonomy on a system that has no inherent one.
- Indexing posits (or equates to positing) a “world of legal ideas” existing outside the textual material of cases and statutes.
- Indexes are updated when a concept has largely pervaded through the system, making them impractical to trace its liminal or diffusion phases.
- The goal of the index might not concur with the goal of the researcher.
Plain-text opens new research strategies

- Work with the text of the document itself instead of working with metadata produced afterwards.
- Natural Language Processing algorithms are still way behind for a productive use in classification.
  - We cannot replace classification by a plain-text process.
- If you work under a hypothesis, plain-text becomes a viable option, because the space that is to explore is limited by the hypothesis, not by the system of classification.
  - We can productively use plain text to explore and test hypotheses.
Plain-text opens legal research to new corpora

• Any textual material is available to study the interaction between law and society.
  – What is the signature of a legal concept?
  – Can you trace the evolution of a legal concept outside of cases?

• Thomas Piketty used Balzac and Jane Austen to illustrate property and capital in the 19th century.

• By looking at series, we can go beyond illustrative use of this material and assess if an institution was influential in the literature of the time.

• If the corpus is very important, one might rely on approximations and control for errors.
Tracing legal concepts

Researching legal corpora

Looking for direct democracy
Get the number of items corresponding to a condition in a corpus: GCorpusAnalytics

- A Python module developed at EPFL to serialize queries on Google websites.
- Serialize on several dimensions:
  - expressions
  - years
  - country
  - patent office
  - court
- With several dimensions, the number of queries to run grows fast.
- Get the number of corresponding results.
- Use full logic of expressions.
Of all the confederations of history, Switzerland bears the closest resemblance in institutions to the United States, ... In these latter days Switzerland has become the standard bearer in all reforms which make for direct democracy and pure ...

The success of the referendum and popular initiative in Switzerland has won over the author to the conviction that, in the face ... The author hence passes to the advocacy of direct democracy for the United States, that is, of the popular right of ...
Direct democracy

```json
{
    "Request": {
        "Corpus": "books",
        "YearDebut": "1850",
        "YearEnd": "2010",
        "TimeInterval": "10",
        "Language": "en",
        "Outfile": "results.csv",
        "Expressions": [
            {
                "Expression": """"direct democracy"
            },
            {
                "Expression": """"direct democracy"
            },
            {
                "Expression": """"direct democracy"
            },
            {
                "Expression": """"direct democracy"
            },
            {
                "Expression": """"direct democracy"
            }
        ]
    }
}
```
## Concepts > Corpora > Direct democracy

<table>
<thead>
<tr>
<th>Timespan</th>
<th>&quot;direct democracy&quot;</th>
<th>&quot;direct democracy&quot; &quot;united states&quot;</th>
<th>&quot;direct democracy&quot; &quot;new england&quot;</th>
<th>&quot;direct democracy&quot; &quot;united states&quot; &quot;new england&quot; switzerland</th>
<th>&quot;direct democracy&quot; &quot;united states&quot; &quot;new england&quot; switzerland</th>
<th>&quot;direct democracy&quot; &quot;united states&quot; &quot;new england&quot; switzerland</th>
<th>&quot;direct democracy&quot; &quot;united states&quot; &quot;new england&quot; switzerland</th>
</tr>
</thead>
<tbody>
<tr>
<td>1850-1859</td>
<td></td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1860-1869</td>
<td></td>
<td>890</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1870-1879</td>
<td></td>
<td>891</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>1880-1889</td>
<td></td>
<td>960</td>
<td>1100</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>678</td>
</tr>
<tr>
<td>1890-1899</td>
<td></td>
<td>1440</td>
<td>1220</td>
<td>7</td>
<td>4</td>
<td>842</td>
<td>950</td>
</tr>
<tr>
<td>1900-1909</td>
<td></td>
<td>1540</td>
<td>1210</td>
<td>2</td>
<td>0</td>
<td>810</td>
<td>798</td>
</tr>
<tr>
<td>1910-1919</td>
<td></td>
<td>3770</td>
<td>2090</td>
<td>1040</td>
<td>5</td>
<td>943</td>
<td>1120</td>
</tr>
<tr>
<td>1920-1929</td>
<td></td>
<td>3290</td>
<td>1720</td>
<td>998</td>
<td>2</td>
<td>830</td>
<td>930</td>
</tr>
<tr>
<td>1930-1939</td>
<td></td>
<td>2710</td>
<td>1550</td>
<td>993</td>
<td>6</td>
<td>832</td>
<td>964</td>
</tr>
<tr>
<td>1940-1949</td>
<td></td>
<td>3720</td>
<td>1850</td>
<td>1030</td>
<td>413</td>
<td>855</td>
<td>995</td>
</tr>
<tr>
<td>1950-1959</td>
<td></td>
<td>6260</td>
<td>2170</td>
<td>1050</td>
<td>414</td>
<td>889</td>
<td>1220</td>
</tr>
<tr>
<td>1960-1969</td>
<td></td>
<td>17700</td>
<td>3750</td>
<td>1240</td>
<td>420</td>
<td>974</td>
<td>1720</td>
</tr>
<tr>
<td>1970-1979</td>
<td></td>
<td>23800</td>
<td>4350</td>
<td>1200</td>
<td>4</td>
<td>1050</td>
<td>2380</td>
</tr>
<tr>
<td>1980-1989</td>
<td></td>
<td>30900</td>
<td>5790</td>
<td>1260</td>
<td>424</td>
<td>1220</td>
<td>3030</td>
</tr>
<tr>
<td>1990-1999</td>
<td></td>
<td>54500</td>
<td>9860</td>
<td>1490</td>
<td>458</td>
<td>1740</td>
<td>6210</td>
</tr>
<tr>
<td>2000-2009</td>
<td></td>
<td>83200</td>
<td>14600</td>
<td>1550</td>
<td>451</td>
<td>2560</td>
<td>9830</td>
</tr>
</tbody>
</table>
Switzerland was the main source of inspiration for direct democracy in the US

- Books on direct democracy mention Switzerland two decades earlier than New England.
- By 1910, almost all Western States have implemented direct democracy by the ballot.
- This validates the hypothesis that Switzerland was the main reference.
Cree was more influential than Sullivan in this diffusion

- Thomas Gliozzo mentions two main influential books introducing the Swiss reference in the American debate:

- Cree is much more referenced in the Google Books corpus than Sullivan.
Tracing legal concepts

Researching legal corpora

Looking for direct democracy
Let’s build something together

• GCorpusAnalytics is available on GitHub [https://github.com/Erispoe/GCorpusAnalytics](https://github.com/Erispoe/GCorpusAnalytics)

• You need to know a bit of Python (but you will be able to use it if it’s explained to you)

• For the moment, it only works with Google Books but expansions are planned for Google Patents and Google Scholar (including legal cases)

• Contact me if you are interested in using it for your work: thomas.favre-bulle@epfl.ch