Legal Data Science

Bart Custers PhD MSc LLM
Associate professor/head of research
eLaw – Center for Law and Digital Technologies
Leiden University

INFORM WORKSHOP
Leiden University
4th December 2018
13:00-13:30
Who thinks I can read your mind?
I cannot read your thoughts, but...

weird rash

weird rash on big toe
weird rash on arm
weird rash comes and goes
weird rash
weird rash on neck
weird rash on chest
weird rash on face
weird rash on back
weird rash on hands
weird rash on my arm
Reading your mind...

What are you thinking when you search for...

- Barbados
- Prozac
- Weird rash
- Lolitas
- Semiautomatic rifles

...on Google?
Why is data collected and analysed?

Making predictions...

- **Examples**
  - Weather forecast
  - Predicting stock markets
  - Law enforcement
    - Predicting terrorist attacks
    - Predictive policing: surveillance
  - Tax authorities
    - Finding fraud
  - Corporations
    - New customers/target groups
    - Marketing new products
    - Credit scores (insurance, mortgages, loans, etc.)
Why is data collected and analysed?

Making predictions

- Examples
  - Weather forecast
  - Predicting stock markets
  - Law enforcement
    - Predicting terrorist attacks
    - Predictive policing: surveillance
  - Tax authorities
    - Finding fraud
  - Corporations
    - New customers/target groups
    - Marketing new products
    - Credit scores (insurance, mortgages, loans, etc.)
Applying data science to law

Topic of this presentation:
- Data science as a new research method...
  ... in the legal domain

How can legal research benefit from data sciences?

This presentation is NOT about regulating data sciences or big data.
Disrupting law school...

- In the USA, it is less desirable to be a lawyer today.
- Law school applications are down 40% since 2005.
- Employment rates for law graduates are shrinking

Advantages in technology and business model innovations are changing the traditional legal model:

- Legal solutions are more standardized and systematized
- Lawyers become more productive due to technological innovations
- Technological innovations offer solutions for non-lawyers
New perspectives

- Legal documents (legislation, case law, policy documents, research papers, etc.) are data:
  - increasingly available in electronic formats
  - new data analyses technologies (viz. big data analytics), such as data mining and text mining
  - increased accessibility of files that were previously only available on paper

- New research methods in the legal domain that are more data-driven.
  - Focus on statistical relations rather than causal relations
  - Less focus on hypotheses and theory
  - Increased efficiency and effectiveness
  - Novel, unexpected results
Data science in legal research

Data science may contribute to answer different types of legal research questions

<table>
<thead>
<tr>
<th></th>
<th>What is existing positive law</th>
<th>How to interpret existing positive law</th>
<th>Sociology of law</th>
<th>Developing legal theory</th>
<th>Developing improved legislation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantitative legal predictions</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Accelerate research</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Improved regulation</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
**Quantitative legal predictions**

- **Ruger et al. (2004):** tournament to predict US supreme court rulings
  - Individual justices:
    - Human experts: 68%
    - Model: 67%
  - Case outcome
    - Human experts: 59%
    - Model: 75%

- **Katz et al. (2014)**
  - Model independent of
    - Justice transition
    - Ideological shift of justices
    - Macro political/economic conditions

Data for the Katz-model:
- Over 60 years
- 7,700 cases
- 68,000 justice votes
Accelerate legal research

- Ravellaw
  https://www.ravellaw.com/search/m?query=privacy

- ROSS
  https://www.youtube.com/watch?v=sbwEO8-Voxo (1’05’’)

Advantages
- Decreasing the likelihood of missing important information
- Increasing accuracy and reliability of research
- Potential of novel, unexpected results
Improved regulation – public acceptance

Not only legal documents are relevant data: example of social media

COOSTO: http://open.coosto.com/nl/

Sentiment analyses:
• public support?
• timing?
**Nudging:**

Use social media data for improved enforcement

- Increase knowledge about how people will behaviorally respond to proposed legislation
- This may help to create legislation that can be enforced more easily...
- ... by nudging people into the expected behavior
Data science does not replace or invalidate other research methods

- Data science may be complementary to other methods:
  - It may disclose **novel, unexpected results**
  - It may **improve efficiency/effectiveness** of research

- Data science, however, provides limited opportunities for:
  - Investigating causality
  - Deeper understanding
  - Theory development
Questions?

Thank you for your attention!

Or contact me later: b.h.m.custers@law.leidenuniv.nl