
Artificial intelligence and judicial systems

CEPEJ, Council of Europe, Strasbourg, 27 June 2018

A.I. and Criminal Justice

Assoc. Prof. Dr. Aleš Završnik

Collegium
Helveticum

ETH zürich



Universität
Zürich

Z

hdk



Content

1. New knowledge in criminal justice
2. New reactions to crime
3. Risks in automated decision-making tools

Content

1. New knowledge in criminal justice
2. New reactions to crime
3. Risks in automated decision-making tools

Criminology & knowledge (Garland 1985):

- statistics as a cornerstone
- scientific understanding of crime v. *ancien régime* superstition

Competing scientific knowledge:

- *psychoanalysis*
- *neo-positivism* (genetics, neuroscience)
- *mathematics* (Amoore 2014): new language for security purposes

New semantics, new language

“meaning extraction”

“sentiment analysis”

“opinion mining”

“computational treatment of subjectivity in text”

New semantics, new language

“sleeping terrorist”

“person of interest”

“escalating behavior”

“signature strike”

New semantics, new language

“sleeping terrorist” (v. suspect)

“person of interest” (v. suspect, accused, convicted person)

“escalating behavior” (v. criminal behavior)

“signature strike” (v. identified person X.Z.)

Consequences

- Crime-relevant knowledge = ***database***
- Proper reasoning = ***algorithms***
- How to tackle crime = ***predictive policing***
- How to adjudicate cases = ***automated justice***

Content

1. New knowledge in criminal justice
2. New reactions to crime
3. Risks in automated decision-making tools

Legal automation

- To predict the outcomes of legal problems
(Ashley and Brüninghaus, 2006)
- To decide on co-financing litigations (e.g. Legalist.us)
- To help human lawyers conduct research more quickly
(e.g. rossintelligence.com)
- To predict the judicial decisions

Automated policing

- Preventive screening of Muslims (Germany)
- Police Twitter analysis (Slovenia)
- Identification of members of violent criminal groups (UK)
- Crime predictive software
 - .PredPol (Santa Cruz, Cal.), BlueCrush
 - .Precobs (Zürich, Munich)
 - . Maprevelation (France)

Example: Automated Inference on Criminality using Face Images

- based on still face images of 1,856 real persons (half convicted)
- Computer vision, pattern recognition with supervised machine learning

(Xiaolin & Xi, 2016)

Example: Automated Inference on Criminality using Face Images



(a)



(b)

Figure 9. (a) The four subtypes of criminal faces; (b) The three subtypes of non-criminal faces.

(Xiaolin & Xi, 2016)

ALGORITHMIC PHRENOLOGY:

Legitimizing, normalizing existing implicit biases

ALGORITHMIC PHRENOLOGY:

Legitimizing, normalizing existing implicit biases

“‘Laundering’ of human prejudice through computer algorithms” (Arcas et al. 2017)

ALGORITHMIC PHRENOLOGY:

Legitimizing, normalizing existing implicit biases

“‘Laundering’ of human prejudice through computer algorithms” (Arcas et al. 2017)

Criminal justice

- pre-trial detention / remand procedure
- evidence gathering / evaluation
- sentencing procedure
- imprisonment (treatment, “AI companions”)
- parole

Content

1. New knowledge in criminal justice
2. New reactions to crime
3. Risks in automated decision-making tools

1. Domain transfer

a) data in criminal justice

b) criminal justice domain embeds values, political decisions

1. Domain transfer

Balancing acts

- e.g. fairness over efficiency
- e.g.: tracking domestic violence: purposes?
- e.g.: “criminal faces”: ignoring the real functioning of a legal system, the nature of law as discourse, the nature of language, confusing normative and empirical levels

2. Building algorithms & datasets

- humans craft them
- may encode errors
- data needs to be collected, cleaned, data sets gathered
- several prior decisions have to be made

3. Functioning of algorithms

Discrimination may be an artefact of the data mining process itself, rather than a result of programmers assigning certain factors inappropriate weight.

Biases may be directly excluded, but can be calculated indirectly (cf. Barocas & Selbst, 2015)

-
- race a proxy for crime (Harcourt 2015)
 - geography a proxy for race (O'Neil 2016)
 - etc.

4. Human v. Machine Biases

- Is de-biasing possible?

De-biasing?

The nature of language

REPORTS | PSYCHOLOGY

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan^{1,*}, Joanna J. Bryson^{1,2,*}, Arvind Narayanan^{1,*}

+ See all authors and affiliations

Science 14 Apr 2017:
Vol. 356, Issue 6334, pp. 183-186
DOI: 10.1126/science.aal4230

4. Human v. Machine Biases

- Is de-biasing desired?

5. Inequality

a) overpolicing

b) underpolicing

- geographic areas

- types of crime

6. Judicial evolution

- Ad infinitum problem: data hunger...

6. Judicial evolution

- Ad infinitum problem: data hunger
- “The eternal past”, vicious circle effect

6. Judicial evolution

- Ad infinitum problem: data hunger...
- “The eternal past” or vicious circle effect
- Self-fulfilling prophecies

7. Place of subjectivity

Erasure of subjectivity and case-specific narrative:

- From narrative to database (Franko Aas, 2004)
- From database to automated justice (Završnik, 2018)

ZAVRŠ NIK, Aleš (ur.). Big data, crime and social control. Abingdon; New York: Routledge.

8. Human rights

.

- constitutional rights
- criminal procedure guarantees
- personal data protection rights

8. Human rights - constitutional rights

- the right to a lawful judge
- the right to natural judge
- the right to fair trial

- principle of non-discrimination

8. Human rights – criminal procedure

- the presumption of innocence
- fair and public hearing
- the right to cross-examine witnesses (Loomis case, US)
- delivering reasons for decision
- standards of proof (“reasonable doubt”...)
- the principle of legality (v. opportunity principle)

8. Human rights – personal data protection

Automated-decision making

e.g. GDPR “suitable measures” (Art. 22):

- . the right to obtain human intervention
- . to express his/ her point of view
- . to contest the decision

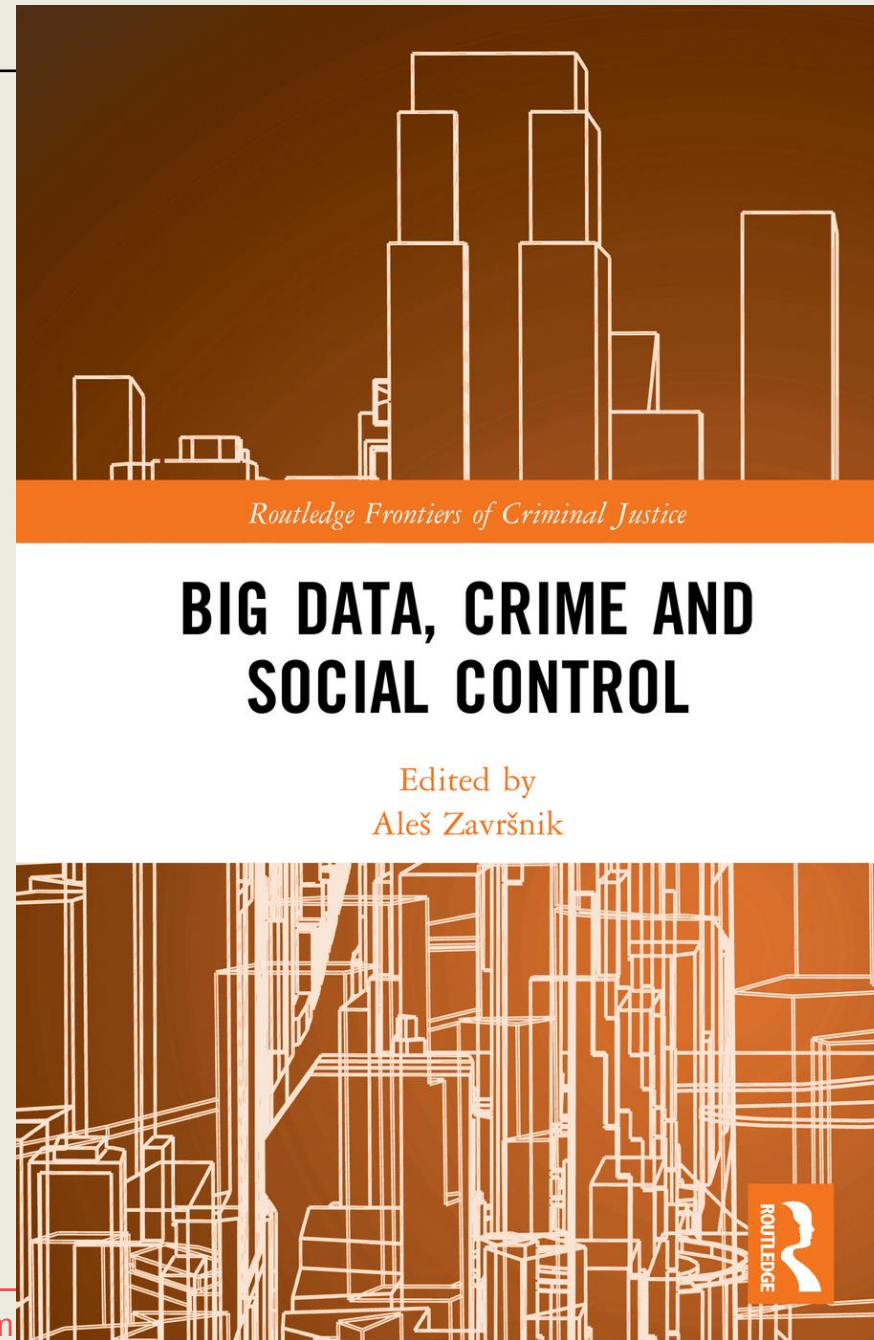
Concluding...

1. Changes in knowledge production

- what is crime and how to (re)act
- human agency

2. Human rights-compliant approach

constitutional, criminal law and personal data protection law



Thank you for you attention.

zavrsnik@collegium.ethz.ch

ales.zavrsnik@pf.uni-lj.si

