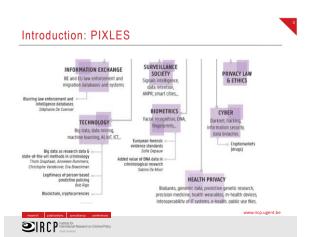


PIXLES is an interdisciplinary knowledge platform and research consortium with a focus on current societal dilemma's and developments in the context of privacy, information exchange, law enforcement & surveillance. The core value of PIXLES is to safeguard privacy while enhancing security.



Privacy, Information Exchange, Law Enforcement & Surveillance





Overview session

- 1. New and emerging data sources for spatio-temporal analyses: Opportunities and threats for big data policing Thom Snaphaan (speaker); Wim Hardyns
- 2. Social capital variables at the neighbourhood level as predictors in a predictive policing model Anneleen Rummens (speaker); Wim Hardyns
- 3. Automated Suspicion: Offender-Based Predictive Policing in the Age of Big Data Bob Rigo (speaker)

New and emerging data sources in criminology: Opportunities and threats for big data policing Thom Snaphaan, MSc. & prof. dr. Wim Hardyns

CPS International Conference 'Street Policing in a Smart Society' Academic panel session: Predictive policing in the big data era 17 September 2019

GREAT GREAT CANADA CRIMINOLOGY OF RECEPCTANT Research on Criminal Policy PIXLES



ww.ircp.ugent.be

n ugent he

Content

- I. Introduction
- II. Big data
- III. Big data policing
- IV. Scoping review: big data in environmental criminology
- V. Conclusion and discussion



I. Introduction

The best predictor of future crime is ... ?

. .

rirch ugent be

PRIOR CRIME

SIRCP Institute for International Research on Driminal P Otor University

I. Introduction

Big data policing is more than predicting the future (# predictive policing)

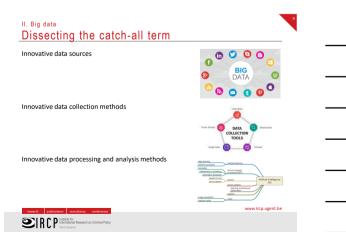
Is big data policing another type of policing? Next to, among others, intelligence-led policing (ILP), problem-oriented policing (POP), community oriented policing (COP), hot spot policing → More like a paradigm shift (which is not limited to the police)

Although big data predominantly have been used for crime prediction purposes, the use of big data is much broader than that

ightarrow Big data as one type of research data

 \rightarrow Crime prevention, policy





II. Big data Definition and characteristics

Myriad of definitions of big data

→ Regularly described based on their characteristics

3V's:

- Volume: enormous quantities of data
- Velocity: created and processed in or near real-time
- Variety: structured, semi-structured or unstructured

Kitchin & McArdle (2016):

"... the key boundary characteristics of Big Data, which together differentiate it from 'small data', are **velocity** (both frequency of generation, and frequency of handling, recording, and publishing) and **exhaustivity**" (p. 8) (captures an entire system, rather than being samples).

viron ugent be

SIRCP Institute for International Research on Driminal Over University

II. Big data Two types of data

Two distinct types of data:

- Data of active sollicitation: made by researchers for research purposes
 - ightarrow 'made' data

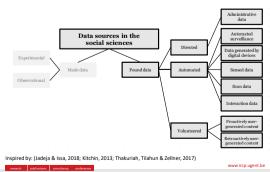
 Data of passive sollicitation: collected for purposes other than research, and researchers do not have any influence on the data collection and data processing

ightarrow 'found' data

Source: (Chen, Ma, Susilo, Liu & Wang, 2016; Connelly, Playford, Gayle & Dibben, 2016)

SIRCP Institute for International Research on Official I

II. Big data Taxonomy of data sources in the social sciences

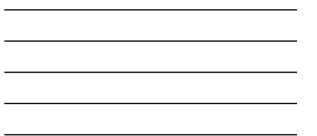


SIRCP Institute for International Research on Criminal Po Over University

II. Big data Taxonomy of data sources in the social sciences

Directed data:





II. Big data

Taxonomy of data sources in the social sciences

Volunteered data (proactively and retroactively user-generated content):







III. Big data policing Black, blue, bright and no data Ferguson (2017): The rise of big data policing

- Black data: distortions of race, transparency and law
- Blue data: (internal) policing data
- Bright data: big data identifies the risk but not necessarily the remedy, and just because big data policing identifies the risk, does not mean every crime problem needs a policing solution



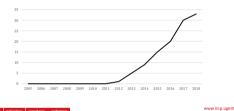
No data: data holes

research	publications	consultancy	conferences
D IR		de for ational Research University	on Criminal Policy

III. Big data policing A quick look at the relative importance

Web of Science: today's premier academic database

Search: ("Big data" AND Policing) in Topic for the years 2005-2018



SIRCP Institute for International Research on Driminal Po Over University

III. Big data policing A quick look at the relative importance

Search: ("Big data" AND Policing) in Topic for the years 2005-2018

18 Law	15 Computer science theory methods	12 Engineering electro Electronic	8 COMPUTER SCIEN ARTIFICIAL INTELLIGENCE	7 KE TELECOMMUNICAT
16 CRIMINOLOGY PENOLOGY	13 COMPUTER SCIENCE INFORMATION SYSTEMS			
		4 Computer science Hardware Architecture	4 Computer science Interdisciplinary Applications	4 Environmental Sciences
research publications consultancy co	nferences			www.ircp.ugent.be
SIRCP Institute for International Research on D	ininal Policy			

III. Big data policing Today's applications of new technologies Table 1 Prevalence of the technologies

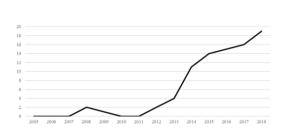
Survey (n=46) regarding the use of new technologies by police forces and law enforcement agencies in 11 different countries (EU + Australia)

Obstacle	Туре	Percentage experienced
Insufficient financing	Organizational	80%
Insufficient availability of technology	Technological	56%
Legal basis not available	Legal	44%
Insufficient overview of available technologies in the market	Technological	36%
Insufficient insight and overview regarding technology in policing	Organizational	29%

International and a set of the and databases Databases/soughing of the and da

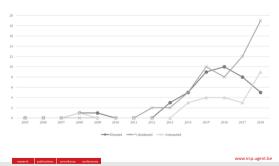
v.ircp.ugent.be

IV. Scoping review: big data in environmental criminology Evolution of the number of included studies



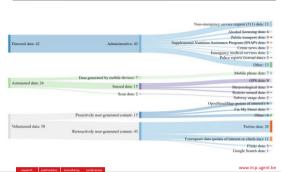
SIRCP Institute for International Research on Driminal Pol Gene Directly

IV. Scoping review: big data in environmental criminology Evolution of the number of data sources used



STRCP Institute for Institute for International Research on Criminal Pole Operativesty

IV. Scoping review: big data in environmental criminology Data sources used



SIRCP

IV. Scoping review: big data in environmental criminology Strengths and weaknesses

Strengths

- Real-time temporal and fine grained spatial character → time and place-specific dynamics
- Frequently open-source availability \rightarrow no or low cost
- Significant improvement of prediction models
 Data available for places where these were previously not available
- Data available for places where these were previously not available
 Unobtrustive measures potentially compensate several blaces and errors present in data
 collected by researchers for research purposes \rightarrow have the potential to better describe
 social phenomena (note: equivocal views, cf. weaknesses)
 Big data enable new data processing methods (e.g., NLP)

Weaknesses

 Larger measurement error, because one may be unaware of how distillated variables from big datasets were measured \rightarrow need to infer the meaning of the data (note: equivocal views, cf. strengths)

ww.ircp.ugent.be

www.ircp.ugent.be

- · Sampling and selection bias: availability and 'digital devide'
- Does not always provide opportunities for longitudinal research

Discussion

- Technology is value-free (neutral)
- · Technology is the means, it is not the end goal
- · Methodology is fallible

Conclusion

(1) It takes two to tango

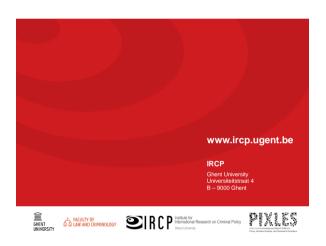
Cooperation and communication between computational (social) scientists and domain experts. For criminology: large opportunities for collaboration between or integration of 'analytical criminology' (scientific realism) and computational social science Also applies to policing!

(2) Framework to assess the quality of data sets and methods (cfr. TSE)

(3) FAT: Fairness, Accountability, Transparency in research and practice

wirch ugent be

research	publications	consultancy	conferences
) R		de for ational Rasearch University	on Criminal Policy



Bibliography

Chen, C., Ma, J., Susilo, Y., Liu, Y., & Wang, M. (2016). The promises of big data and small data for travel behavior (lak human mobility) analysis. *Transportation Research Part C, 68,* 285-299.
Connelly, R., Playford, C. J., Gayle, V., & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. *Soci Science Research, 59,* 1-12.
Cothett, R., & Marx, G. T. (1991). Critique: No soul in the new machine: Technofallacies in the electronic monitoring movement. *Lustice Quarterly, 9(3),* 399-414.
Custers, B., & Vergouw, B. (2015). Promising policing technologies: Experiences, obstacles and police needs regarding law enforcement technologies. *Computer Law & Security Review, 31,* 518-526.
Jadeja, B., & Issa, T. (2013). Big Data: A new technology trend and factors affecting the implementation of big data in Australian industries. In: G. Skourletopoulos, G. Mastorakis, C. X. Mavromoustakis, C. Dubre & E, Pailis (Eds.), *Mobile Big Data: A new technology trend models to technologies*. (LAW Notes on Data Engineering and Communications Technologies, Vol. 10, pp. 259-288). Cham: Springer International Publishing.
Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography, 33,* 262-267.
Kitchin, R., & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data Society, 31,* 1-10.
Thakuriah, P. V., Tilahun, N. Y., & Zellner, M. (2017). Big data and urban informatics: Innovations and challenge discovery, In: P. V. Thakuriah, N. V. Tilahung M. Zeilner (Eds.), *Secing clies through big data: Research, methods and applications in urban informatics* (pp. 114-5). Cham: Springer.

ww.ircp.ugent.be

Making the most of the available data: Social capital variables at the neighbourhood level as predictors in a predictive policing model

Anneleen Rummens Wim Hardyns

17 September 2019 Predictive policing in the big data era CPS international conference 'Street policing in a smart society'



Contents

I. What is predictive policing?

II. (Big data) sources for predictive policing

III. Social capital research in Ghent: SWING and SCAN

IV. Predictive modelling using SWING Conclusions

SIRCP Institute for International Research on Otherinal Policy Over University

I. What is predictive policing?

- Predictive policing: "the use of historical data to create a spatiotemporal forecast of areas of criminality or crime hot spots that will be the basis for police resource allocation decisions with the expectation that having officers at the proposed place and time will deter or detect criminal activity" (Ratcliffe 2014, p. 4)
- Recent development: use of big data and predictive analysis in criminology
- Evolution to small units of analysis (street segments, grids) and complex machine learning methods
- More dynamic than methods such as hotspot analysis



I. What is predictive policing?



I. What is predictive policing

Main objectives of predictive policing:

- Short-term:
 - Use available resources more efficiently
 - Proactively target crime
- Long-term:
 - Cost-efficiency
 - Decrease crime rate

SIRCP Institute for International Research on Driminal P Gene Unwersty

II. Predictive policing using the SWING data

• SWING: Social capital and Well-being in Neighborhoods in Ghent (2011-2015) (Hardyns et al. 2015)

- Extensive study of relationship between social capital, crime and health
 Collected data on among others social and physical disorder, social trust, informal control, ...
- Doctoral research at Ghent University: Performance of different data sources and methods for predictive policing?
- Retrospective analysis: predicting crime for a period that has already happened, so that we can compare the predicted zones to the real crime zones
- Using the SWING data, a predictive model for home burglary in Ghent was created

rcp.ugent.be

SIRCP Institute for International Research on Criminal Pole Oper University

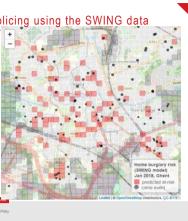
II. Predictive policing using the SWING data

Including the SWING variables improved our base model (only crime history variables) with 1% for both recall and precision (a limited increase)

Recall: 20,18%
(% incidents correctly predicted)
Precision: 17,00% (% predictions which are true events)

N incidents = 122 N predictions = 100

SIRCP Institute for International Otors Diver



II. Predictive policing using the SWING data

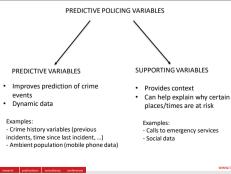
Advantages and disadvantages of using social capital variables in predictive policing

Advantages	Disadvantages		
 Provides information on residents' concerns 	Difficult to acquire dynamically		
Provides inside information	 Can be misleading for predictive purposes 		
 Social capital has been shown to have a strong relationship with crime 			

<u>Conclusion</u>: social capital variables (as collected by the SWING study) have low predictive value, but can be useful as supporting variables

SIRCP





III. Selecting variables for predictive policing

- One of the main characteristics of predictive policing, is its use of (big data) sources
- => More advanced systems even provide real-time integration
- New technologies open up new possibilities: ANPR, mobile phone data, ...
- Some applications are also looking at civilian inputs, e.g. 'citizen science' apps

=> This kind of data can also be of interest in the context of predictive policing, for example, with regards to fear of crime, well-being, social control, cleanliness, ...

SIRCP Institute for Other Conception Research on Christel Policy Citer Christel

III. Selecting variables for predictive policing



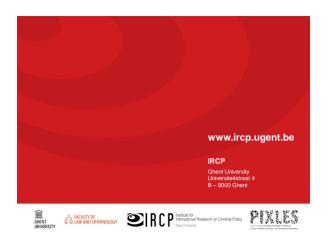
Such an approach could also help to make data more dynamic, thus increasing its predictive value

SIRCP Institute for International Research on Driminal F

Conclusions

- Making the most of predictive policing also requires making the most out of the available (big) data
- Predictive vs supporting variables, e.g. social capital data
- Use of big data & new technologies
- General guidelines for selecting variables for predictive policing, but specific context is important
- \Rightarrow Explorative analysis needed before implementing predictive policing \Rightarrow Test different configurations
- Important questions to ask:
- Which data resources are available?
- Which data could be interesting to include?
 For prediction?
- For prediction?For support?

SIRCP Institute for Informational Research on Driminal Policy Over Drivershy

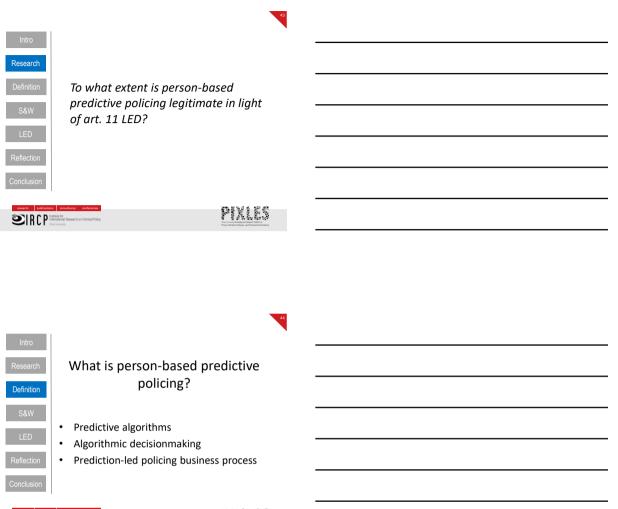




Intro Research Definition S&W LED Reflection Conclusion

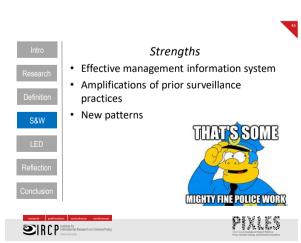
"To what extent are person-based predictive policing tactics legitimate in light of what EU data law is and ought to be in a privacy-friendly mass surveillance society?"

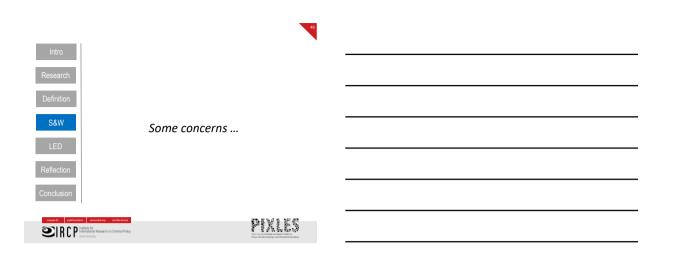


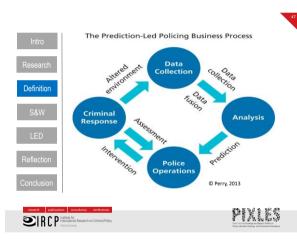


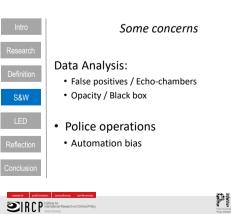


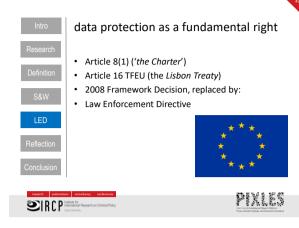


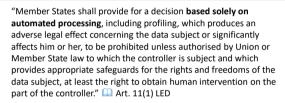
















"Member States shall provide for a decision based solely on automated processing, **including profiling (...)**"

- Predictions (vs. decision)
- Some form of automated processing (e.g. data mining)

```
SIRCP Institute for
International Research on Drininal Polic
Oper University
```

"Member States shall provide for **a decision** based solely on automated processing, including profiling (...)"

- What constitutes a decision?
- Process matters
- Stages in the algorithmic process



SIRCP Institute for Informational Research on Driving Policy Otoci University PIXLES Der Stande auf verschieders Neise indere internet in der seiter

"Member States shall provide for a decision **based solely on automated processing**, including profiling, (...)"

Solely

- without any human intervention?
- Ambiguous

SIRCP Institute for International Research on Ofminel Pol PIXLES

"Member States shall provide for a decision **based solely on automated processing**, including profiling, (...) "

What constitutes human intervention?

Nominal (human rubber stamping) or meaningful human intervention?





"Member States shall provide for a decision **based solely on automated processing**, including profiling, (...)"

Unauthorised law-making by Art. 29?

	N
\odot	
European Commission	
ARTICLE 29	

- "not any human involvement"
- "with an influence on the result"
- "meaningful"





"(...) a decision (...) which produces an adverse legal effect concerning the data subject or significantly affects him or her (...) "

Adverse legal effect (\leftrightarrow GDPR)

- No definition
- Discriminatory effects?

SIRCP Institute for International Research on Driminal Pol Control Diversity PIXLES

"(...) a decision (...) which produces an adverse legal effect concerning the data subject or significantly affects him or her (...) "

Significant effect

- Less clear
- "substantial"

SIRCP Institute for International Research on Criminal Pole

PIXLES

"(...) prohibited unless (...) appropriate safeguards for (...), at least the right to obtain human intervention on the part of the controller."

- Human intervention?
- Automation bias







"(...) prohibited unless (...) **appropriate safeguards** for the rights and freedoms of the data subject (...) "

Some clarity?

- to express point of view
- to obtain an explanation
- to challenge the decision

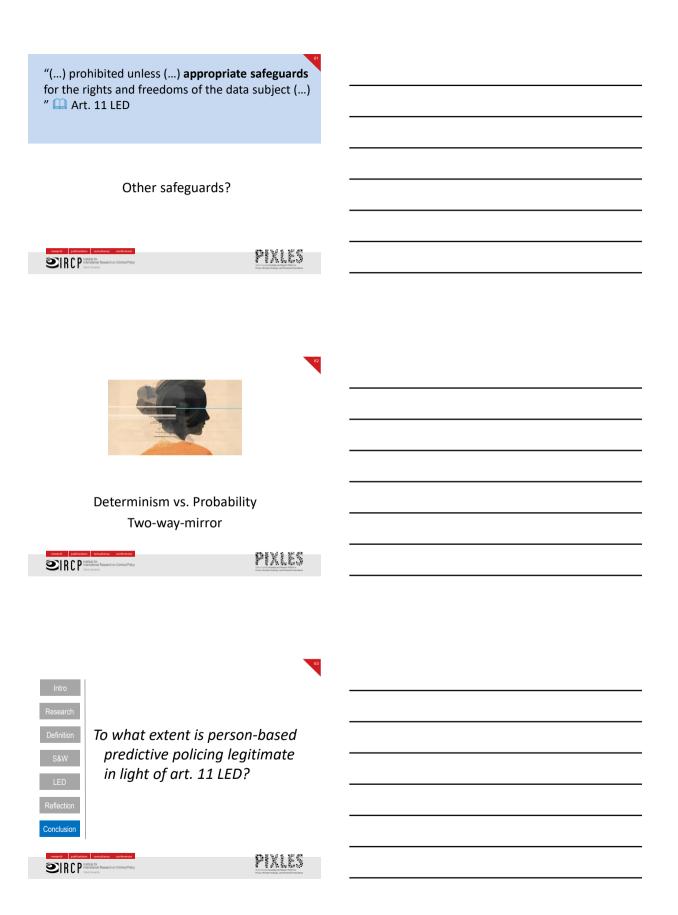


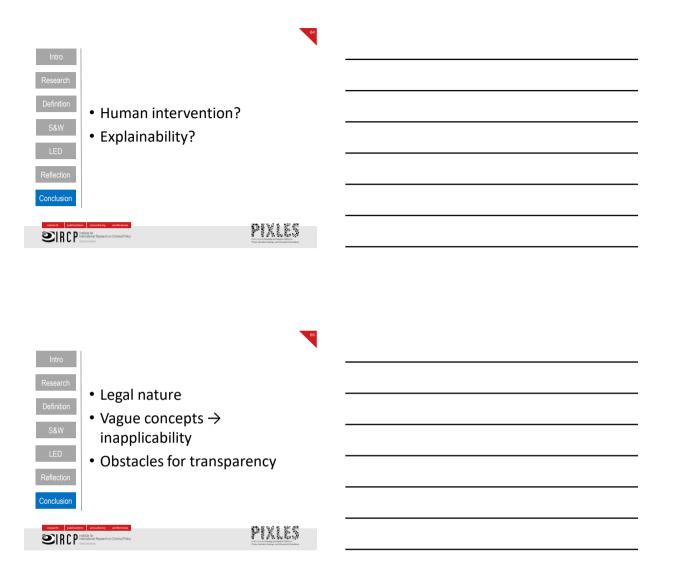


However,

- Recital/Directive \leftrightarrow art. 15 GDPR?
- Trade secrets
- Need for flexibility (recital 44)
- Black box

SIRCP Institute for Other Editional Research on Climinal Pol Other Editional Viewsky









Discussion: Predictive policing in the big data era Chair: prof. dr. Wim Hardyns CPS International Conference 'Street Policing in a Smart Society' 17 September 2019



Discussion

Technology is value-free (neutral)

- The question is not whether or not we can predict the future. The question is more fundamental: what do we want to do with this knowledge?
- Based on linking data and data fusion, we can provide you a social credibility score, but do we want this?
- Based on new and emerging data collection methods, we can fine any non-paying parked car, but do we want this? And we can fine every car that drives too fast, but do we want this?

SIRCP Institute for Institute for Other University

Discussion

Technology is the means, it is not the end goal

"There is no soul in the new machine" (Corbett & Marx, 1991)

GUNS DONT KILL PEOPLE. STUPID PEOPLE WITH GUNS KILL PEOPLE.

SIRCP Institute for International Research on Criminal Pole Oper University

Discussion

Methodology is fallible

- There are some things big data can never quantify, but
- there are also quantities that human can never phathom

and



