Towards an evidence-based model for big data policing

Prof. Dr. Wim Hardyns
ERC Consolidator Grantee

18th International Conference on Location Based Services
Ghent, Belgium
20/11/2023
# Table of Content

1. The run-up to
2. From predictive policing to big data policing
3. BIGDATPOL: #ERCCoG
4. Future valorisation
1. The run-up to
2. From predictive policing to big data policing
Predictive policing: what’s in a name?

Predictive analysis techniques are used to:

1. predict offenders
2. predict victims
3. predict WHERE and WHEN there is a high risk of new criminal acts

‘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)

➢ See: https://link.springer.com/article/10.1007/s10610-017-9361-2
Examples applications

Hunchlab (sold to SoundThinking)

Geolitica (formerly PredPol)

Crime Anticipation System (Netherlands)
1. Which **statistical models** are useful?

1. Which **statistical models** are useful?

Ongoing study: comparing different ML methods across different EU cities

- Comparing prediction performance across EU cities
  - Brussels, Belgium
  - London, UK
  - Vienna, Austria

- Comparing prediction performance across ML methods
  - Ensemble Neural Network
  - Random Forest
  - K-Nearest Neighbors
2. What is the influence of predicted area and predicted time window?

See: [Link to article](www.sciencedirect.com/science/article/abs/pii/S0169207020300558)
3. To what extent is predictive policing context-specific?

3. To what extent is predictive policing context-specific?

- Structural area characteristics
- Spatiotemporal clustering – stability (crime, crime opportunity, etc.)
- Amount of data – number of crime incidents (examples of positive class)
- Quality of data – differences across settings/contexts
- Methods used – differences across settings/contexts

4. What opportunities do new (big) data sources offer for predictive policing?

See: [www.mdpi.com/2220-9964/10/6/369/htm](www.mdpi.com/2220-9964/10/6/369/htm)
4. What **opportunities do new (big) data sources offer for predictive policing?**

Automated data – Mobile phone data (N = 595,858,852)
4. What **opportunities do new (big) data sources offer for predictive policing?**

Automated data – Mobile phone data (N = 595,858,852)
4. What opportunities do new (big) data sources offer for predictive policing?
5. How can we **evaluate** predictive policing in practice?

Randomised Controlled Trial – Matched pairs

- **Use of predictive policing**
  - Test neighbourhoods
  - Matched on demographic and socio-economic characteristics

- **Use of business-as-usual approach**
  - Control neighbourhoods
3. BIGDATPOL: #ERCCoG
Since 2015, Professor Wim Hardyns was one of the first criminologists to conduct innovative research on big data and predictive policing. Today, knowledge about big data policing in Europe is still fragmented, with a lack of interdisciplinarity and a lack of scientific evaluations. In his ERC-funded big data policing project (BIGDATPOL), Wim Hardyns uses historical and real-time data to predict when and where the risk of new criminal acts is high.

The overarching goal of this ERC project is to integrate statistical-methodological, criminological, legal and ethical conditions into a single evidence-based 3D model. This model will be tested by several randomised controlled experiments in European settings. The approach of this project is innovative and radically different because it is transparent in terms of predictive algorithms, effectiveness, and legal and ethical safeguards. The ultimate goal of the evidence-based model is to provide both academia and law enforcement practice with guidelines and recommendations for studying, applying and implementing big data policing. This ambition is challenging and innovative, as it will be the first time that interdisciplinary research on this topic has been conducted in Europe, and even worldwide. Wim Hardyns therefore intends to present the results as a European benchmark for big data policing.
Big data policing:

**STEP 1**
BIG DATA COLLECTION

**STEP 2**
PREDICTIVE ANALYTICS

**STEP 3**
ACTIONABLE OUTPUT

Project objective:
Building an evidence-based 3D model for big data policing:

- RCTs
- Statistical Methodological
- Legal and Ethical
- Criminological
PHASE 1
- Data base
- Expert network
- Typology
- Data collection

PHASE 2
- Track 1: statistical-methodological
- Track 2: criminological
- Track 3: legal and ethical

PHASE 3
Result:
Evidence-based big data policing model
Challenges and opportunities

Experimenting with (big) data sources

- Social media data
- Ambient population data (e.g., mobile telephone data, ANPR)
- Satellite data

Experimenting with (big) data techniques

- Deep learning models
- Computer vision
- Natural Language Processing (NPL)

Effectiveness question not yet sufficiently answered, mainly due to lack of solid number of evaluation studies

Attention to ethical aspects and privacy: privacy-by-design !!!
4. Future valorisation
Scientific and socio-economic valorisation

i4S Business Development Center UGent

- **Cluster Technology**
- **Cluster Digitalization**
- **Cluster Privacy**

**Areas**
- **Security**
- **Crime**

**Partners**
- Private security services
- Data technology companies
- IT companies
- European LEA’s
- Privacy experts
Wim.Hardy@UGent.be
https://www.linkedin.com/in/wim-hardyns-7250a129/
Acknowledgement

Funded by the European Union (ERC, BIGDATPOL, 101088156). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.


Pauwels, L. (2002). De ene buurt is de andere niet: exploratie van mogelijkheden tot contextualisering van geregistreerde criminaliteit op buurtniveau. VUBPress.


