ColdCase: Solved & Unsolved

The use of new digital tools and data science techniques in cold case/homicide research

Tatjana Kuznecova, MSc
Jaap Knotter, PhD

CEPOL conference, Vilnius, 8-10 of June 2022

Project outputs

SAXION
UNIVERSITY OF
APPLIED SCIENCES

WP1. Data collection and (smart) database development

SAXION UNIVERSITY OF APPLIED SCIENCES

I. Data collection

Automated and semi-automated:

- Web scraping, retrieval of articles from various portals and databases
- OCR for newspaper archives

~ 11 000 articles from various sources

Manual:

- Extraction of detailed variables from articles based on ESC12 (de Kock, 2014)
- >200 variables extracted in a template

~ 300 homicide cases processed for years 2006-2015

POLITIE Politieacademie

WP1. Data collection and (smart) database development

II. Research on automatic processing of text with AI (NLP)

1. *Automatically retrieve components from text (limited success)*
2. Generate a description summary of an article
3. Distinguish articles about murder cases from other topics

NLP methods like BERT algorithm

SAXION UNIVERSITY OF APPLIED SCIENCES

POLITIE Politieacademie

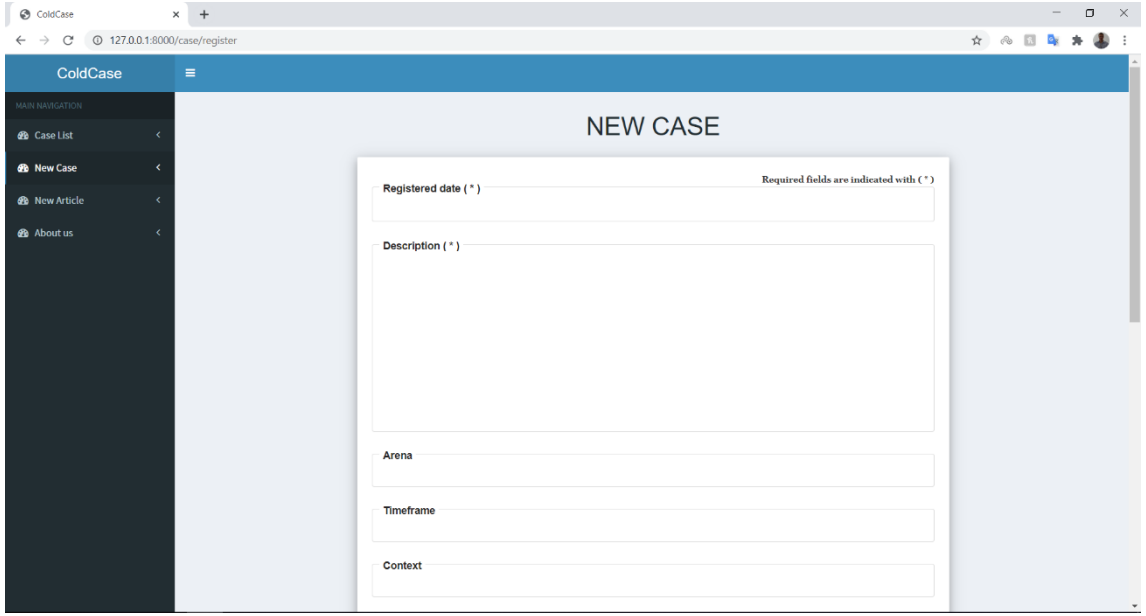
WP1. Data collection and (smart) database development

III. Search tool API

- Search and retrieve news articles from the web (prototype)

IIV. Database platform (prototype)

- Web-based Database platform with a User Interface for storage and editing of data
- *Connect DB with AI module (test for future developments)*



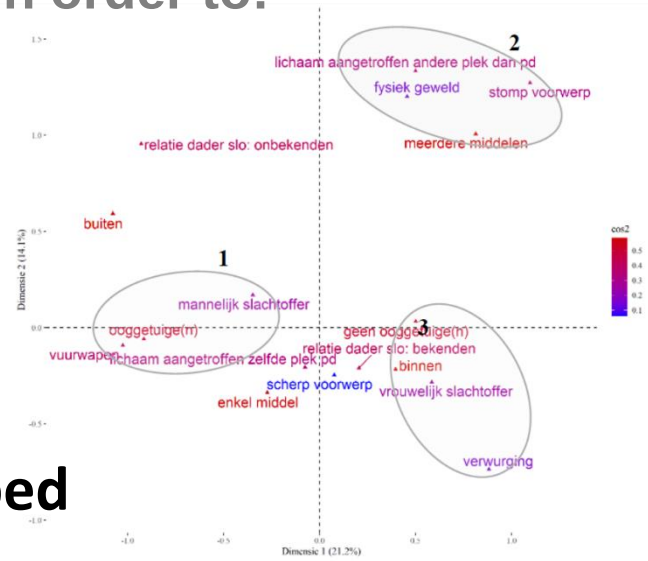
SAXION UNIVERSITY OF APPLIED SCIENCES

POLITIE Politieacademie

WP2. Data analysis

- **Can we identify meaningful relationships and patterns in the historical homicide data?**

- **Can we learn from those patterns in order to:**
  - Fill in the gaps in information
  - Help develop investigative scenarios



- **2 MSc and 1 BSc theses developed**

SAXION UNIVERSITY OF APPLIED SCIENCES

WP2. Data analysis:

MSc thesis:
Hanneke van de Mortel, University of Leiden

SAXION UNIVERSITY OF APPLIED SCIENCES

Predict interpersonal relationships between victim and perpatrator based on some case characteristics

Tabel 14

Correlaties tussen modus operandi en soorten familierelaties

Onafhankelijke variabelen		Filicide	Parricide	Siblicide	Partner-doding	Overige familierelaties
Plaats delict	Plaats delict	2.36[b]	1.93[b]	4.43[b]	11.79*[b]	11.56*[b]
	Openbare ruimte	-	-	-	2.93[a]	-
	Openbare weg	-	-	10.90[a]	.14[a]	.16[a]
	Woning	-	3.58[a]	.38[a]	3.08**[a]	4.69**[a]
	Horecagelegenheid	-	-	-	-	-
	Overig plaats delict	-	.60[a]	-	.43[a]	.39[a]
Doodsoorzaak	Doodsoorzaak	115.50**[b]	12.27[b]	3.64[b]	26.39*[b]	24.64[b]
	Schotwond	-	.33[a]	-	.48[a]	.25**[a]
	Steekwond	-	1.86[a]	-	1.00[a]	1.20[a]
	Snijwond	-	6.63*[a]	-	.31[a]	2.27[a]
	Geweld stomp object	-	3.19[a]	-	.56[a]	2.23[a]
	Verstikking	-	-	-	3.04[a]	.55[a]
	Verdrukking	-	-	-	-	-
	Verhanging	-	-	-	-	-
	Verbranding	-	-	-	-	-
	Verwurging	4.37[a]	-	-	1.42[a]	.80[a]
	Vergiftiging	-	-	-	-	-
	Overrijden	-	-	-	-	-
Gebruikte	Wapengebruik	23.11*[b]	12.48[b]	4.12[b]	18.76[b]	33.07**[b]
	Verstikking	2.12[a]	-	-	1.08[a]	1.25[a]
	Scherp voorwerp	-	3.46[a]	-	.95[a]	1.06[a]
	Stomp voorwerp	-	2.95[a]	-	.83[a]	2.01[a]
	Fysiek geweld	8.78**[a]	-	-	.67[a]	.76[a]

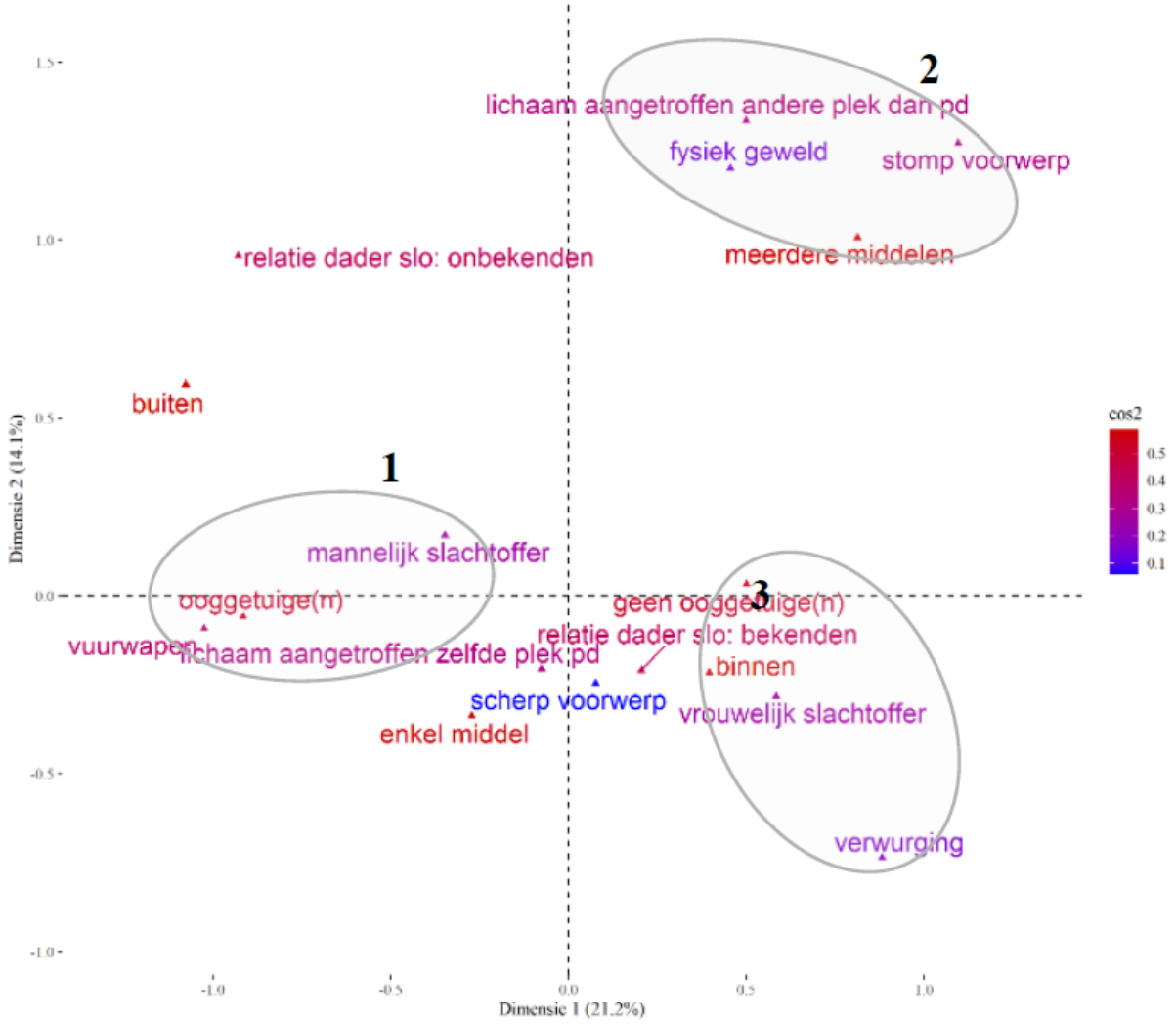
WP2. Data analysis:

Multivariate analysis

MSc thesis:
Rob Schipperheijn,
Politie Academie

SAXION
UNIVERSITY OF
APPLIED SCIENCES

Co-occurring characteristics in homicide cases



WP2. Data analysis

BSc thesis:

Izzy van der Veur,

Hogeschool van Amsterdam

SAXION
UNIVERSITY OF APPLIED SCIENCES

Link homicide types to socio-economic, geographic and demographic factors

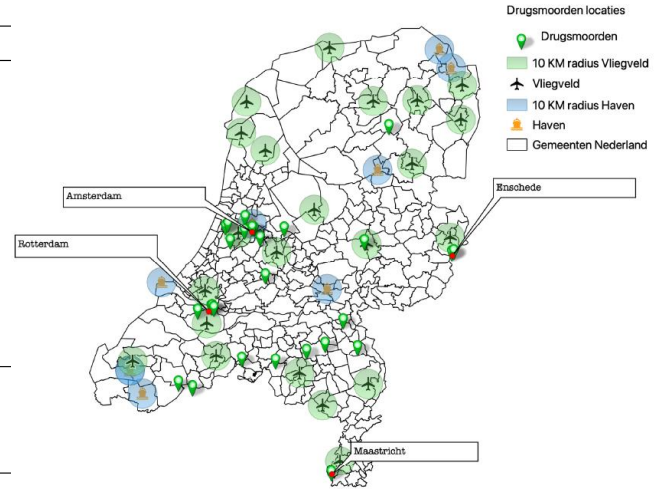
Tabel 17: Odds Ratio tussen sociaaleconomische, geografische en demografische kenmerken en moorden binnen het criminele circuit

Sociaaleconomische, geografische en demografische kenmerken	criminele circuit		
	Liquidatie	**drugs**	**anders**
GEMINKOMEN	0.327	3.800	0.862
GEMWONING	0.917	1.292	0.766
GINI	3.394	0.286	0.986
PROLEEFTIJD	1.065	0.977	0.926
GEMWERKELOOS	0.278	4.083	0.783
PERARMOEDE	1.541	0.406	2.406
VERDELINGMAN	2,000	0.559	0.774
GEMHUISHOUDEN	0.489	2.076	1.100
PEROPLEIDING	0.333	3.600	0.789

Tabel 18: Pearson Chi-squared correlatie tussen sociaaleconomische, geografische en demografische kenmerken en moorden binnen het criminele circuit

Sociaaleconomische, geografische en demografische kenmerken	Criminele Circuit		
	Liquidatie	**Drugs**	**Anders**
GEMINKOMEN	5.589*	7,666**	0,057
GEMWONING	0,038	0,295	0,186
GINI	5,001*	5,519*	<0,000
PROLEEFTIJD	0,014	0,002	0,11
GEMWERKELOOS	4,429*	6,093*	0,087
PERARMOEDE	0,767	3,213	1,228
VERDELINGMAN	1,498	1,054	0,125
GEMHUISHOUDEN	2,383	2,32	0,023
PEROPLEIDING	4,530*	6,363*	0,111

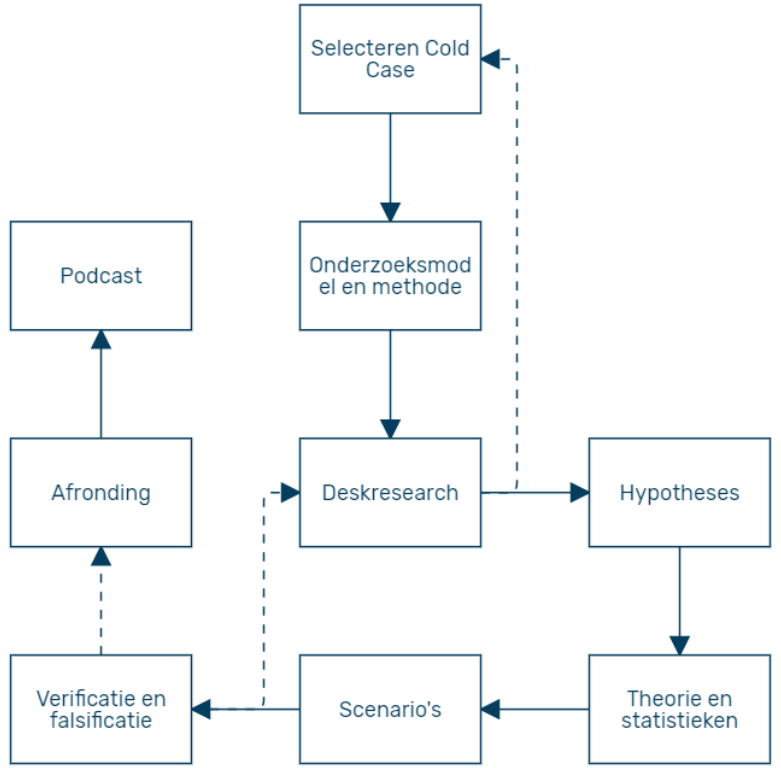
***. P< 0.01 (2-tailed). *. P< 0.05 (2-tailed).*



MOORD atlas

WP3: Digital Cold Case Blueprint

- Cold Case evaluation structure
- Digital documentation of decisions, evidence, data sources
- Track decisions and work out scenarios
- Maintain a clear overview of investigative process



❖ Currently educational version

WP4-5: Application in education and training

SAXION UNIVERSITY OF APPLIED SCIENCES

Use developed technology and knowledge in education:

Educational program:

Cold Case Minor course (Saxion University)

- Students learn about new methods and tools
- Students participate in data collection
- Students learn to work with open-source data
- Students test tools and provide feedback
- ..while working on cold case reconstructions

POLITIE Politieacademie

Conclusions

SAXION UNIVERSITY OF APPLIED SCIENCES

POLITIE Politieacademie

1. Automation of data processing and interpretation is still difficult
    - Training data!!!
    - Context interpretation
2. Do not overestimate the power of AI: use in combination with human interpretation capabilities
3. More and better data is needed to explore complex relationships and patterns
4. Data for AI training. Labelling.
5. Need for data quality control and revision mechanism
6. Close collaboration with Law enforcement professionals is essential
7. Knowledge exchange between research groups -> work together!
8. Ethical issues regarding predictive AI have to be addressed
9. **AI and Tunnel Vision: preventing or reinforcing??**

SAXION HOGESCHOOL

POLITIE Politieacademie

Future directions

1. Further data collection
    - Explore volunteer engagement
2. Further develop AI components
    - Automation of data collection and processing
    - Predictive AI -> *e.g. fill in gaps in crime storylines*
    - Generate storylines/scenarios from components
4. Research on ethical and privacy issues
5. New partnerships and funding applications
6. Collaboration with volunteer/citizen groups (Bureau Dupin, Netherlands)

SAXION HOGESCHOOL

POLITIE Politieacademie

Contact us

Tatjana Kuznecova:

t.kuznecova@saxion.nl

Jaap Knotter:

j.c.knotter@saxion.nl

SAXION HOGESCHOOL

POLITIE Politieacademie