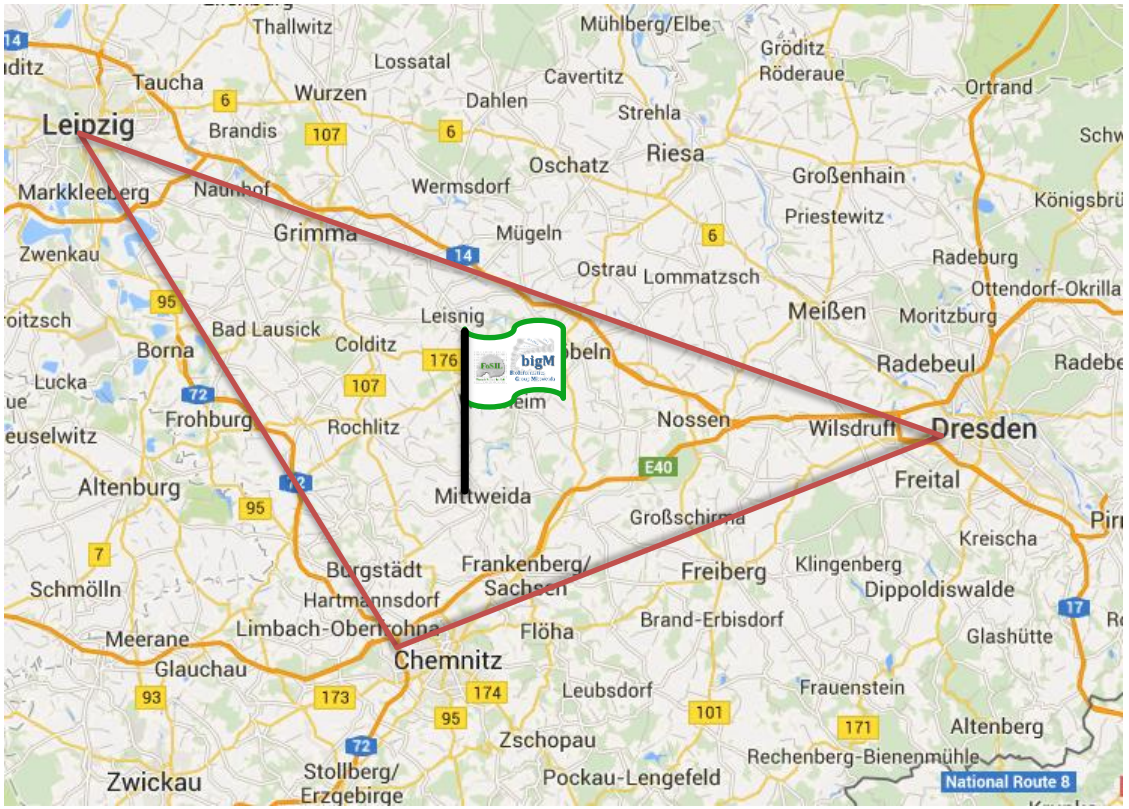
„Something, that is allegedly secure
is not necessarily secure,
Something, that is allegedly known
might turn out to be unknown.
Appearance can be deceptive,
our senses can deceive us.
Even though experience and knowledge
can limit errors,
reality also limits those.“

Inspired by Berthold Brecht

Author: unknown

Mittweida

Schleswig Holstein
Mecklenburg Vorpommern
Hamburg
Bremen
Niedersachsen
Sachsen Anhalt
Brandenburg Berlin
Nordrhein-Westfalen
Hessen
Thüringen
Sachsen
Rheinland Pfalz
Saarland
Baden-Württemberg
Bayern



Leipzig
Dresden
Chemnitz
Mittweida
Freiberg
Zwickau
Altenburg
Frohburg
Borna
Grimma
Wurzen
Oschatz
Riesa
Meißen
Radebeul
Freital
Nossen
Wilsdruff
Großenhain
Priestewitz
Radeburg
Ottendorf-Okrilla
Moritzburg
Mügeln
Ostrau
Lommatzsch
Leisnig
Colditz
Rochlitz
Bad Lausick
Burgstädt
Hartmannsdorf
Frankenberg/Sachsen
Limbach-Oberfrohna
Flöha
Brand-Erbisdorf
Klingenberg
Dippoldiswalde
Kreischa
Glashütte
Frauenstein
Altenberg
Rechenberg-Bienenmühle
Leubsdorf
Zschopau
Pockau-Lengefeld
Stollberg/Erzgebirge
Glauchau
Meerane
Schmölln
Lucka
Markkleeberg
Zwenkau
Taucha
Brandis
Thallwitz
Lossatal
Dahlen
Cavertitz
Strehla
Mühlberg/Elbe
Gröditz
Röderaue
Ortrand
Wermsdorf
Königsbrü
National Route 8

HOCHSCHULE MITTWEIDA
UNIVERSITY OF APPLIED SCIENCES

Fraunhofer SIT

FoSIL Forensic Science Inv. Lab

The Communication of Criminal Groups on the Internet

New Approaches in Predictive Policing

Dirk Labudde

Mittwoch, 1. August 2018

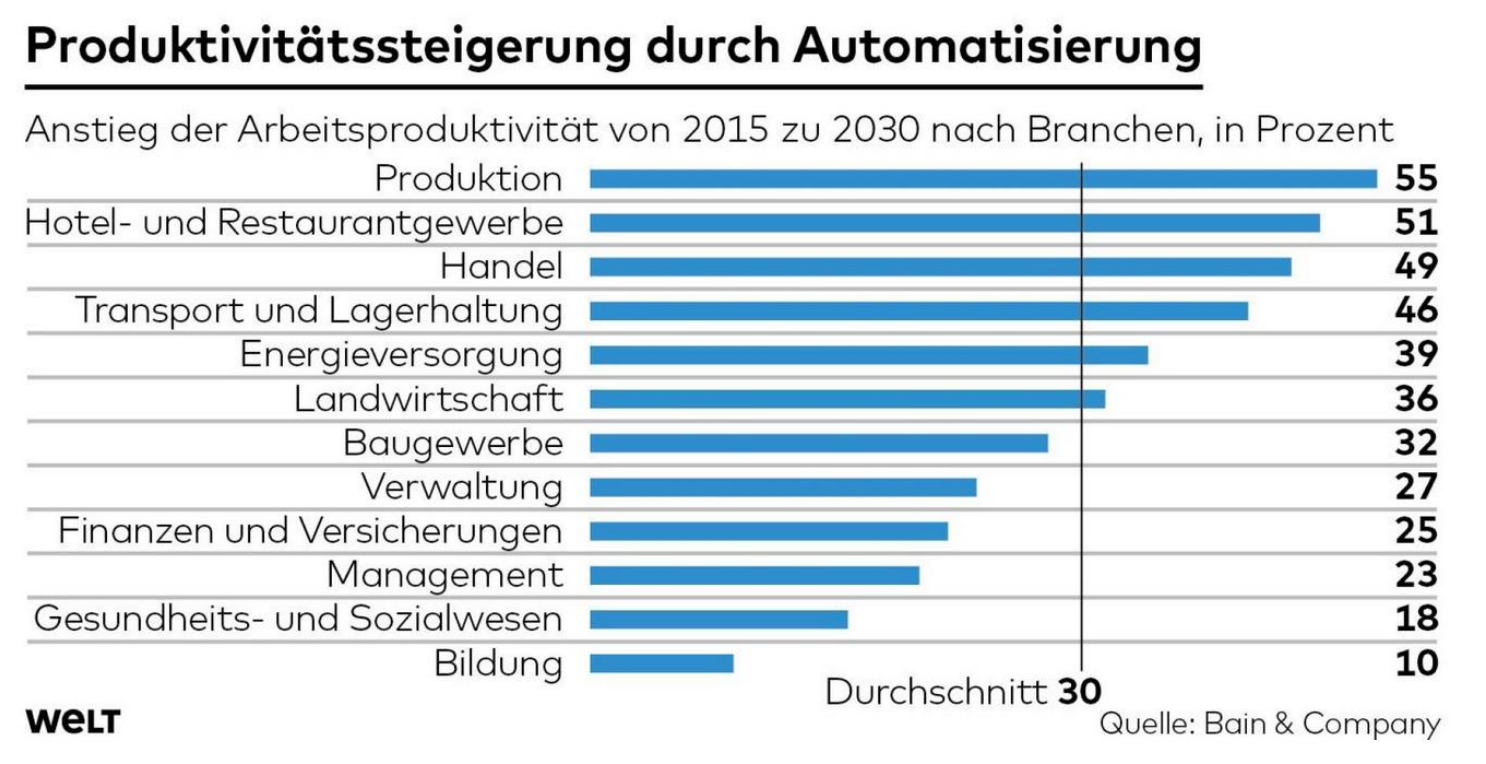
FoSIL

Digitization of society

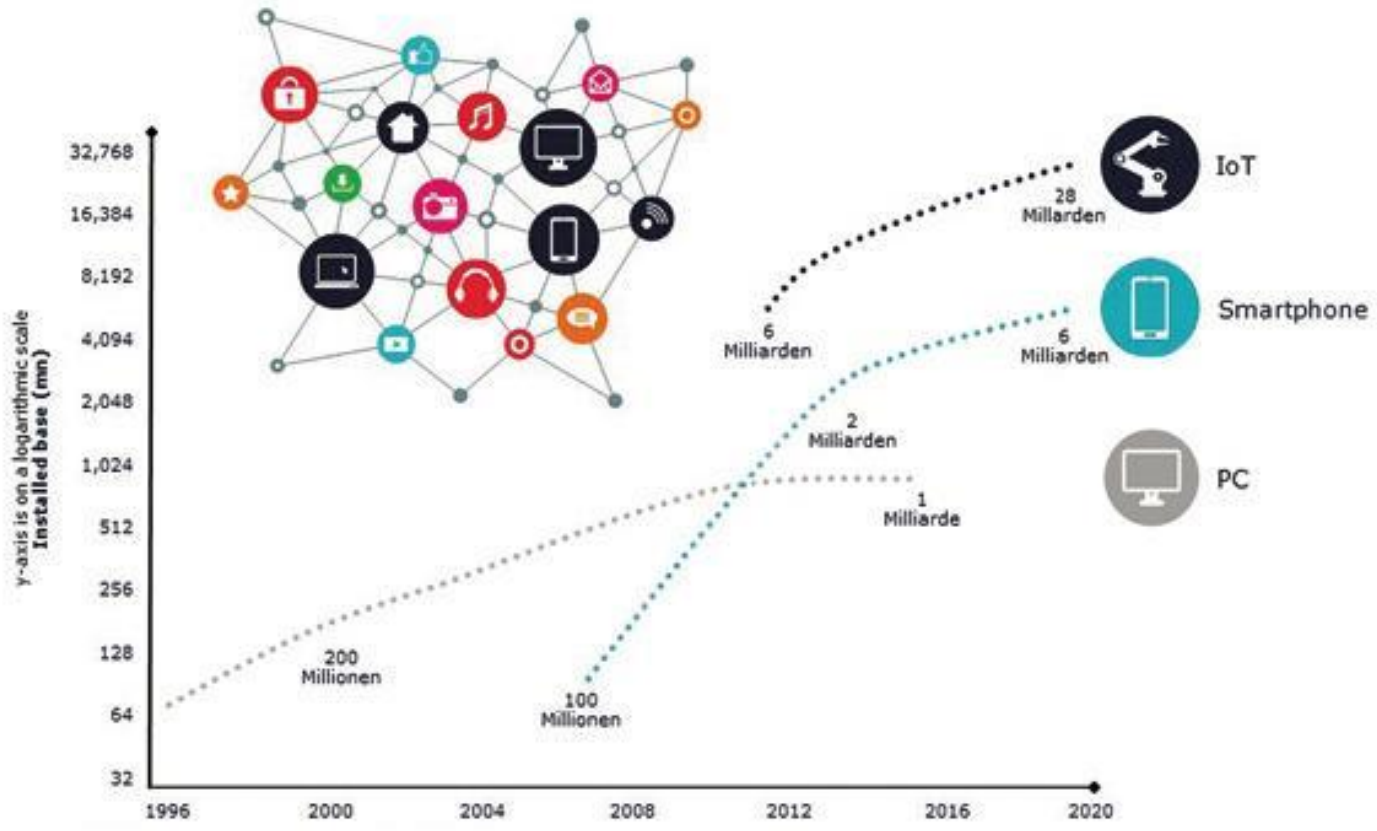
Digitization is in full swing. It affects us all - and makes for a profound change in every area of life. The digital transformation offers great opportunities for a better quality of life, revolutionary business models and more efficient management.

"Automation will reshape entire economies, turmoil labor markets and change the rules of the game in many industries"

Karen Harris, Austin Kimson and Andrew Schwedel (2018)



Digitization of society

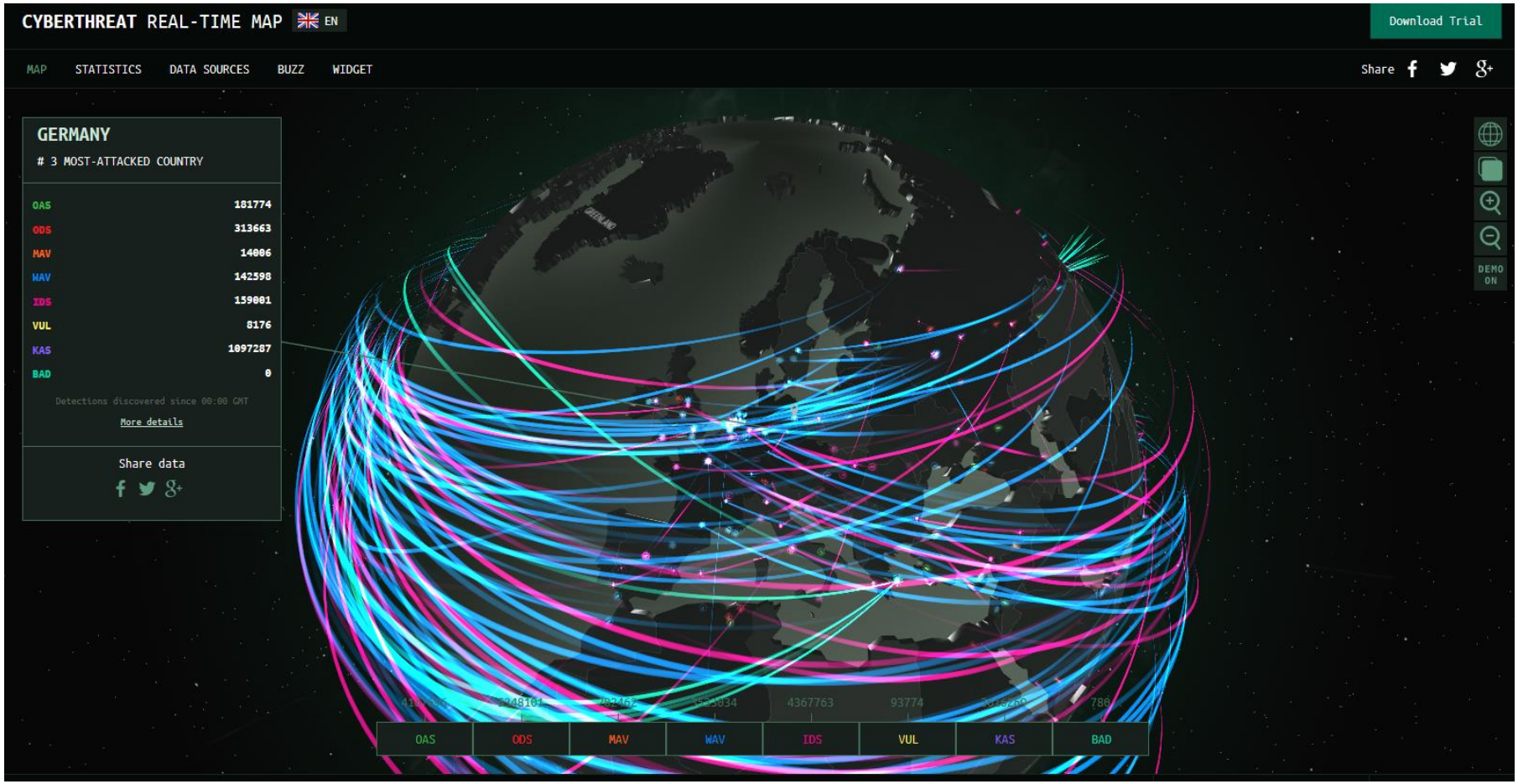


With the proliferation of mobile devices, the networking of our society has grown rapidly. The IoT exponentially increases this tendency.

Modern society is globally networked, we communicate in seconds with friends, acquaintances and business partners worldwide. However, the **positive possibilities** of using the Internet also have **negative side effects**: cybercriminals have a variety of opportunities.

Crime moves to the Internet, new crime phenomena arise.

Crime vs revolutionary business


CYBERTHREAT REAL-TIME MAP
EN
Download Trial
MAP
STATISTICS
DATA SOURCES
BUZZ
WIDGET
Share
GERMANY
3 MOST-ATTACKED COUNTRY
OAS 181774
ODS 313663
MAV 14006
WAV 142598
IDS 159001
VUL 8176
KAS 1097287
BAD 0
Detections discovered since 00:00 GMT
More details
Share data
DEMO ON
4367763
93774
OAS
ODS
MAV
WAV
IDS
VUL
KAS
BAD

MedEvolve leaked over 200,000 patient information due to server vulnerabilities

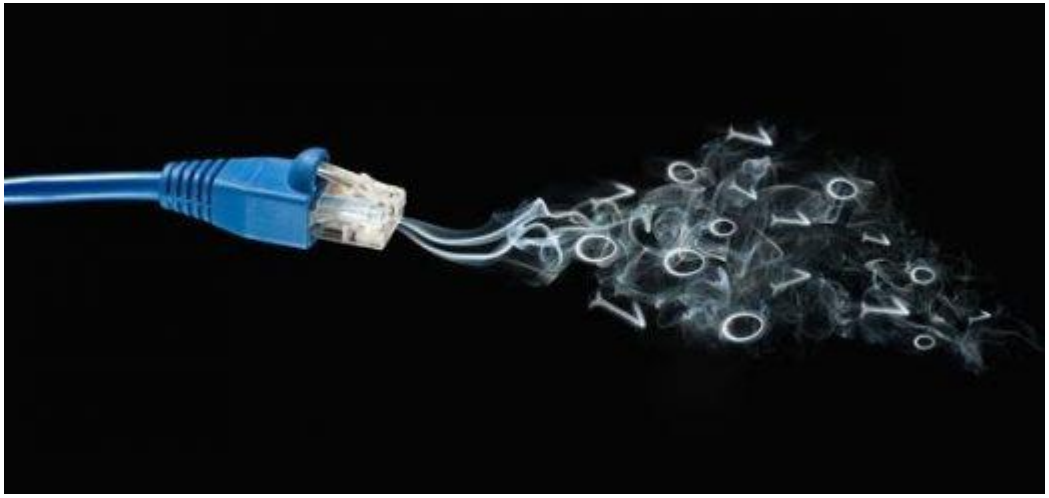
MedEvolve is a medical software company based in Arkansas, USA. According to a press release issued on July 10, 2018, MedEvolve faced a data breach that exposed the **patient's personal information**. On May 11, 2018, when the company's staff found a file containing patient data on an FTP server, they noticed the vulnerability. It is worth noting that anyone can access this file.

Although databreach[.]net clearly announced that the exact amount of data leaked was 205,000, no such data was disclosed in MedEvolve's press release.

https://securityonline.info/medevolve-leaked-over-200000-patient-information-due-to-server-vulnerabilities/

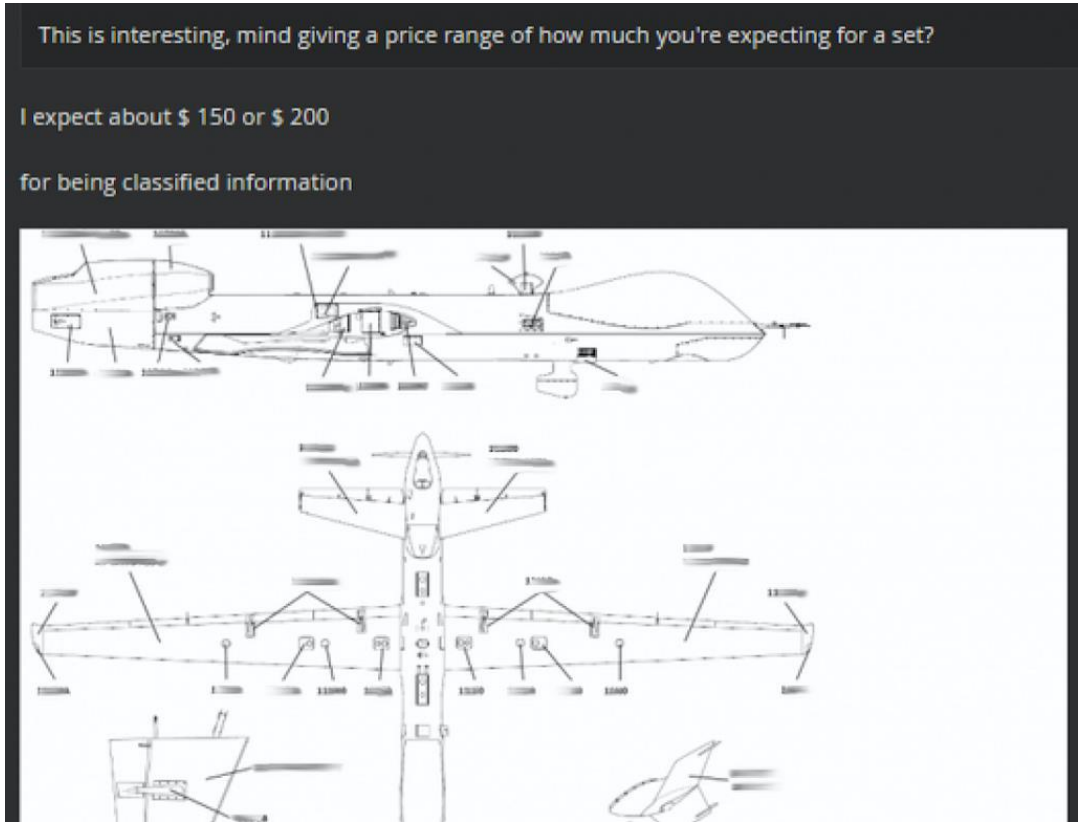
Login information for a large number of Mega accounts has been compromised

Mega, a company founded in New Zealand that provides **online cloud storage and files hosting services**, is currently found to have thousands of account credentials on its platform that have been publicly posted online. The leaked information is provided as a text file. It is understood that this text file contains more than **15,500 usernames**, **passwords, and file names**, which means that these accounts have been abnormally logged in, and the file name in the account is also crawled.



https://securityonline.info/login-information-for-a-large-number-of-mega-accounts-has-been-compromised/

Hackers sell US military drone documents for $150

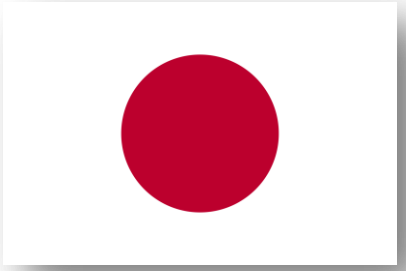


https://securityonline.info/hacker-sell-us-military-drone-documents-for-150/

200 million Japanese netizens' personal data offered on the underground market

In a survey released on Thursday, May 17th, the **cybersecurity company FireEye** stated that the company's security team has discovered a set of data sets being sold on underground hacking forums. These datasets involve a large number of sensitive information includes **personal identity information** (PII) of more than **200 million Japanese netizens**.

According to the report description, these data sets were discovered in early 2017, and a cyber hacker suspected of living in China's Zhejiang Province is advertising for the sale of these data sets. The contents of the data set include name, login credentials (account number and password), email address, date of birth, phone number, and home address.



https://securityonline.info/200-million-japanese-netizens-personal-data-offered-on-the-underground-market/

Australia's largest bank Commonwealth lost 12 million user bank data

Foreign media BuzzFeed reported that **Australia's largest commercial bank**, the Commonwealth Bank of Australia (CBA), confirmed that two storage tapes containing customer's name, address, account number, and transaction details record from 2004 to 2016 were identified in a data center transfer mission. Its subcontractor Fuji-Xerox was lost. Including at least **12 million users of bank transaction data**.

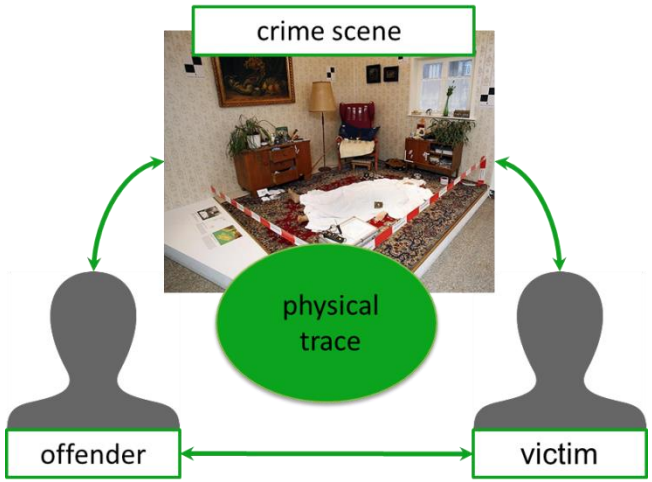
https://securityonline.info/australias-largest-bank-commonwealth-lost-12-million-user-bank-data/

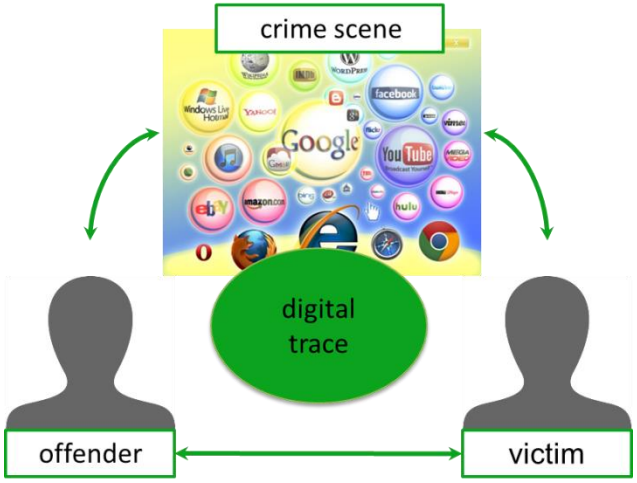
Crime moves to the Internet, new crime phenomena arise.

- Cybercrime -

Cybercrime covers the offenses that are directed against the Internet, data networks, information technology systems or their data or that are committed by means of this information technology.

Systems forensics - Basic principle





Dr. Edmond Locard (1877 - †1966)

Locard's exchange principle

holds that the perpetrator of a crime will bring something into the crime scene and leave with something from it, and that both can be used as forensic evidence

link real und virtual world

MAC-, IP-Adress, net information Photos/pictures

Systems forensics – traces (data)

- Analyses
- Validation
- **Evaluation**

circumstances of a crime

Physical or digital traces

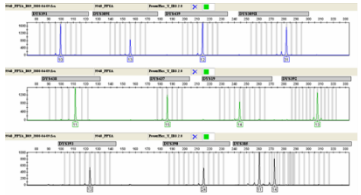
evidence

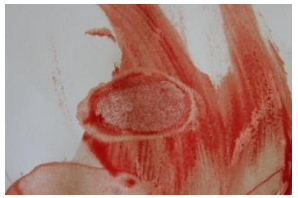
- fingerprint
- DNA traces
- traces of blood and pattern

- **Texts**
- **Short massages**
- **Chats**
- **…**

comparison of traces:

- Methods for the comparison
- Definition of similarity
- Evaluation of the results

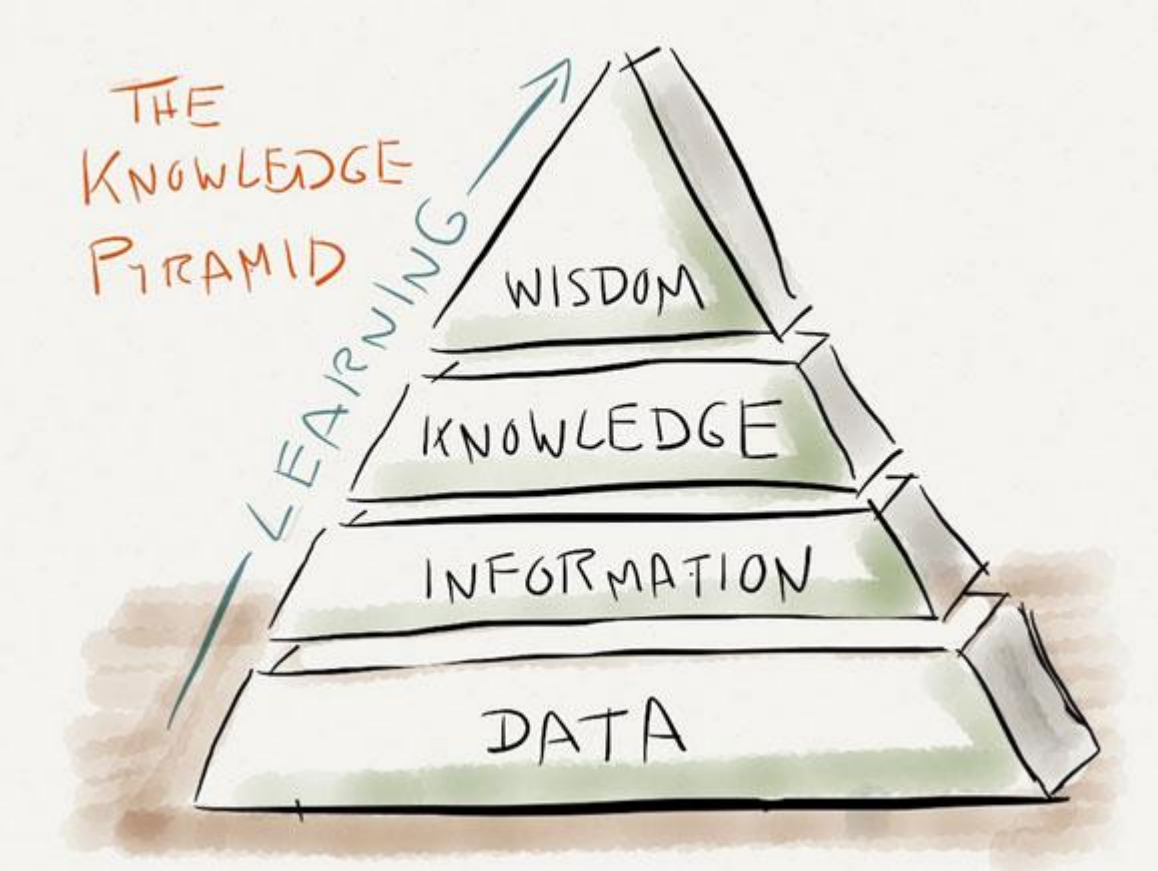




Digital Data and Digital traces

Human Individual Digital Footprint

Data … Information … Knowledge



System understanding

Models Modeling

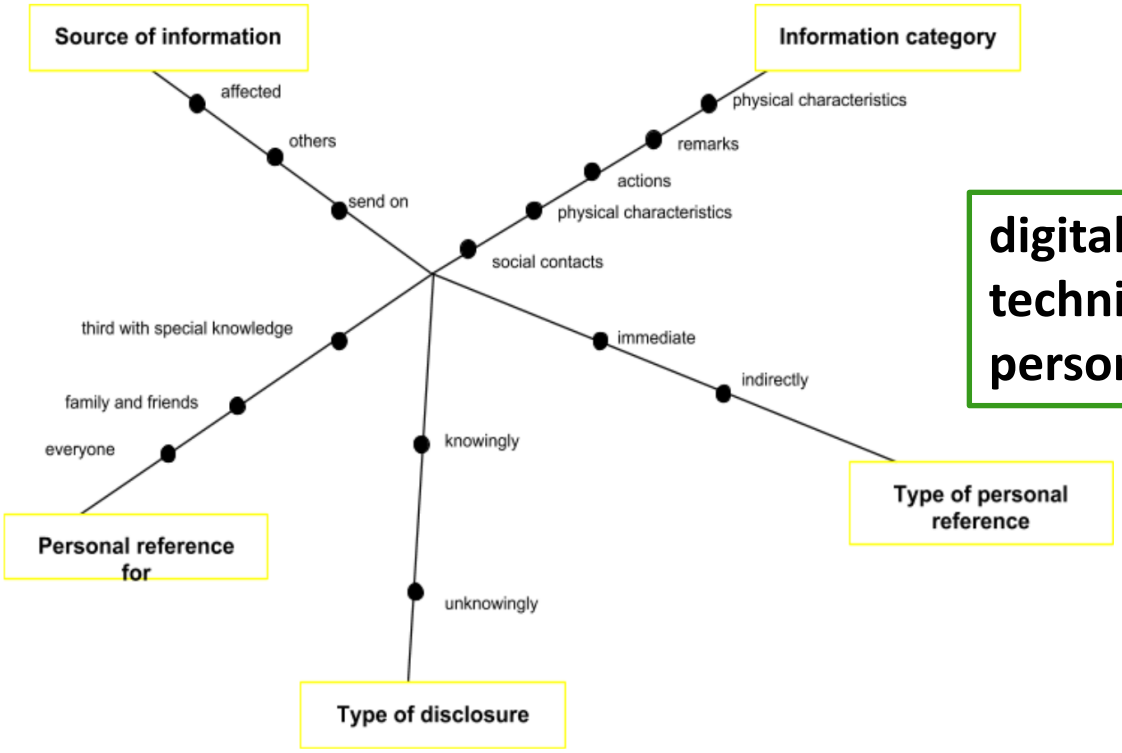
relationship

Data, counts, numbers
the lowdown

Modeling Models <–> real System

Our identity – our data

The **Identity** is the unique personality structure of a human being: **who am I, who am I referring to, who refers to me, what do I define myself and what makes me special?**



digital Identity: Any form of technical data that belongs to a person.

personal Data: The 5 Dimensions

Modern communication

Communication and content have changed due to the digitalization.

Digital identity and real identity

Modern communication and crime

Modern ways of communication is not only used in order to perform criminal acts, yet especially to plan criminal activities.

?

Consideration of digital communication for the modeling and simulation of criminal offences.

Combating crime

Digitization of society

4 Dimensions of BigData

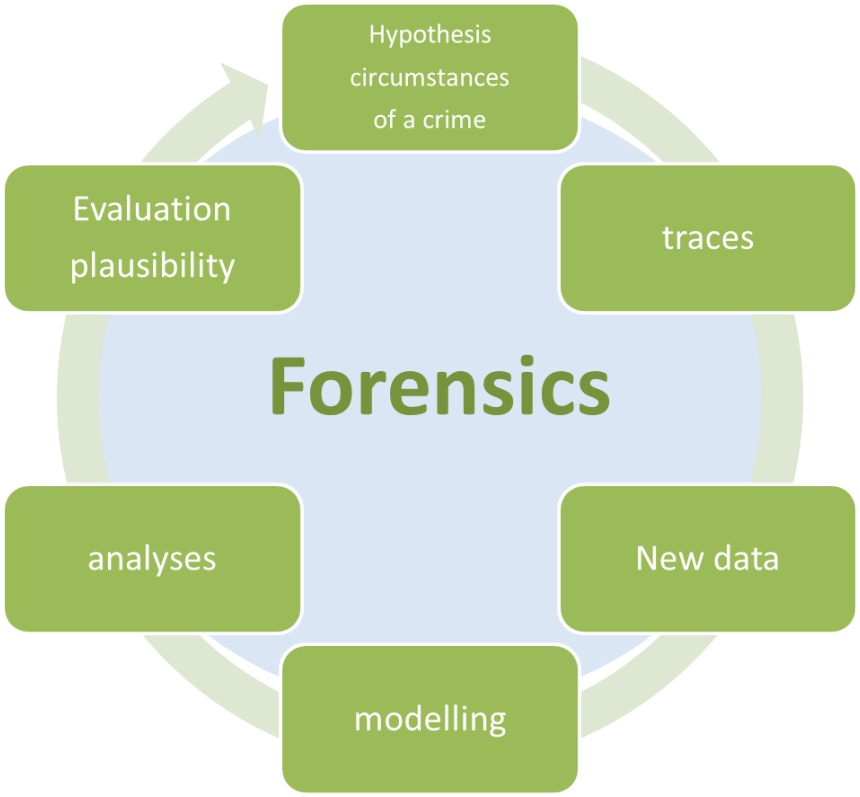
Volume	Variety	Velocity
mounds of data	data and data formats	Data streams

accuracy → uncertain data: **reliability management** and **predictability of inaccurate** data types

Short messages | **Digital Texts** | **Social Network**

Big Data analyzation -- prediction of system states

Cybercrime - systems forensics

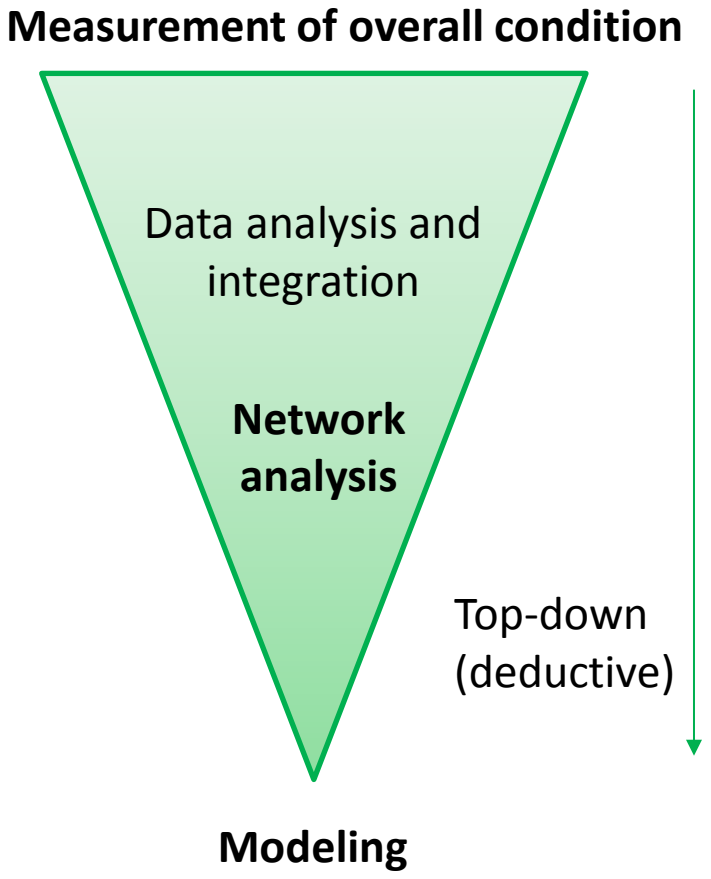


Top-Down - Approach

Approaches for modern forensics from the system theory
Methods in information gathering

Top-down (deductive)

- Creating experimental data
- Analysis of data and formulation of hypotheses (iterative)
- Validate hypotheses by means of experiments (iterative)
- **Goal: Identification and characterization of new mechanisms leading to a better understanding of the complex relations in a „populated urban structure“**
- Integration in the resilience process



Top-Down - Approach

A **Simulation** is an approach used for the **analysis** of systems which are too complex to be analyzed theoretically or with formulas. This is mostly the case for **dynamic system behavior**. In a simulation, experiments are performed with a model in order to gain knowledge about the real system.

- Discovering of network structures on different levels
- Gathering of the temporal and spatial dynamic of „system“ components using different (environmental) conditions
- Development of detailed mathematical models
- Visualization of networks and the processes taking place within those networks

-> Understanding of the processes as a whole

Can we use the *Human Individual Digital Footprint* in the modern forensics?

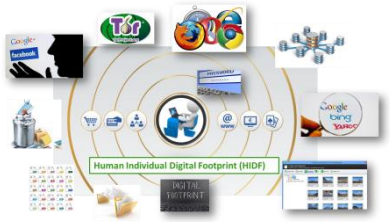
Digital communication and predictive policing

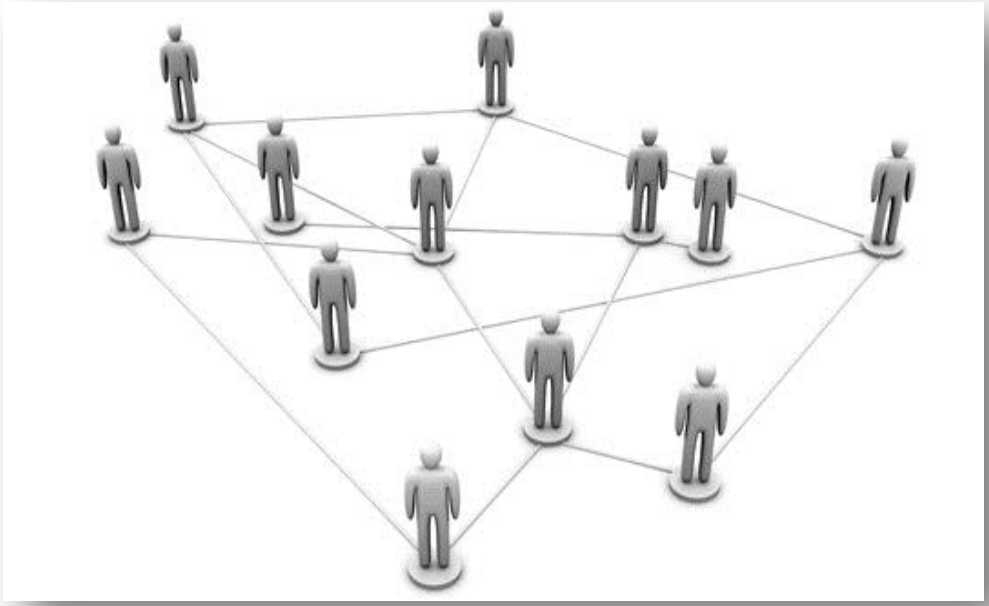
Communication

Communication of Criminal Groups on the Internet

Communication

Human Individual Digital Footprint (HIDF) Network



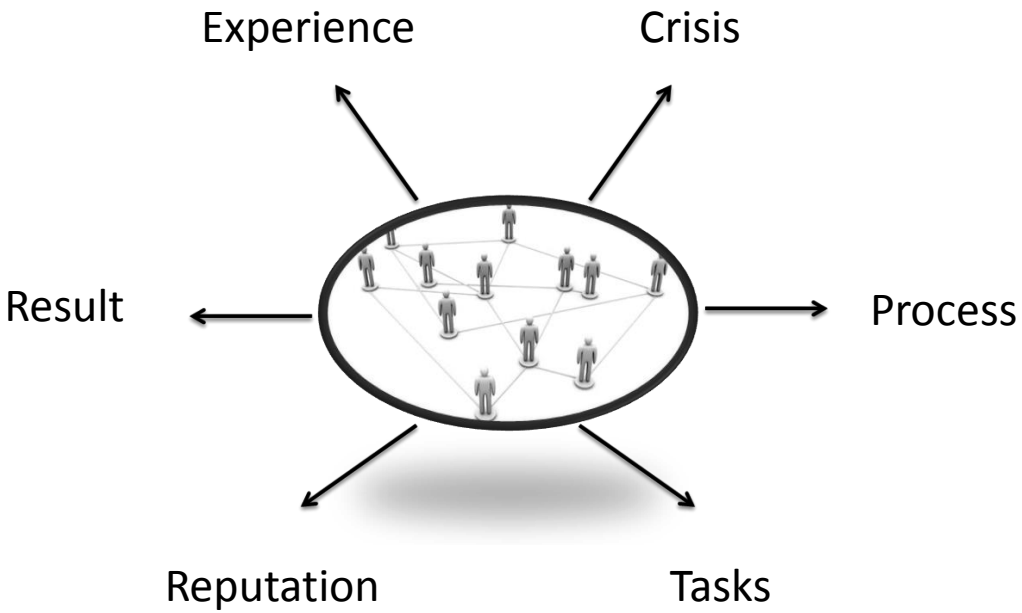


- Communication
- Information transfer
- Feelings
- Emotions
- Sentiments
- Statements
- ….

- **Analyses**
- **Validation**
- **Evaluation**

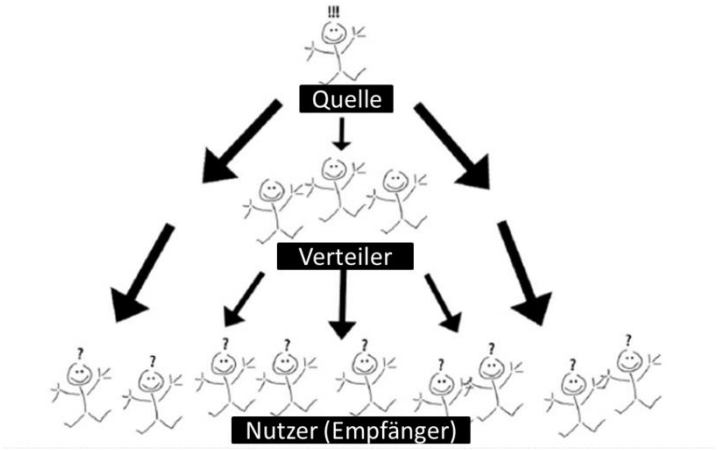
HIDF in groups, sub-groups

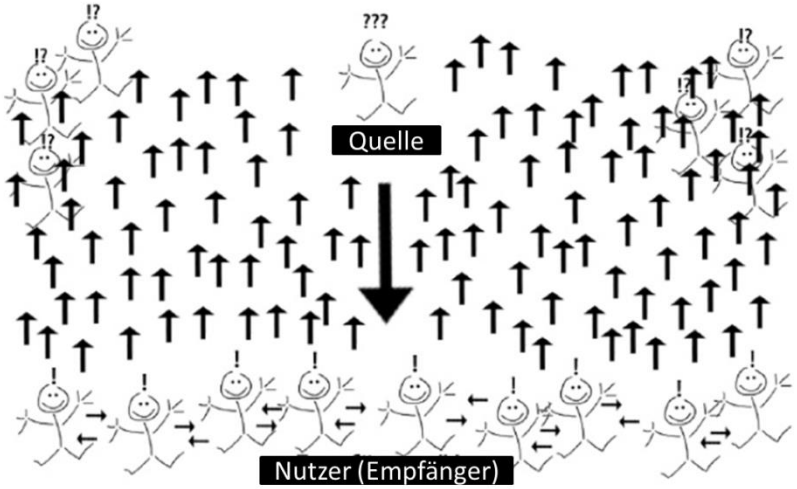
Groups in the net



Virtual groups, where the members mainly (but not exclusively) communicate via computer-aided communication.
In the context of the virtual group, the spatially dispersed members collaborate for a certain amount of time to complete defined tasks.

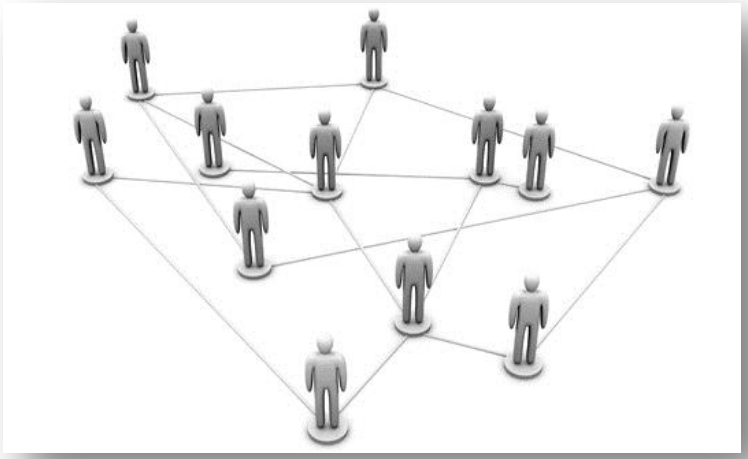
Communication





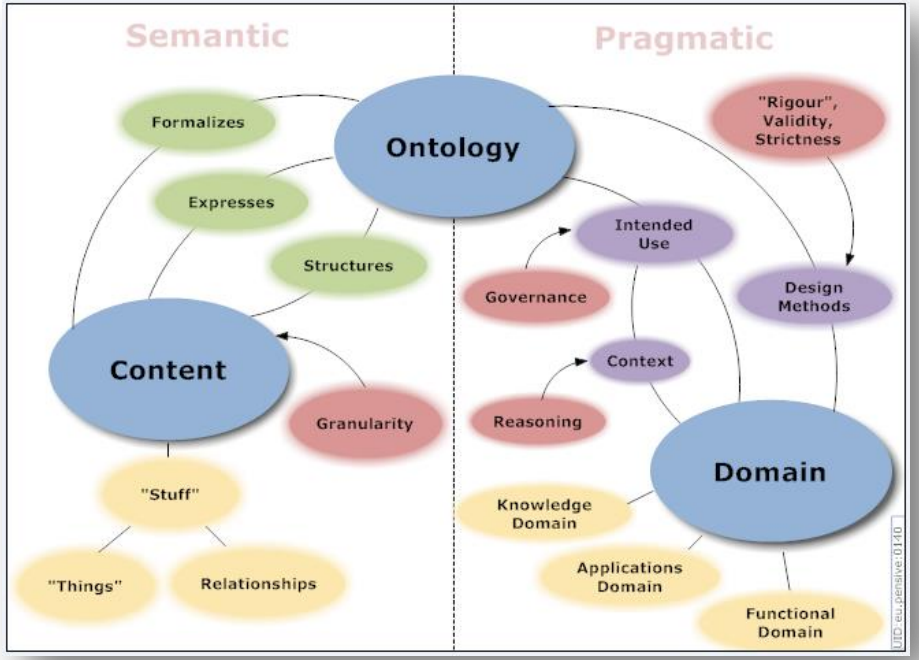
Computer-mediated communication, which is proceeds via computer networks. In the narrower sense computer-aided communication is understood to be only text communication, as it is typical for exchanging via e-mail, mailing list, newsgroup or chat, for example, and is theoretically modeled using various models.

Communication of Criminal Groups on the Internet

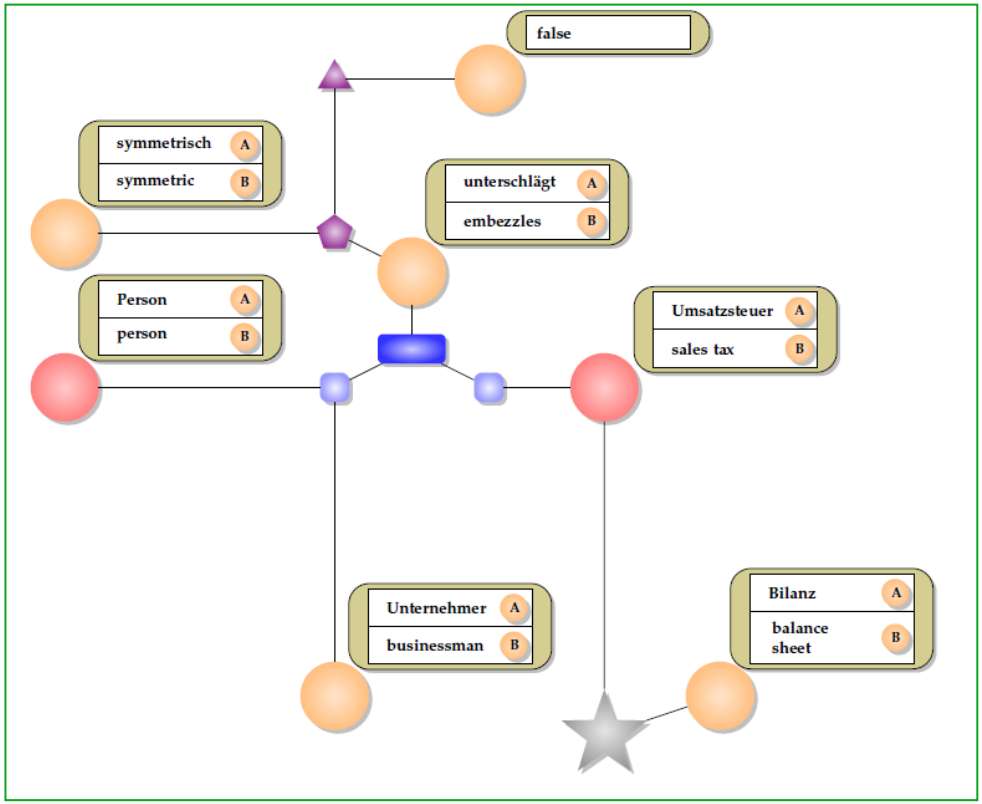


An ontology is a specification of a conceptualization.

Definition of syntax of terms and symbols in a network of associations



Ontologies and Semantics

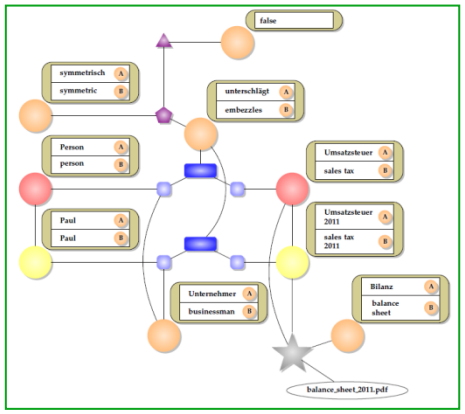


- they define terms and symbols referring to a syntax and an association network
- can support the information extraction process in different ways
- can support the visualization of results

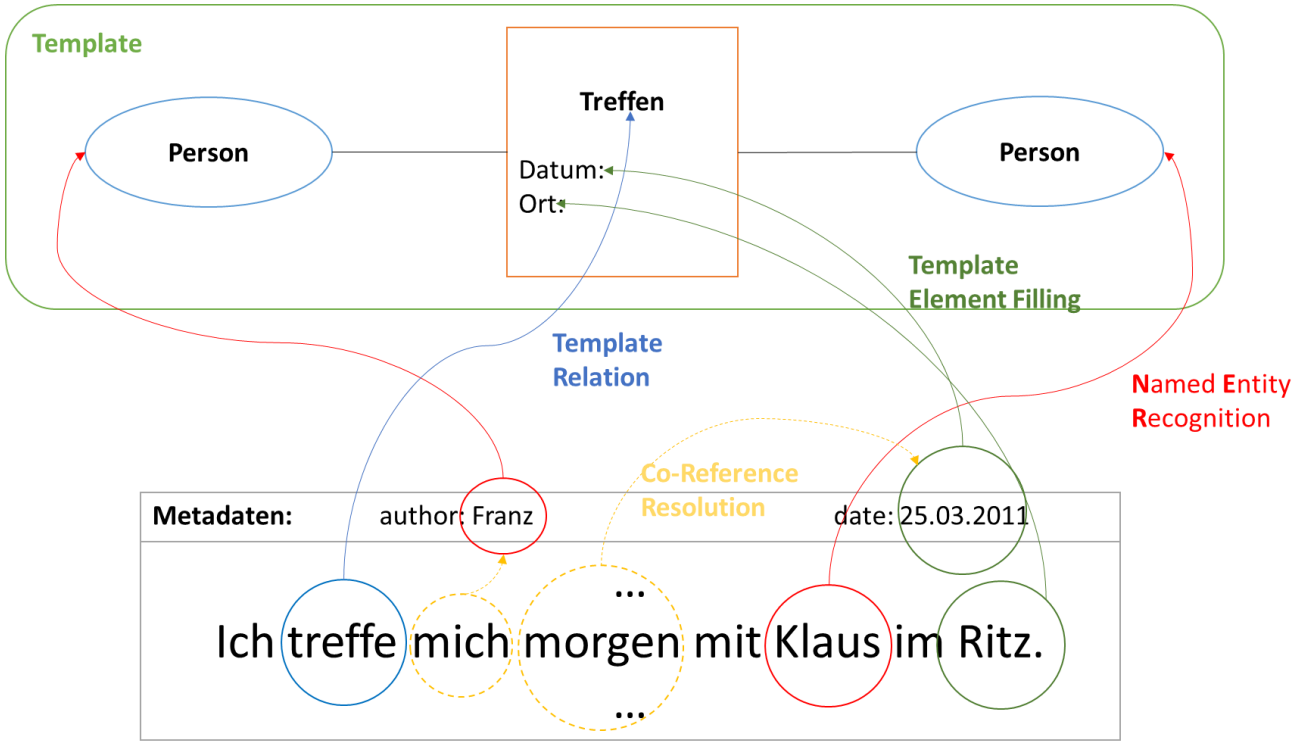
- oriented in the Topic Map ISO-standard
- readable for humans and processable for computers

Ontologies and Semantics

Forensische Topic Map
instantiation



Information Extraction



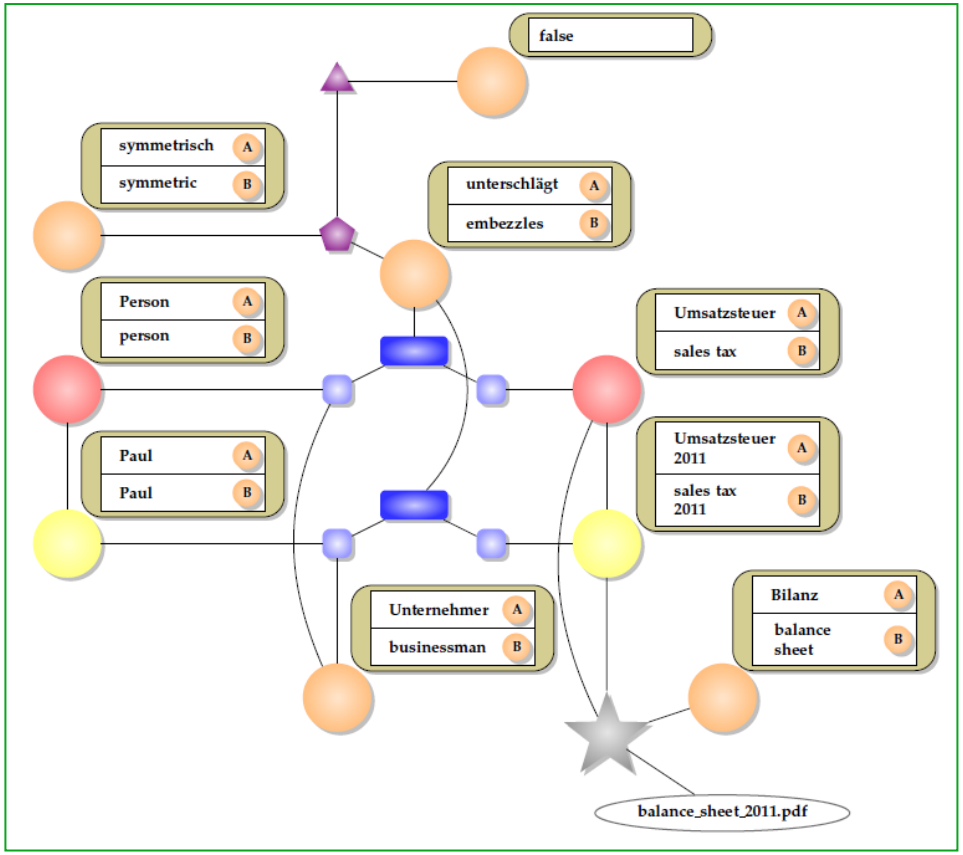
Ontologies and Semantics

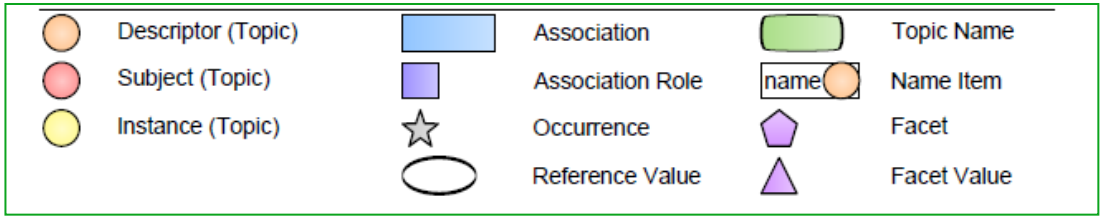
Forensische Topic Map

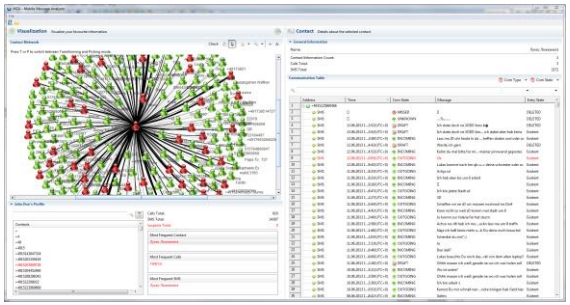
instantiation

↓

Question Answering System

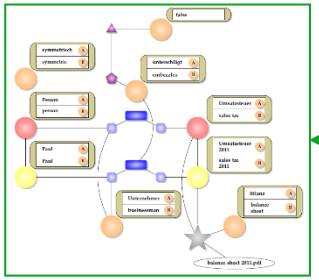


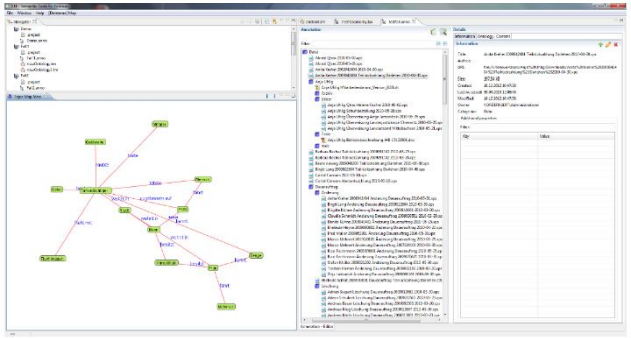




Mona

Mobile Message ANALYZER



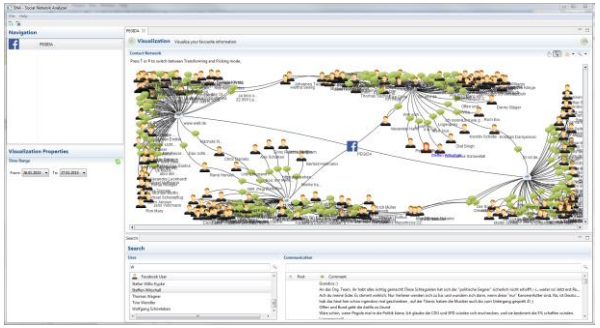


SemanTA

Semantic Text ANALYZER

AVATAR

A Victim Analysis Toolbox for Anatomic Reconstruction



SoNA

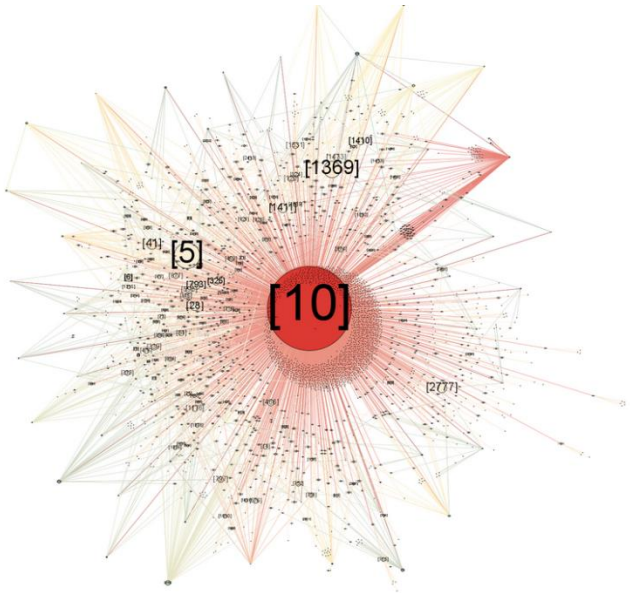
Social Network ANALYZER

Measurement of communication

Thursday, July 26

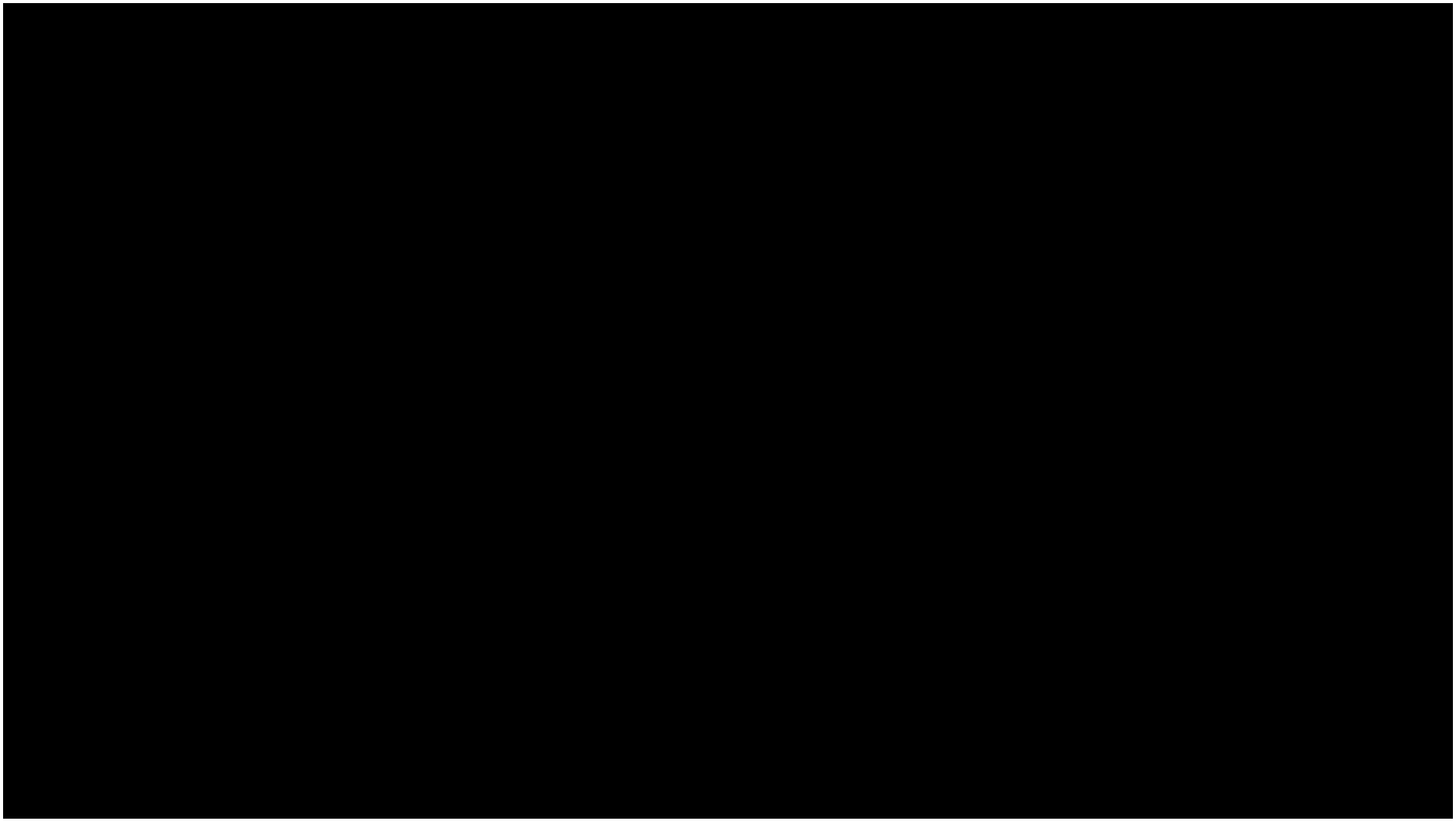
IMMM 3: Information Mining and Management II
Session chair: Ayodeji Oyewale

Opinion Leaders in Star-Like Social Networks: A Simple Case?
Michael Spranger, Florian Heinke, Hanna Siewerts, Joshua Hampl, Dirk Labudde



- Influence
- Opinions
- communication

Predictive policing



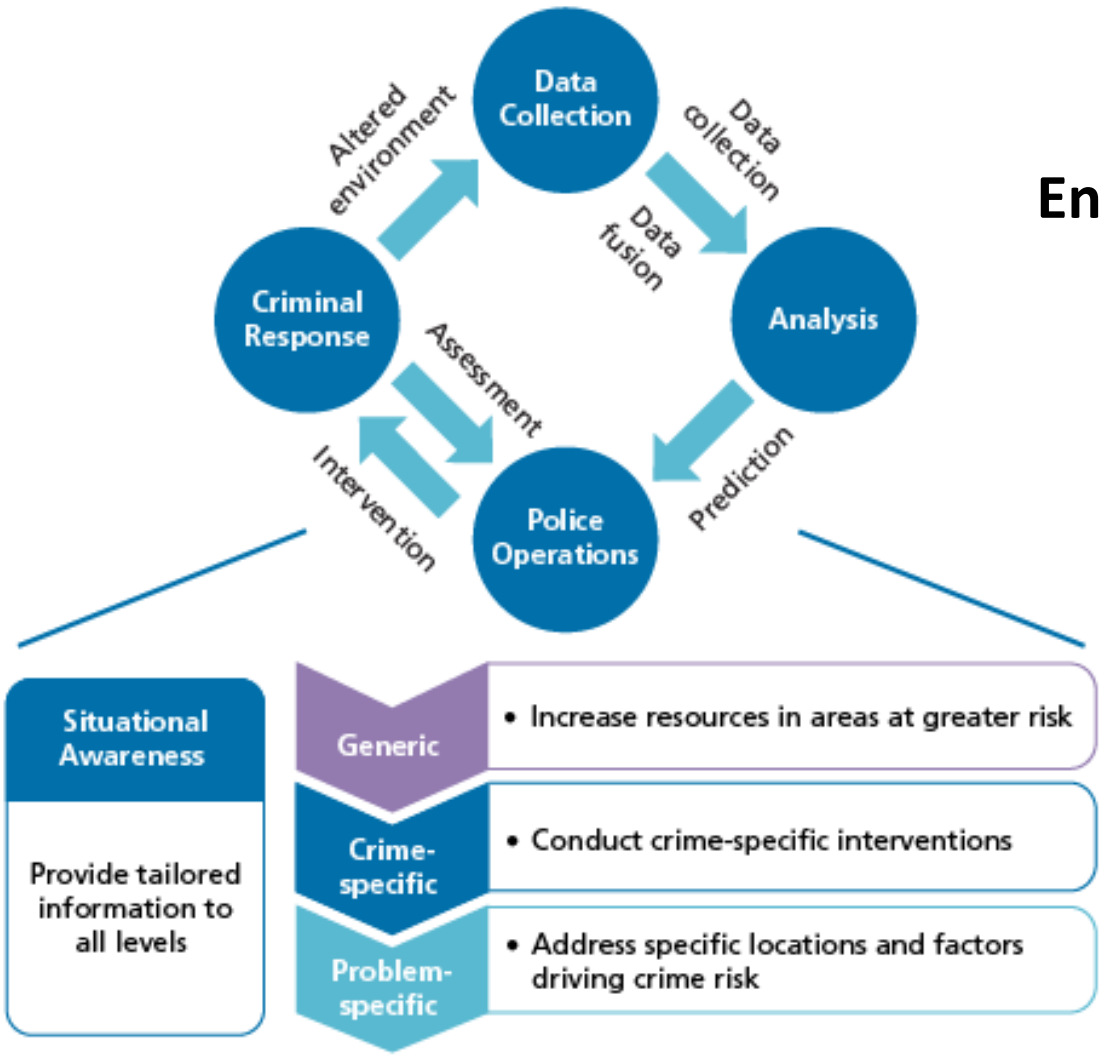
Advertising … but

- temporal aspect
- spatial aspect
- Aspect of Data collection
  - Aspect of statistical prediction
  - Aspect of the intervention

Information about future crimes

Predictive policing is a “multi-disciplinary, law enforcement-based strategy that brings together advanced technologies, criminological theory, predictive analysis, and tactical operations that ultimately lead to results and outcomes – crime reduction, management efficiency, and safer communities.” (Uchida 2014)

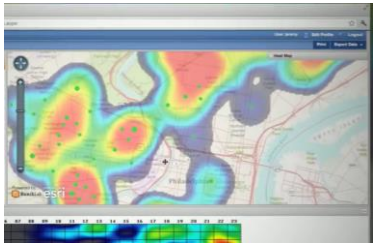
Predictive policing



End-to-End-Process

Mathematical models and assumptions

Mathematical models and assumptions

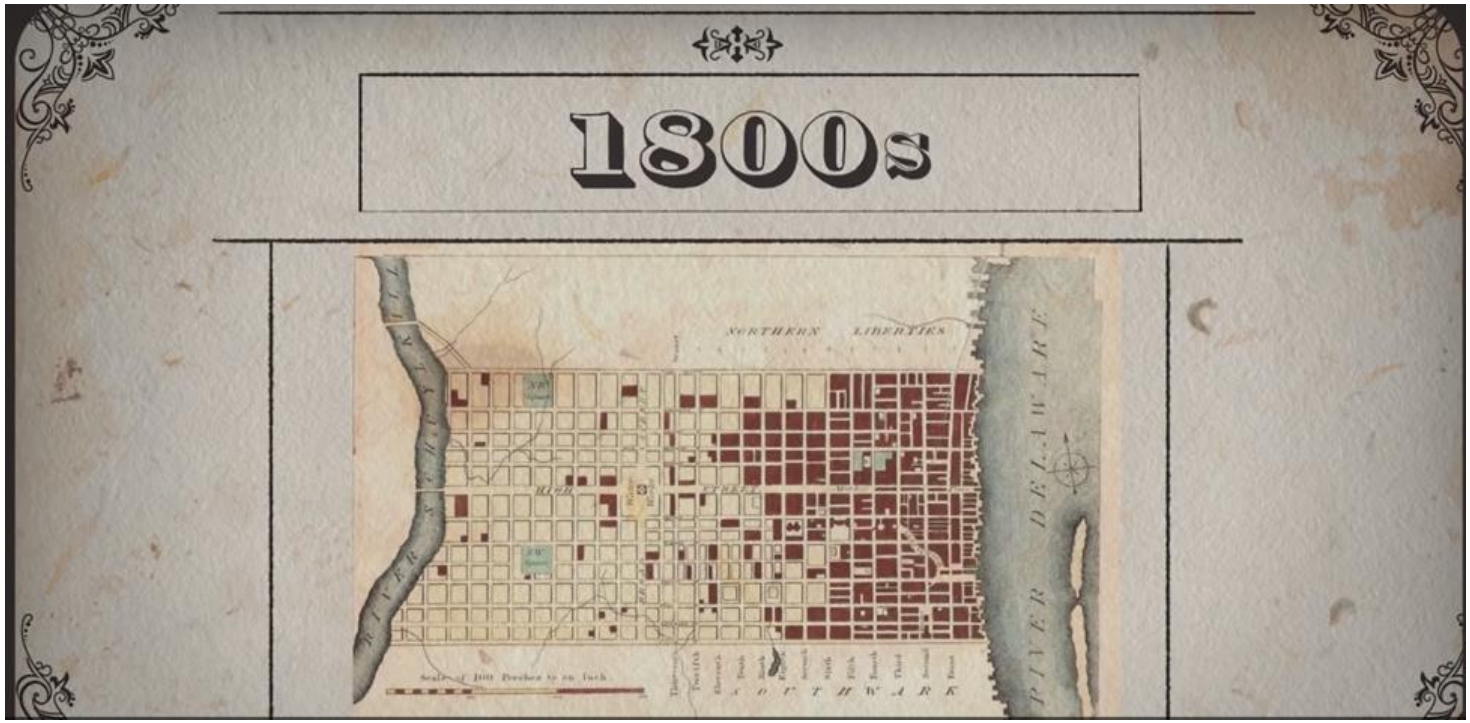


- Approaches from epidemiology (diffusion)
- Crime Sociology "Repeat Victimization“
- „Routine-Activity-Theory“
- Rational Choice
- Lifestyle Approach
- Broken Windows-Theory

Geographic Information Systems
+
space-based algorithms and modeling
=
Probabilities of crime

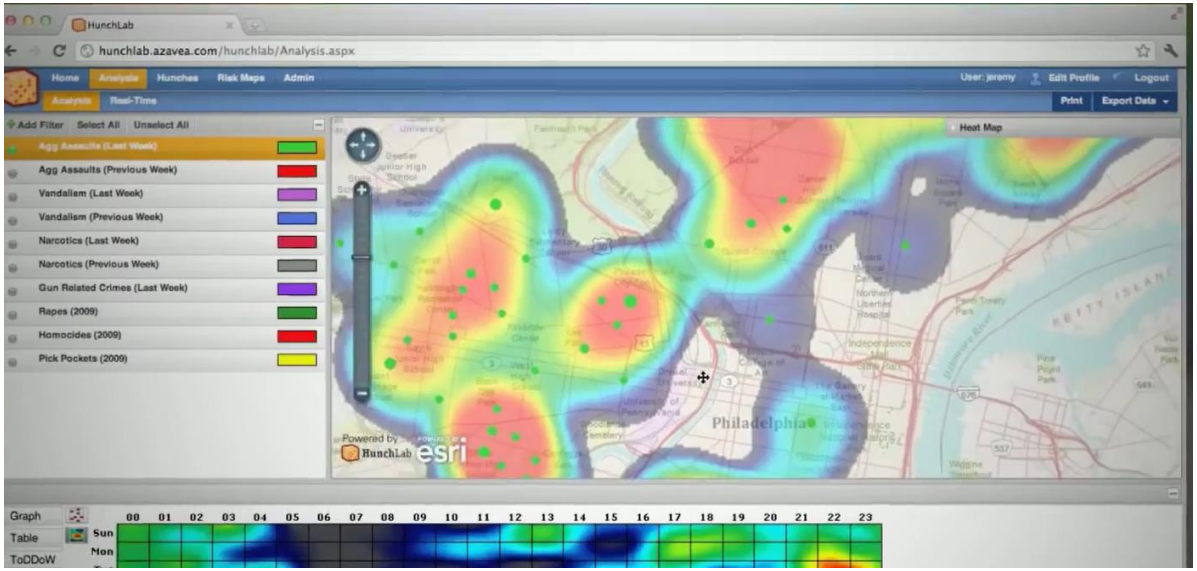
→ Search for patterns

BigData – Predictive Policing



First „statistic of criminal activity“ – people mark crimes on a map.
Basic idea – divide in quadrants (grid)

BigData – Predictive Policing



Number of criminal offences and their „conditions“

All still awake?

Near-repeat pattern analysis

Question:

When a criminal event takes place, how does the risk develop in the surrounding area?

→ Repeated attacks at the same or nearby places

$$P(A|x,t) \rightarrow P(B|x+\Delta x, t+\Delta t)$$

What influence does an assault A at the place x at the time t have at a later point in time $t + \Delta t$ and a nearby place $x + \Delta x$.

Relations for time and space can be derived from statistics.

Near-repeat pattern analysis

Odds Ratios

	0 to 15 days	15 to 30 days	30 to 45 days	45 to 60 days	60 to 75 days
Same location	2.6803	0.9033	0.6894	0.4331	0.6208
0 to 121.92 meters	1.2824	1.1158	0.8223	0.9218	1.0476
121.92 to 243.84 meters	1.1851	0.9325	1.0807	0.9207	0.8254
243.84 to 365.76 meters	1.1203	0.9448	1.0811	0.9923	0.851
365.76 to 487.68 meters	1.065	1.0754	0.9932	0.9271	0.8171
487.68 to 609.6 meters	1.149	1.0396	0.9577	1.002	0.9482
609.6 to 731.52 meters	1.0259	1.0772	0.9504	0.9865	1.0612
731.52 to 853.44 meters	1.0309	1.0922	1.046	0.9247	1.0202
853.44 to 975.36 meters	1.0423	1.0727	1.0303	1.0683	0.9943
975.36 to 1,097.28 meters	1.0666	1.0307	1.0631	1.0236	0.9784
1,097.28 to 1,219.2 meters	1.0027	0.9971	0.9834	1.0447	1.04
> 1,219.2 meters	0.9943	0.997	0.9993	1.0004	1.0019

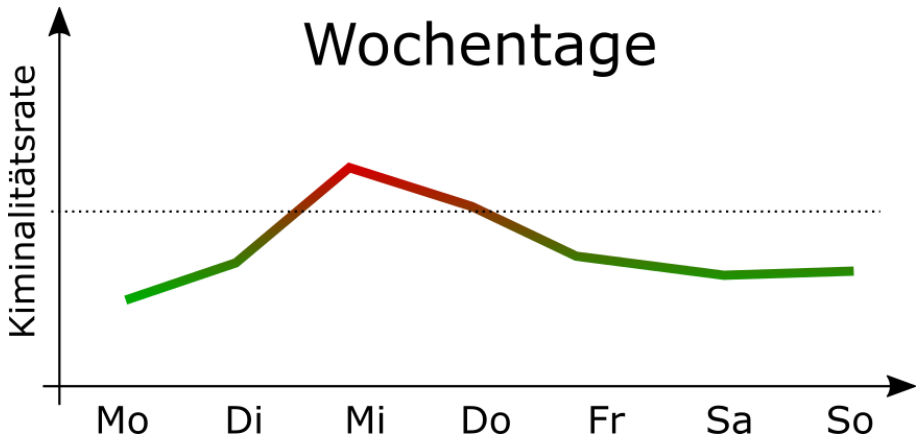
Near-Repeat Patterns in Philadelphia Shootings. Jerry H. Ratcliffe and George F. Rengert, Security Journal 2008.

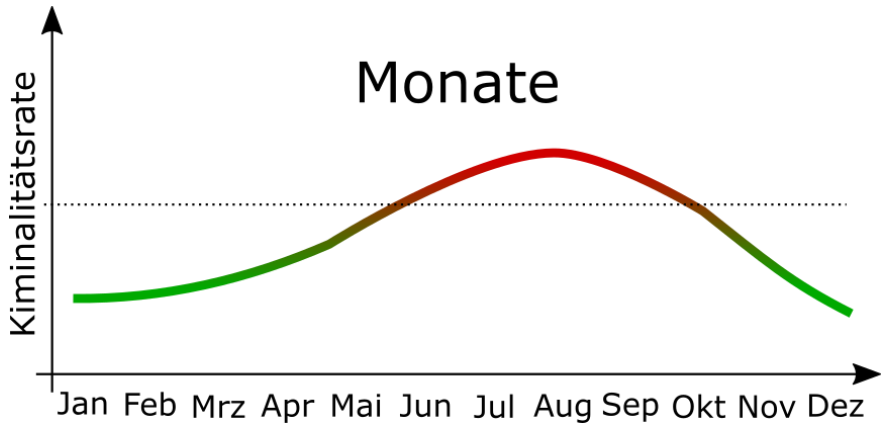
Cyclic-load forecasting

Question:

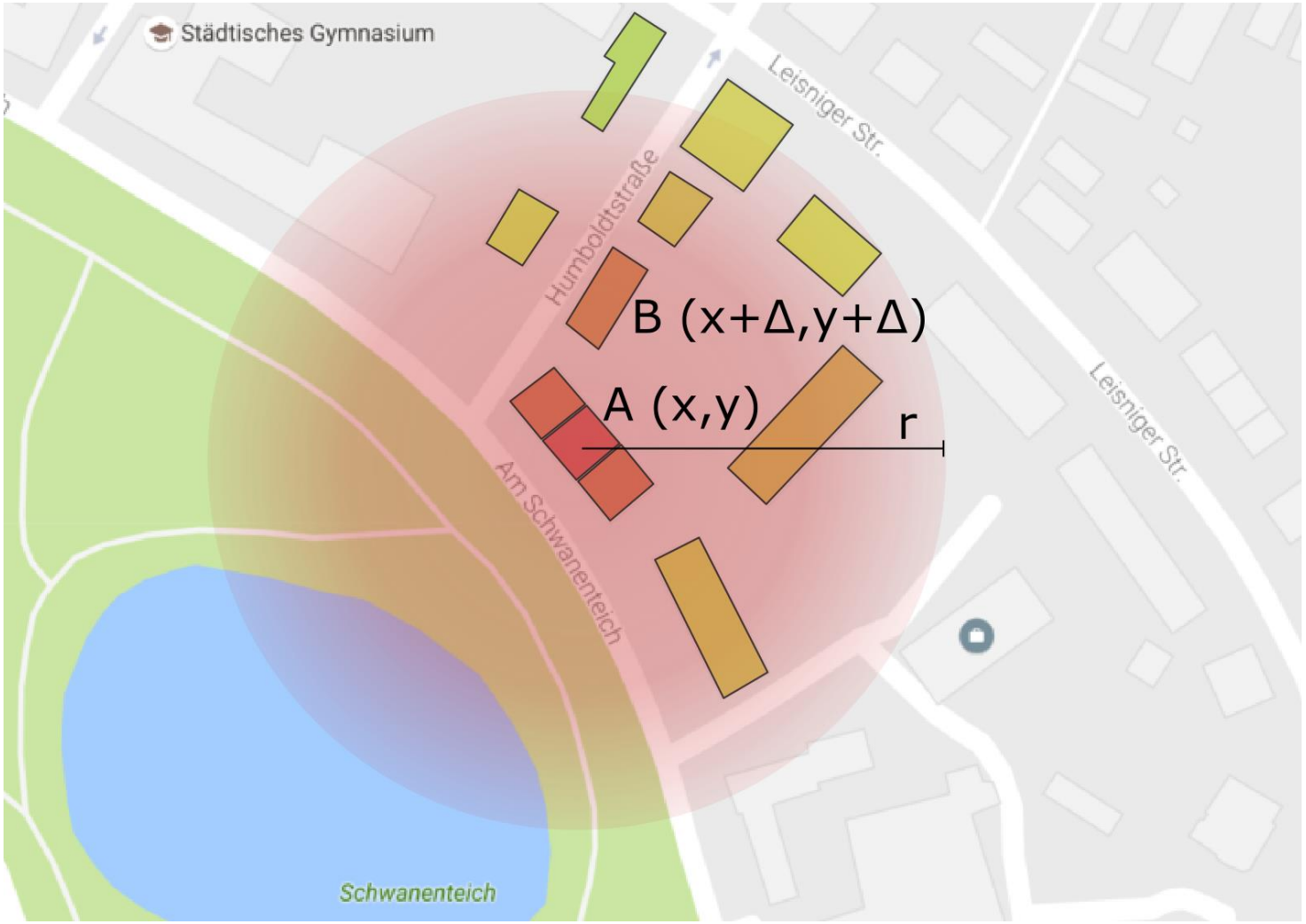
Are there certain days in the year, certain days of the week or certain times in the day in which more criminal incidents are to be expected?

$$\boldsymbol{P}(\boldsymbol{Event}|Point\ in\ time) = P(\boldsymbol{Event}|Month) + P(\boldsymbol{Event}|Day) + P(\boldsymbol{Event}|Day\ of\ the\ week) + P(\boldsymbol{Event}|Time\ of\ day) + \ldots$$

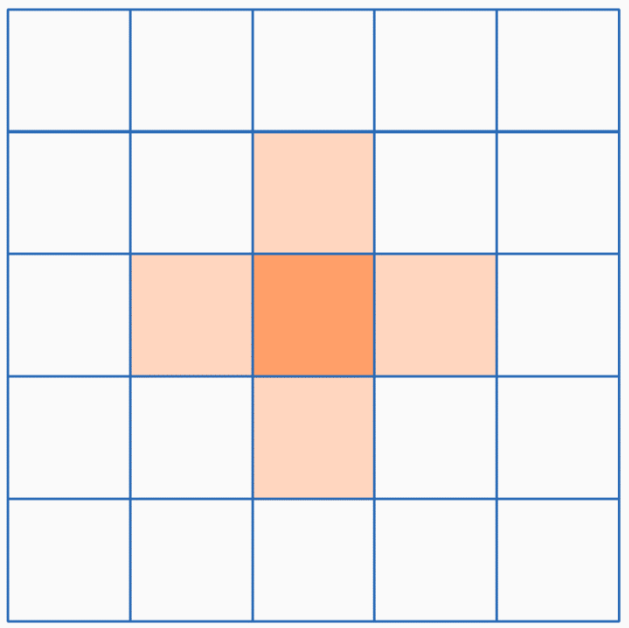




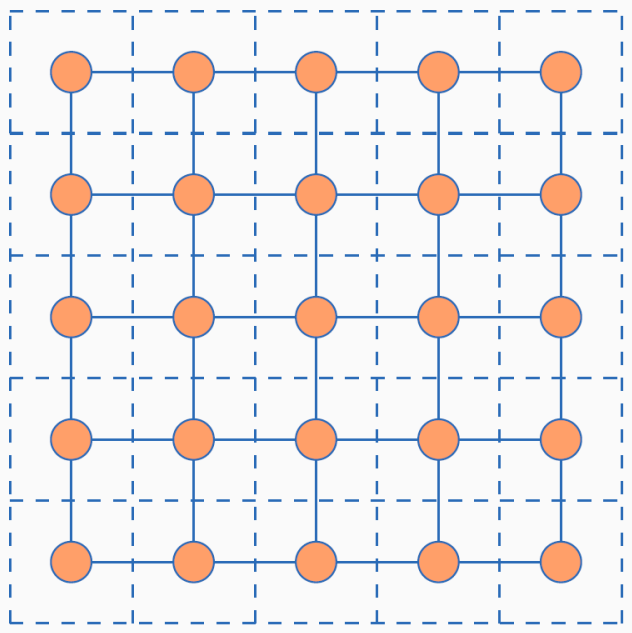
Result


Städtisches Gymnasium
Leisniger Str.
Humboldtstraße
B (x+Δ,y+Δ)
A (x,y)
r
Leisniger Str.
Am Schwanenteich
Schwanenteich

Automata



Definition of neighbors by means of a graph



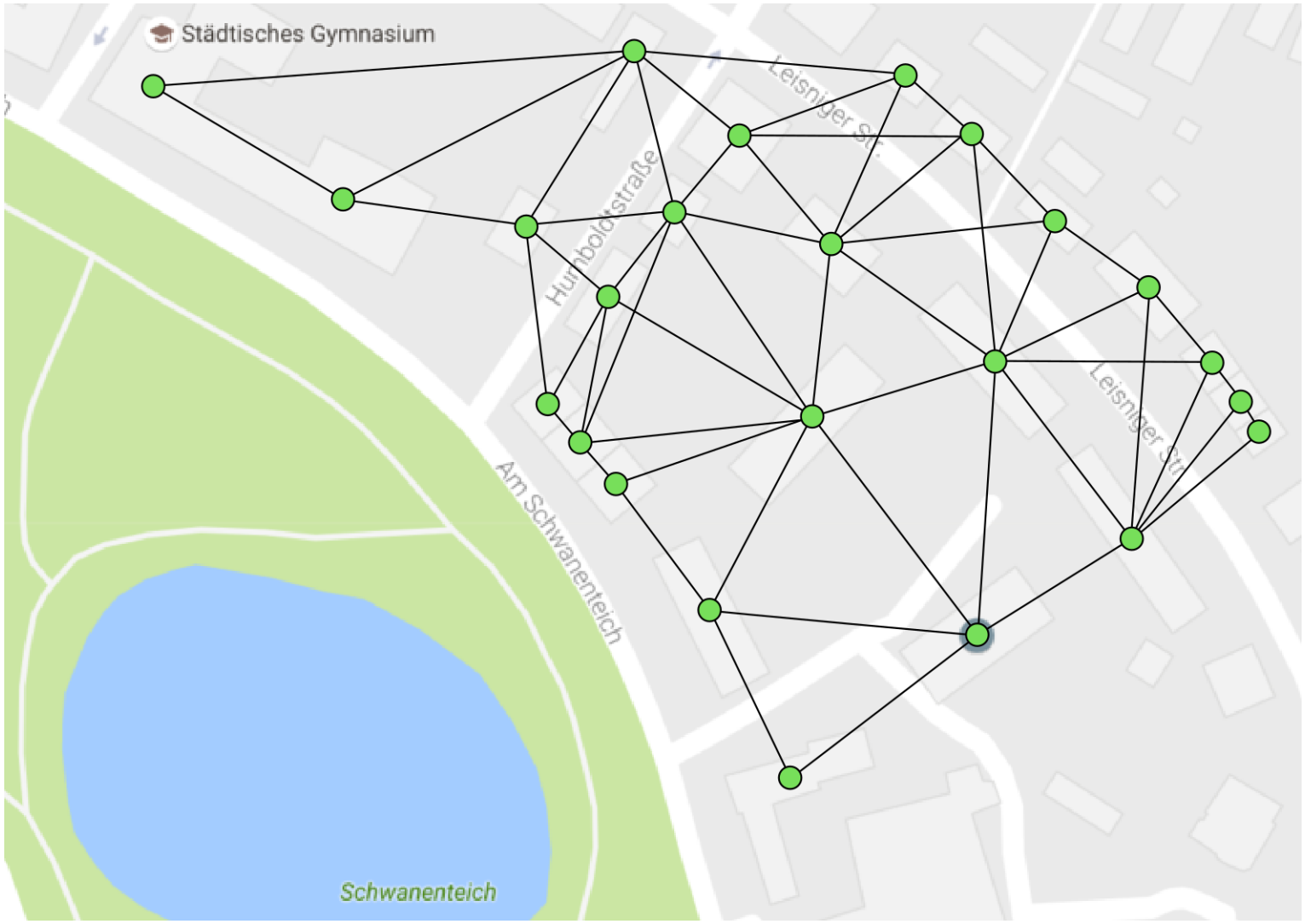
Cellular Automata

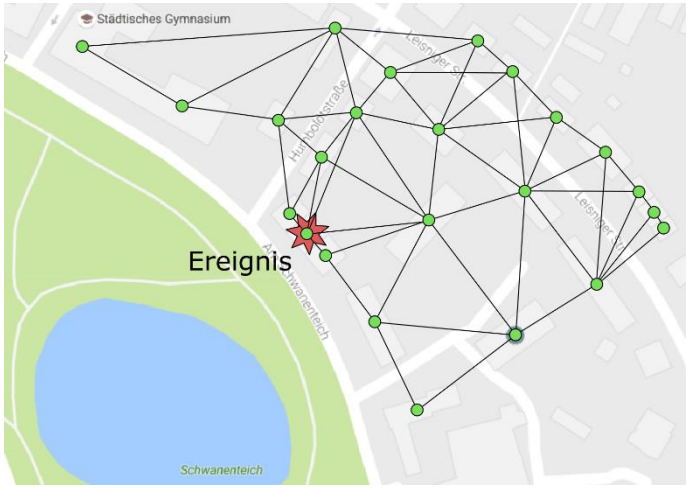
- Interaction with neighboring cells in a fixed grid
- Neighborly relations have to be determined at the beginning

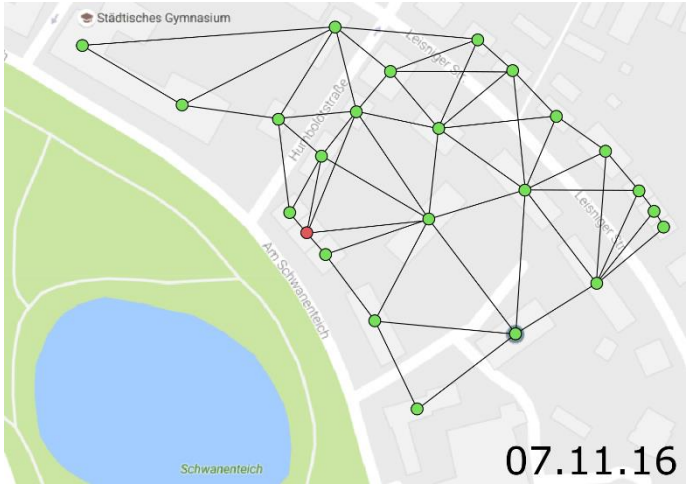
Graph Automata

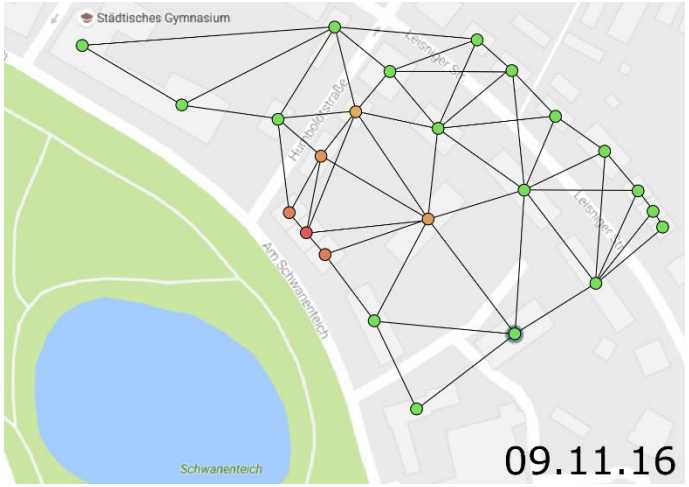
- Neighbor geometry can be flexibly modified
- Neighborly relations can change at any time

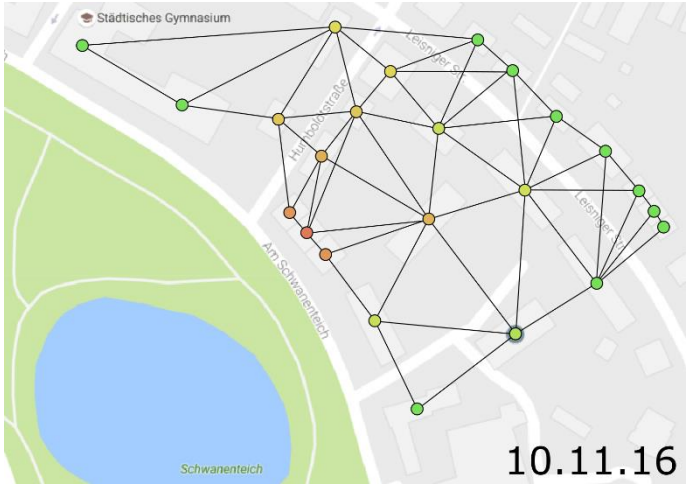
Graphen Automaten





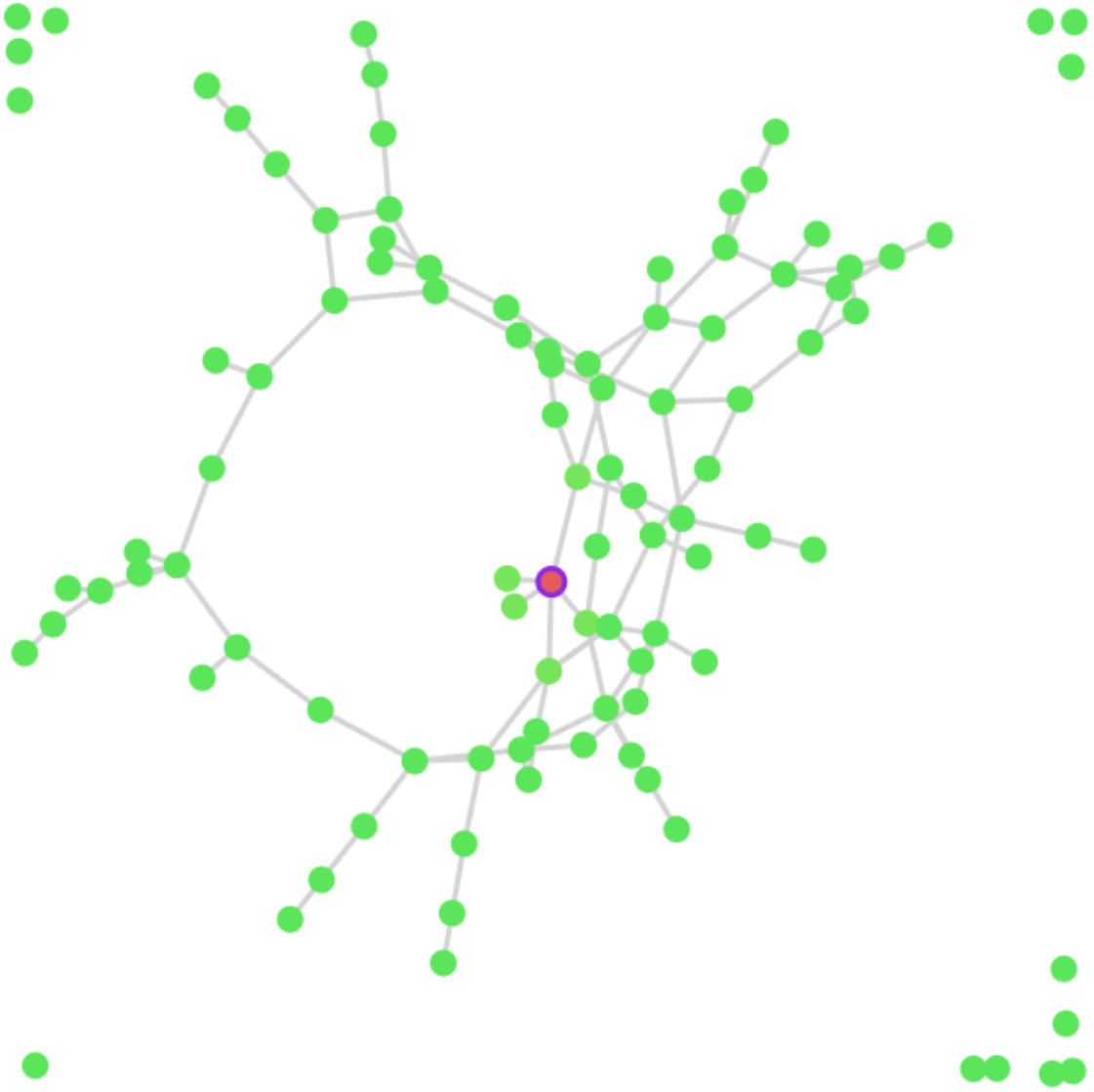




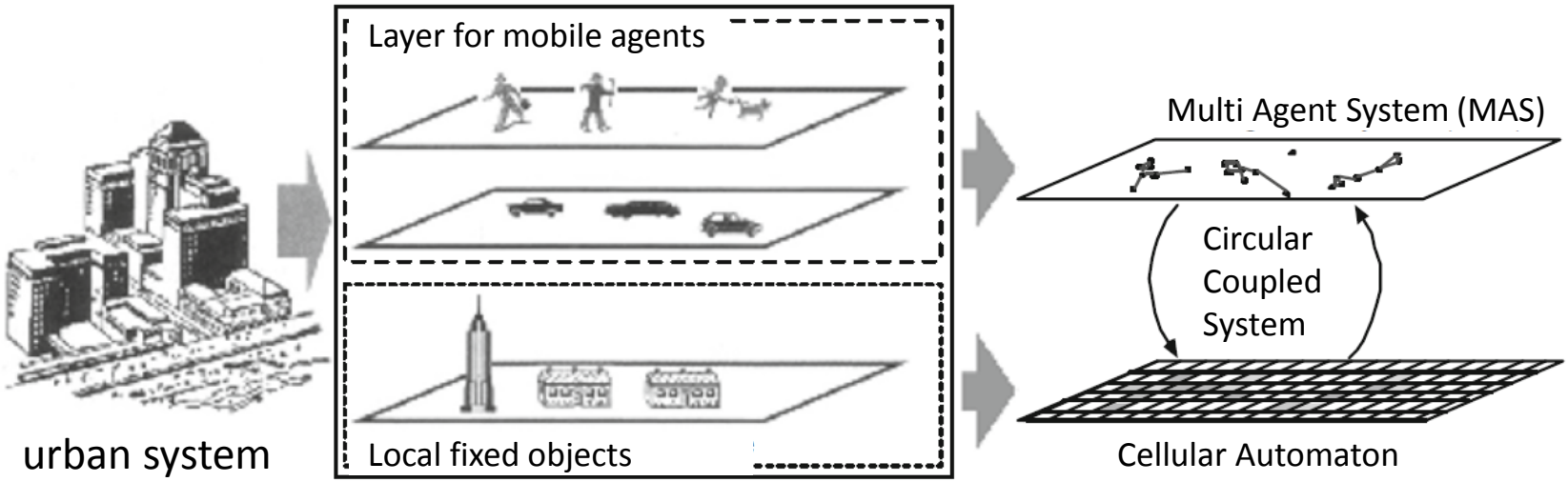


Securit of a house in an urban structure with a well defined crime rate

Simulation Urban Structure



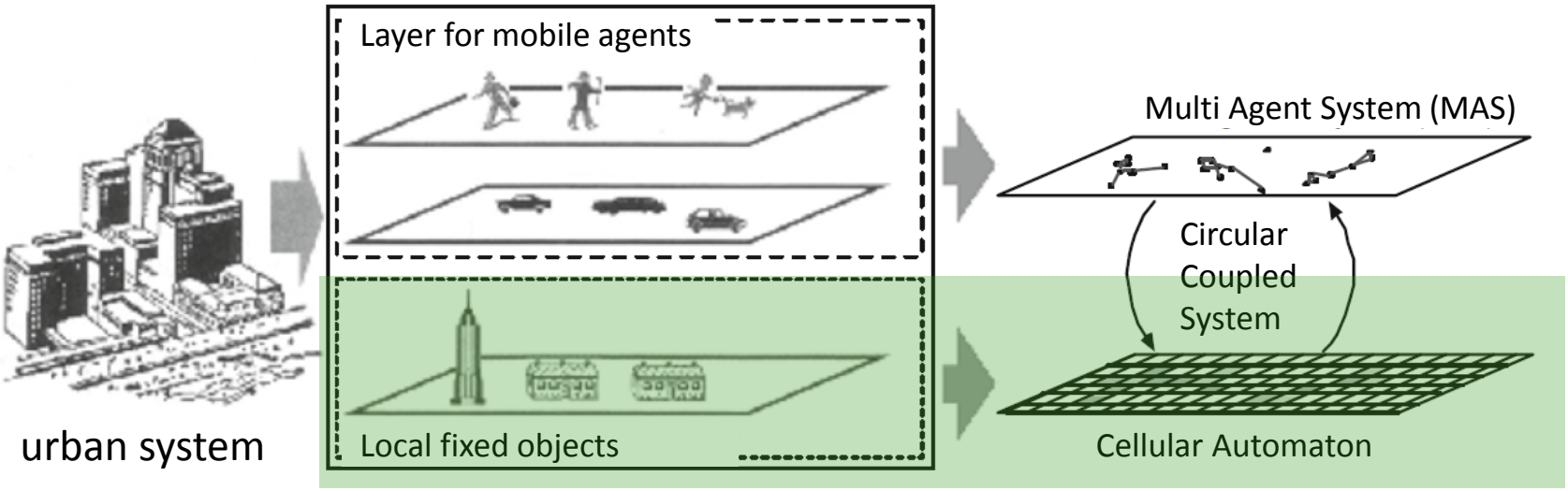
Multi-Agent-Systems



Cellular automata and multi-agent systems can be used for the simulation of the dynamic of **spatial processes**.

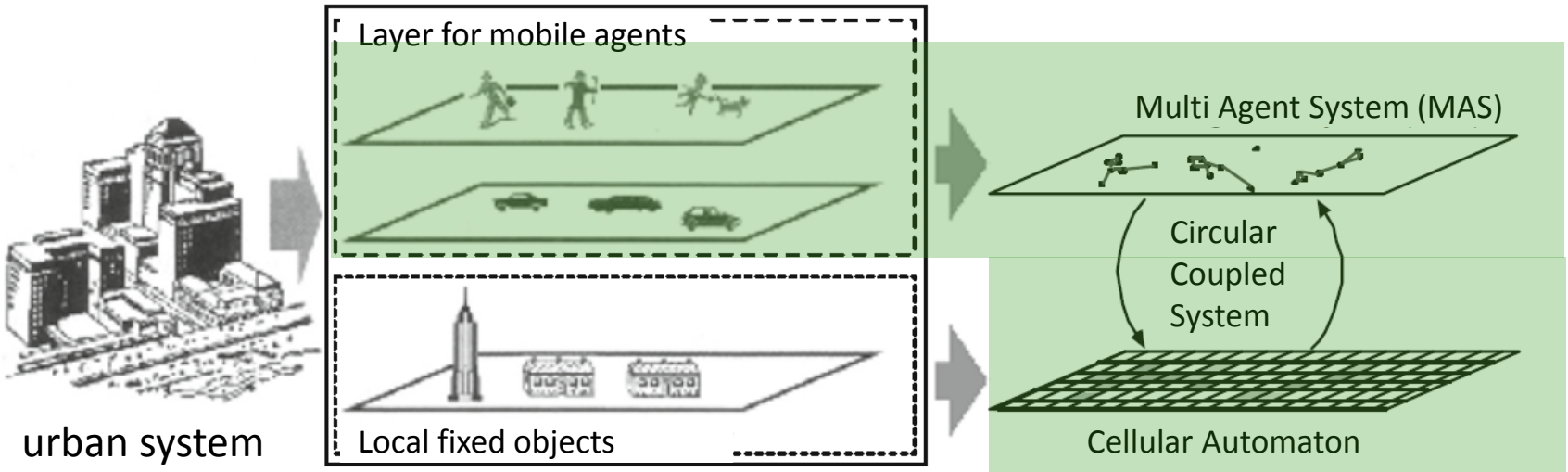
Cellular automata

Cellular automata model spatial elements such as streets, parcels of land, and buildings. Those are treated as locally fixed objects, whose state can change at a certain point in time. In a first step, these elements are transferred to a cell of a regular grid and saved as the status of this cell.

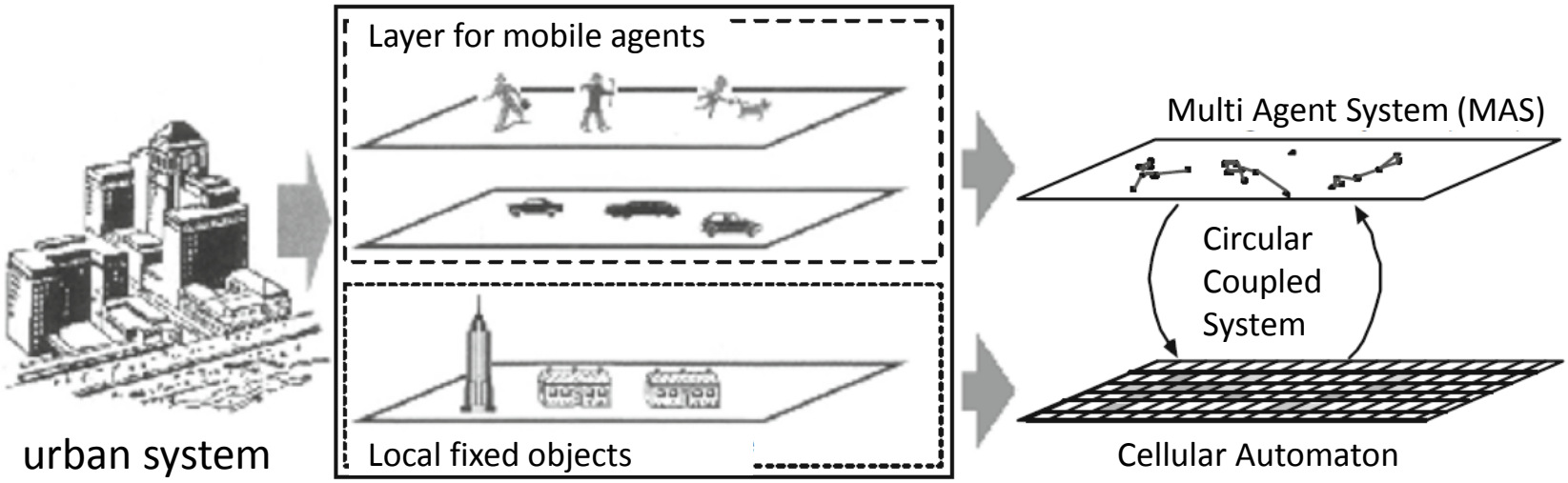


Multi-Agent-Systems (MAS)

A second level includes the individual and collective urban actors, which will be called agents. In comparison to the cells, agents are mobile and can move freely through the cell grid, the cellular room. It is possible to define different types of communication between the agents and between the agents and the cell.



FACS (Free Agents in a Cellular Space)



Model of an urban system. The system is separated in different levels, which are represented either by the cells of a CA or the agents of the MAS.

Communication as additional input for predictive policing

Social networks -
Sentiment analyses of groups in social networks

Generation of landscapes

Social networks

Sentiment Detection:

- sub-area of text mining
- automatic analysis of texts with the aim of identifying an **expressed attitude** as positive or negative

trend analysis

Statistical analysis starts from a basic set of concepts (or n-grams), which is associated positive or negative sentiments.

- Occurrences of positive or negative sentiments
- Comparison
- Designation of trends, opinions Statements

Inside the Web --- feelings, beliefs and personal opinions!

Generation of person landscapes from Facebook

PEGIDA – Patriotische Europäer gegen die Islamisierung des Abendlandes
(*Patriotic Europeans Against the Islamisation of the West*)

"Slogan"

- Asylindustrie – alysum industry
- Lügenpresse – lies press
- Volksverräter – public traitor
- Der Untergang des Abendlandes
- „the decline of the West"
- Islamisierung … Islamic

Social networks

Generation of landscapes from Facebook

https://www.facebook.com/pages/PEGIDA/790669100971515

likes

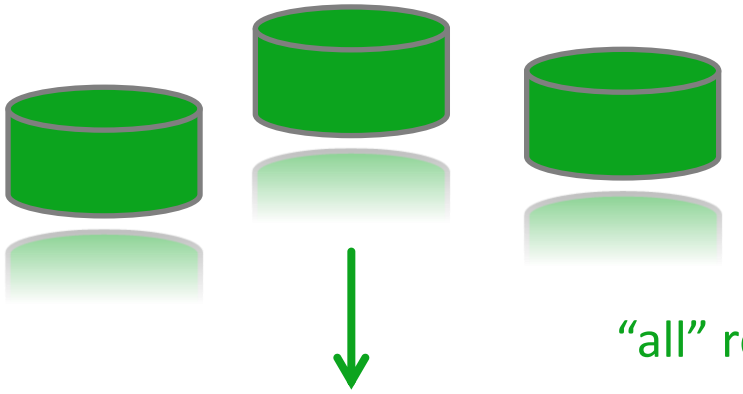
Comments

Likes, Comments and Returns

Social networks

Definition of a general **Data Model** – based on the Graph API

Information from FB via API –> Data model

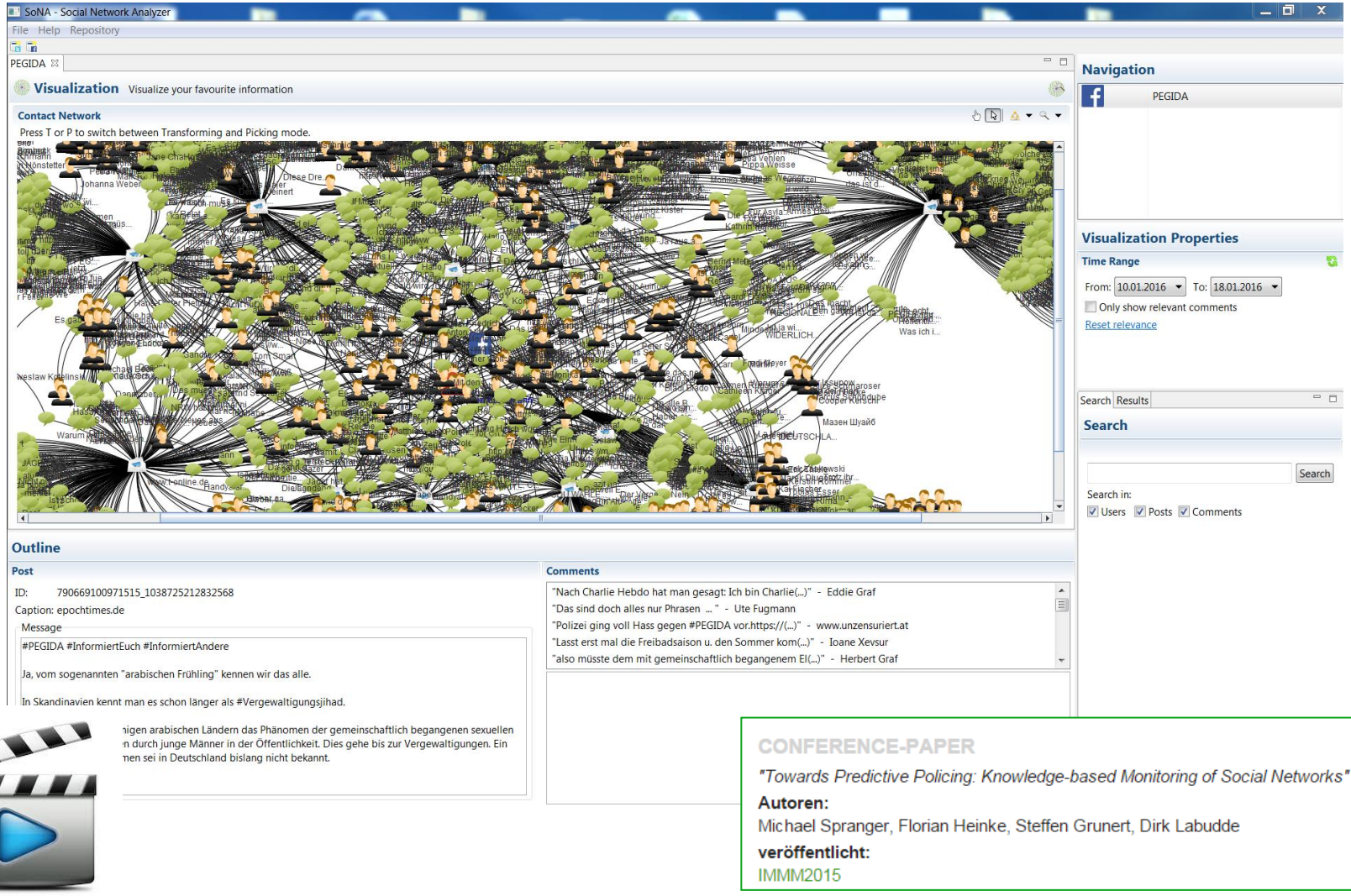


“all” relevant Data (Subsets)

Search relations and correlations

and sentence

Social networks



CONFERENCE-PAPER

"Towards Predictive Policing: Knowledge-based Monitoring of Social Networks"

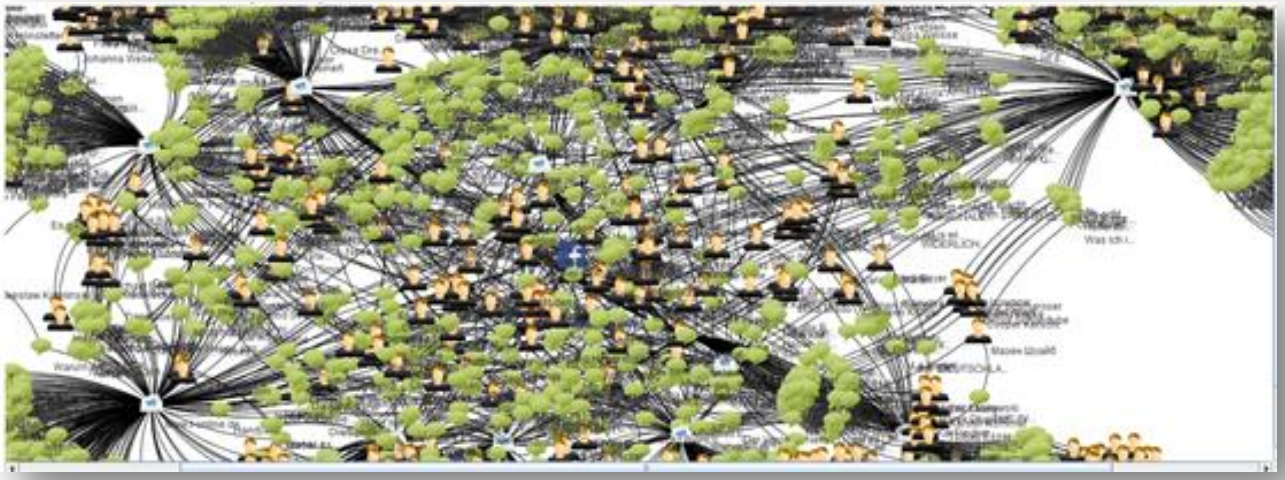
Autoren:

Michael Spranger, Florian Heinke, Steffen Grunert, Dirk Labudde

veröffentlicht:

IMMM2015

Social networks – Detection of Information



All information are labeled by time and user!

Analyze der opinion in der Group

a. positive
b. negative
c. neutral

Sentiment Analyze

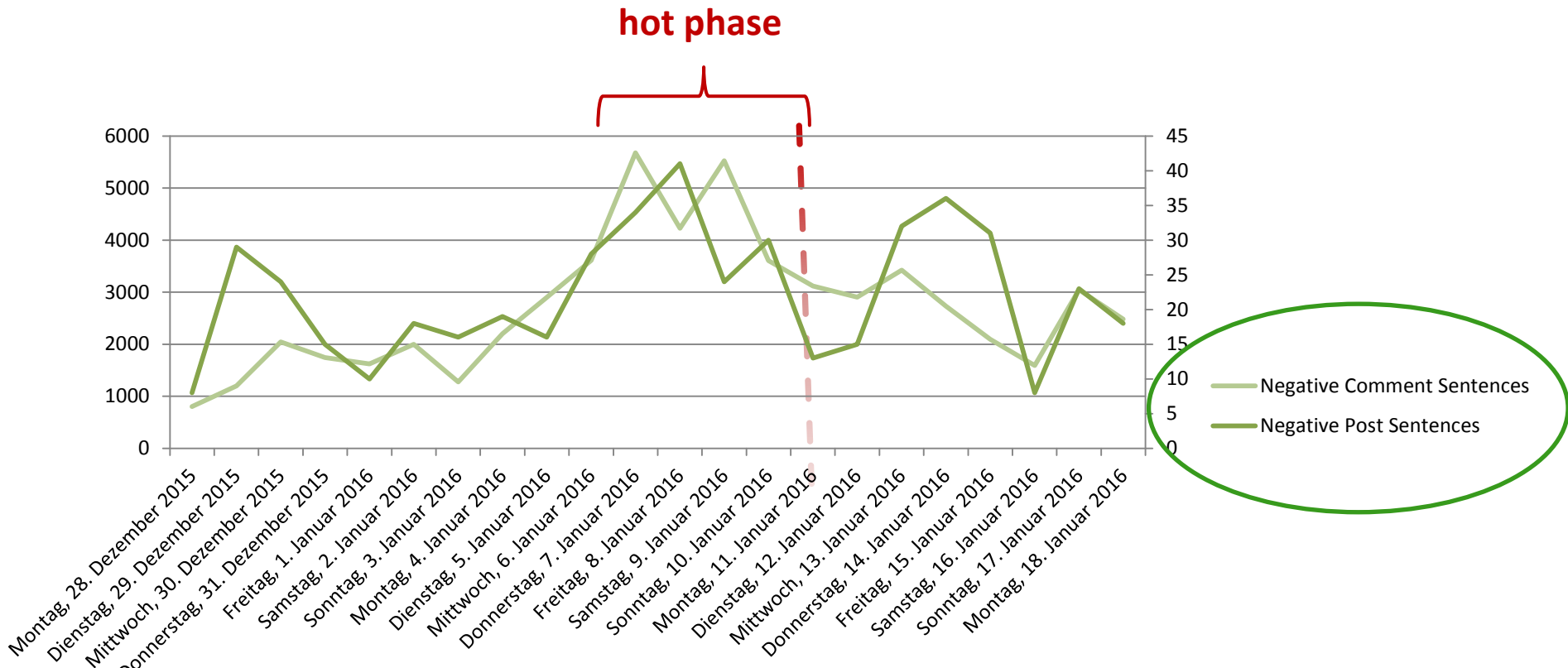
→ Time and geographical Monitoring

Social networks – Detection of Information

- User activities
- Topics
- Subgroups

- User-specific content
- Sentiments of the network

Opinions and moods (Sentences)



Digitized traces - geographical Monitoring

Digital traces

Digitized traces - geographical Monitoring

Foto BigData

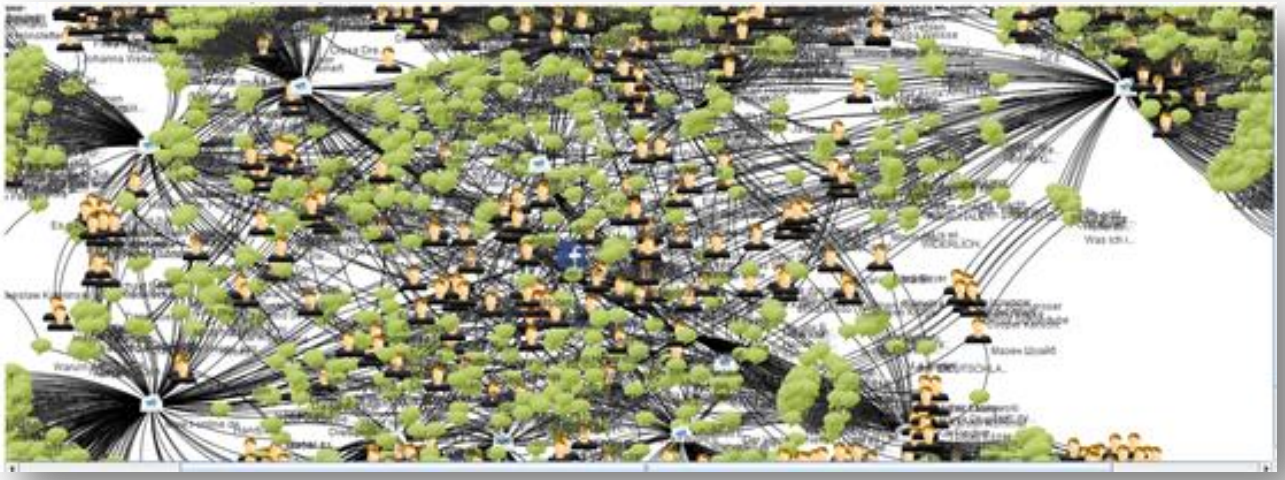
Video BigData

Social networks – Detection of Information

Content

Meta-data

User-list



All information are labeled by time, locations and user!

Sentiment Analyze

+

Posible geo-information

Digital traces

Digital traces

Digital traces

Digital traces

Digital traces

Digitized traces

Digitized traces

Digitized traces

Digitized traces

Digitized traces

Prediction

BigData from communication – Predictive Policing

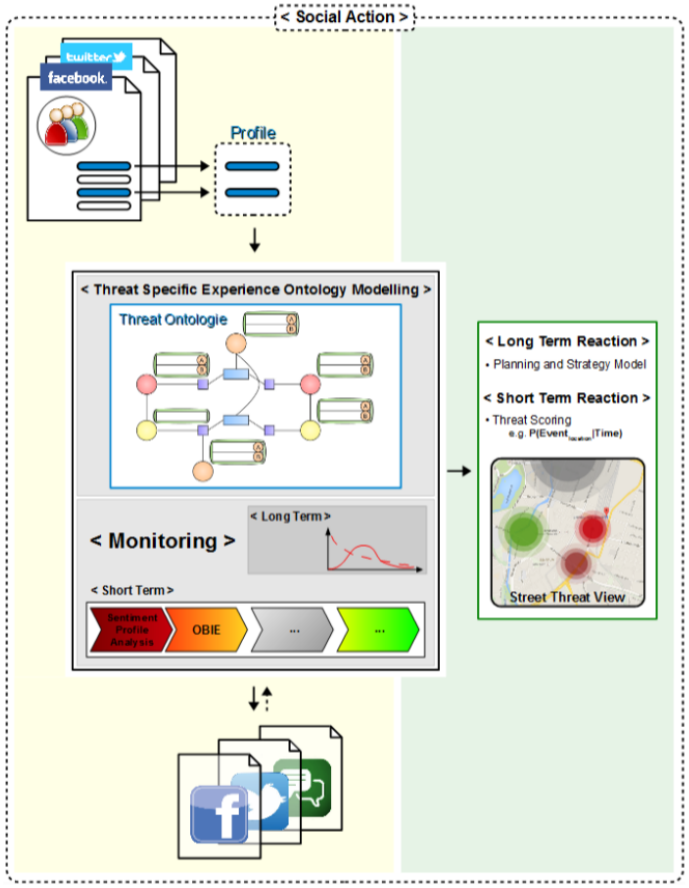
BigData – Predictive Policing

Towards Predictive Policing: Knowledge-based Monitoring of Social Networks

Michael Spranger, Florian Heinke, Steffen Grunert and Dirk Labudde
University of Applied Sciences Mittweida
Mittweida, Germany
Email: {*name.surname*}@hs-mittweida.de

***Abstract*—Increasing the resilience of the society against disorders, such as disasters, attacks or threatening groups, is one of the biggest challenges. Recent events highlight the importance of a resilient society and steps which are required to be taken in resilience engineering. *A priori* the optimal way to handle such adverse events is to prevent them, or at least provide appropriate courses of preparation. The essential requirement for every kind of preparation is information about relevant upcoming events. Such information can be gained for example from social networks and can form the basis for a long-term and short-term strategic planning by security forces. For that purpose, an application framework for knowledge-based monitoring of social networks is proposed, which is able to predict short-term activities as well as the long-term development of potentially dangerous groups, and discuss its basic concepts.**

***Index Terms*—forensic; text processing; resilience engineering**



BigData – Predictive Policing

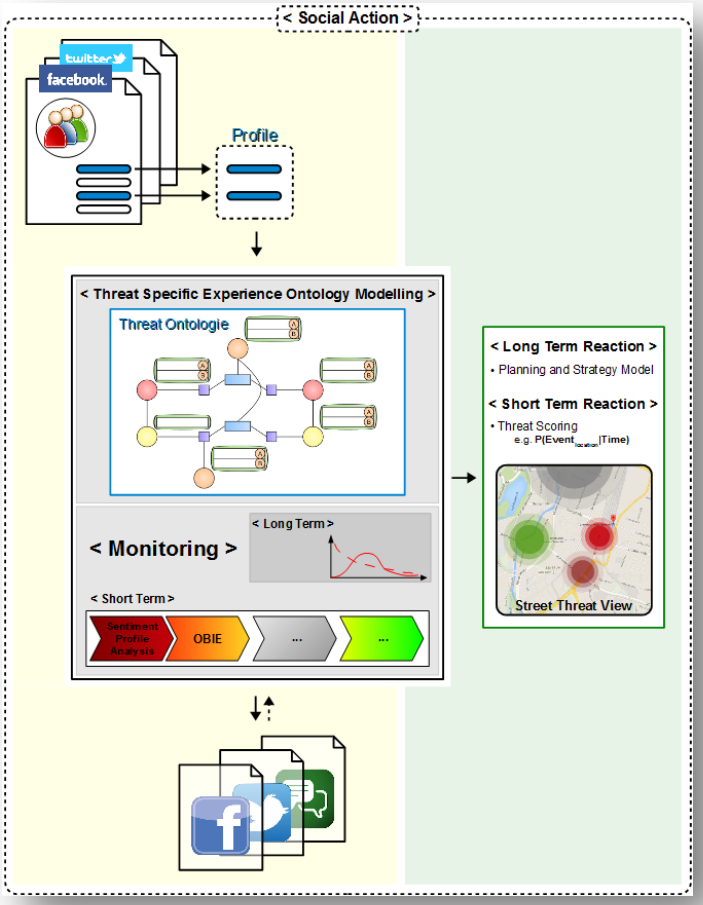
- Extraction of profiles for monitoring
- Extraction of post or comment content relating to the threat ontology and a sentiment analysis

→ enables short-term reaction

- Simulation of temporal development of groups and hot-spots

→enables long-term resource and strategic planning

- Increasing resilience



Opinions and moods (Sentences)

rights rioters in Leipzig-Connewitz

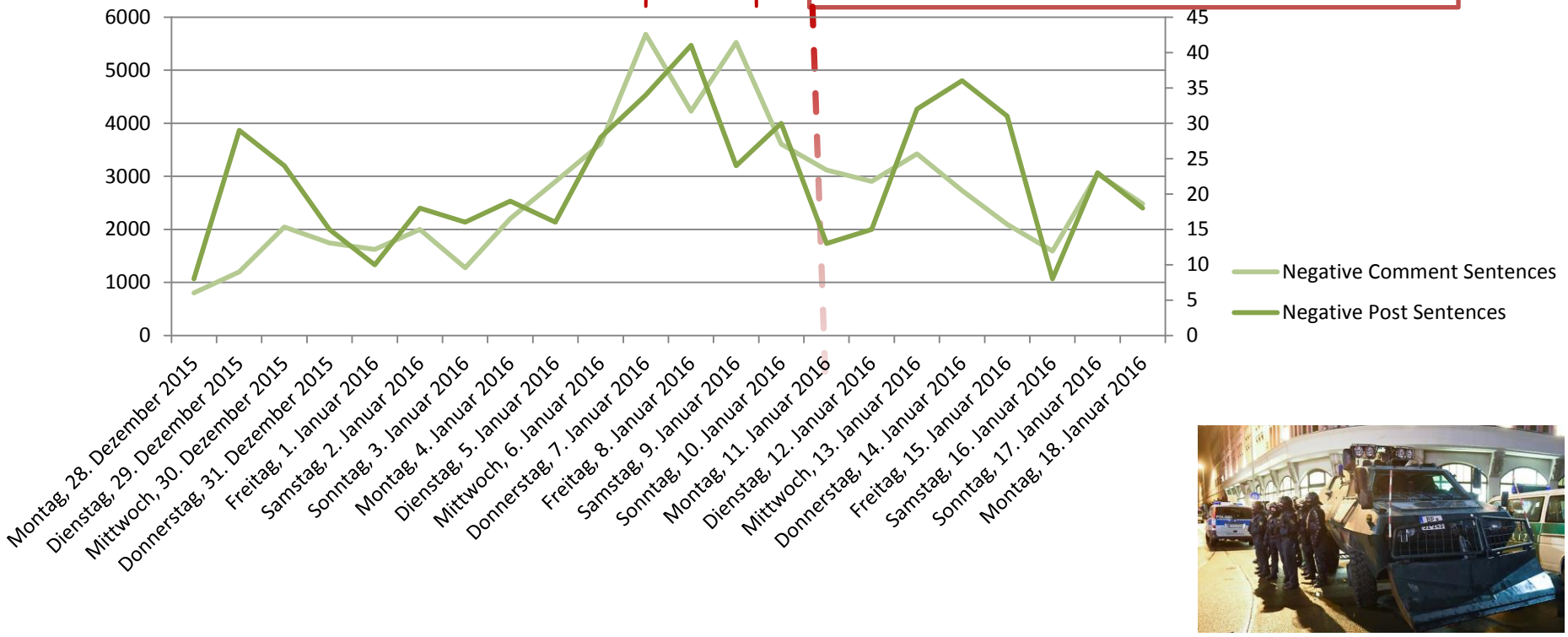
hot phase

Montag, 11. Januar 2016

Entsetzen in Leipzig

Rechte Randalierer verwüsten Connewitz

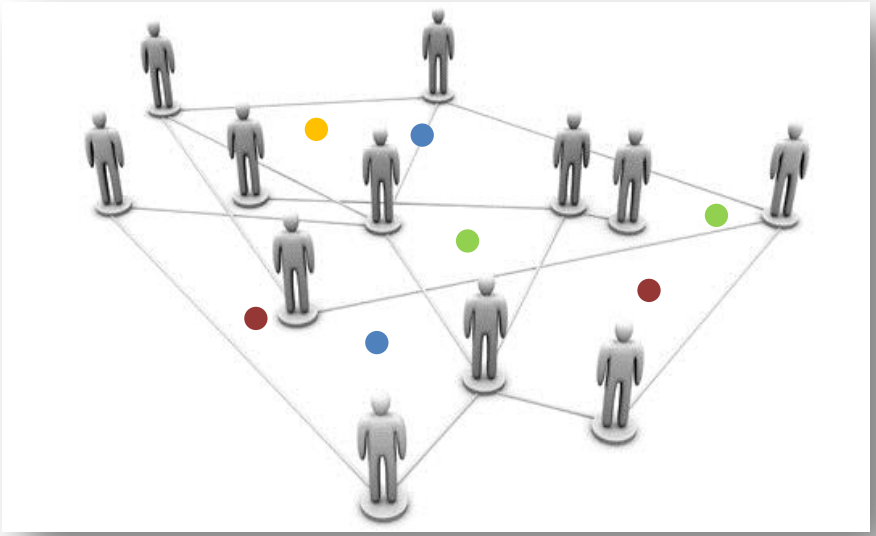
Während in der Leipziger Innenstadt das fremdenfeindliche Legida-Bündnis demonstriert, nutzen rund 250 Hooligans die Gelegenheit, in Connewitz einzufallen. Sie ziehen eine Schneise der Verwüstung durch das linksalternative Viertel, bevor die Polizei die Lage unter Kontrolle bekommt.



Social networks – Predictive Policing

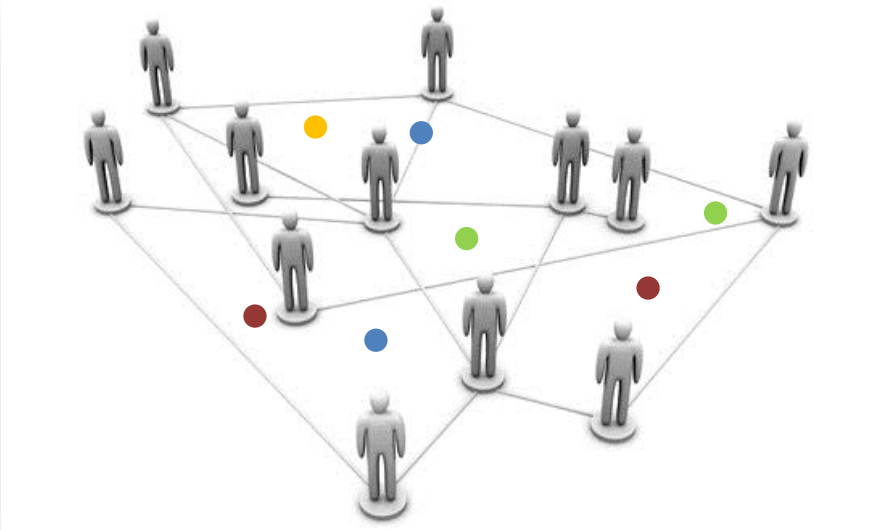
statistical testing units

„energy agents“



- comprised by a set of independent **statistical testing units**
- the role of each unit is **to test the degree of change between sentiment statistics obtained for two points in time**

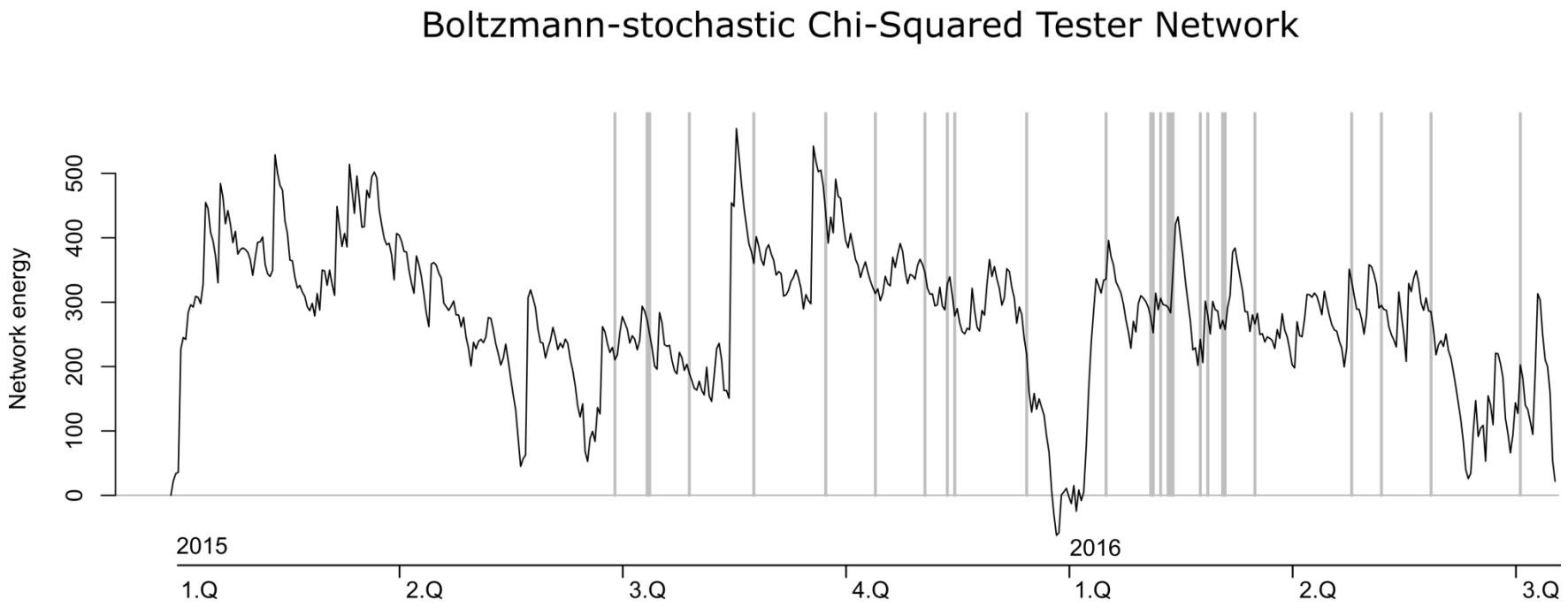
Social networks – Predictive Policing



- an energy value is assigned to each unit
- for each observation (sentiment and comment frequencies), it is decided whether a given unit conducts a statistical test **on the new data depending on its energy value**
- **in the test, the new data is compared to the data the unit processed during the last test!**
  → the network memorizes the data and underlying dynamics!

Social networks – Predictive Policing

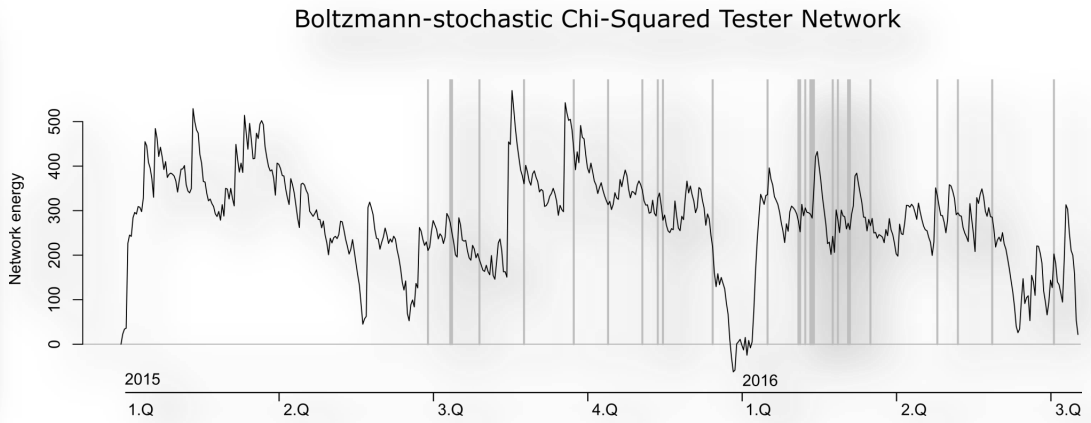
Sentiment analyses of user comments made on officials PEGIDA's facebook page
→ per day 'temper' tracking



The networking energy is the sum of all unit energy values
→ Representation of 'network awareness'

Social networks – Predictive Policing

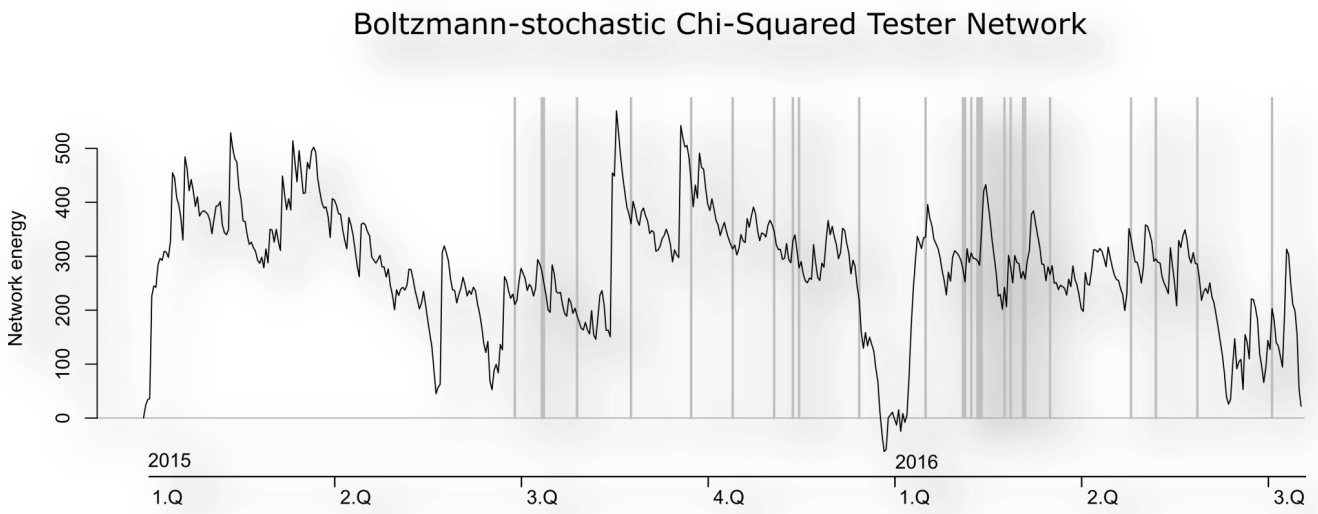
Sentiment analyses of user comments made on officials PEGIDA's facebook page
→ per day 'temper' tracking



Are energy peaks correlated to major criminal incidents with anti-Islam and anti-refugee background?

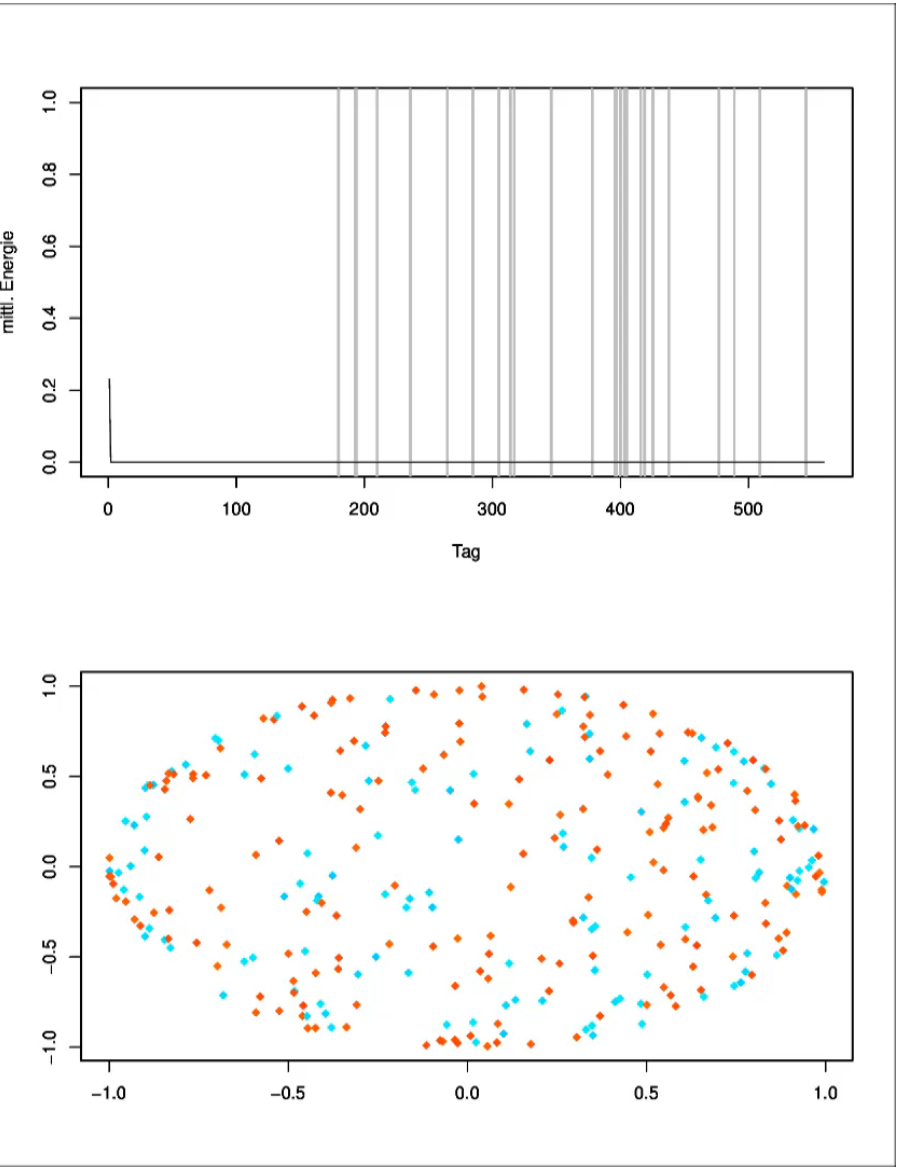
The networking energy is the sum of all unit energy values
→ Representation of 'network awareness'

Social networks – Predictive Policing



Are energy peaks correlated to major criminal incidents with anti-Islam and anti-refugee background?

Social networks – Predictive Policing

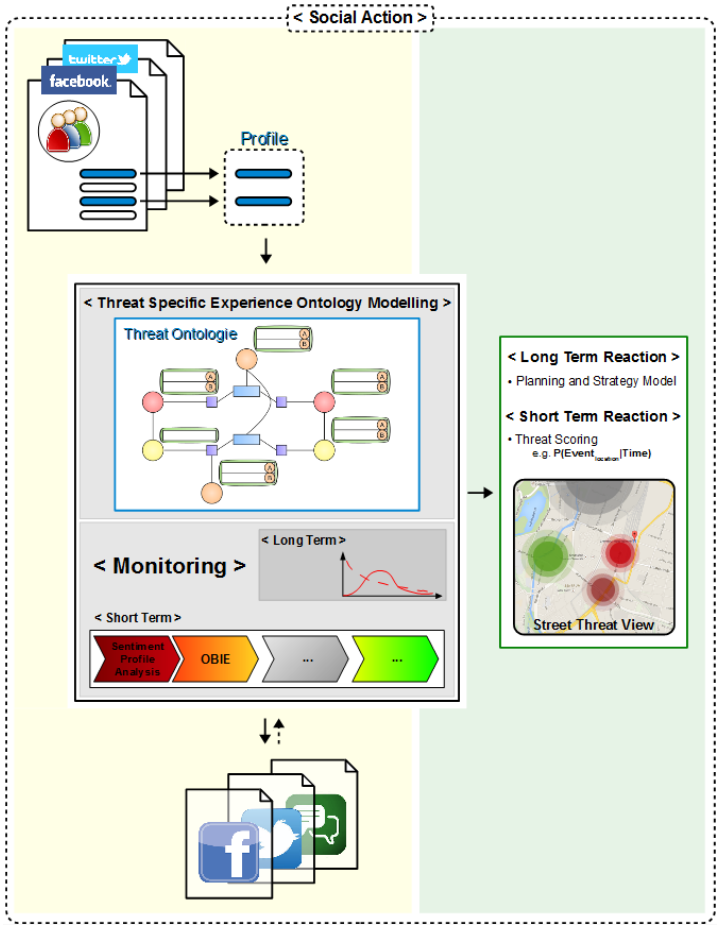


The networking energy is the sum of all unit energy values
→ Representation of 'network awareness'

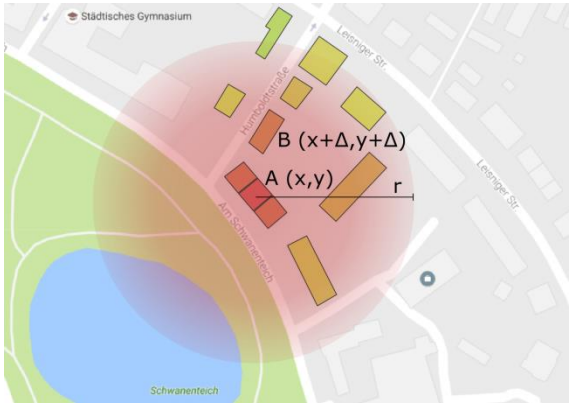
Possible events

predictive policing 2.0

BigData – Predictive Policing 2.0



Number of criminal offences and their „conditions“



Combination of two statistically approaches:

Real world and Digital world

Data from social networks: communication behavior of “virtual” groups

Is this the future of policing?

Is everything allowed?

Wednesday, July 25 - 15:45 - 17:30

How Dark is the Darknet? -- Dangers and Possibilities for the Digital Society

Time for discussions

FEEL FREE TO ASK QUESTIONS

VISIT US AT: www.bioforscher.de/FoSIL