



Risiken durch den Einsatz von IT und KI für Beschäftigte — Anforderungen aus technischer Sicht an einen modernen Beschäftigtendatenschutz

Prof. Dr. Hannes Federrath

Präsident der Gesellschaft für Informatik (GI) e.V.

Universität Hamburg, Sicherheit in verteilten Systemen (SVS)

Beirat Beschäftigtendatenschutz, Bundesministerium für Arbeit und Soziales, Berlin, 11. November
2020

Der Arbeitsbereich Sicherheit in Verteilten Systemen (SVS)

- Unsere Forschungsthemen (Auswahl)
 - IT-Sicherheitsmanagement, und -Grundschutz, ISO 27001
 - Privacy im Internet, Schutz vor Beobachtung, IT-Forensik
 - Sichere und datenschutzfreundliche Vernetzung von Fahrzeugen
 - Sicherheit und Datenschutz in mobilen Systemen
- Beiträge und (interdisziplinäre) Ergebnisse
 - Begleitung von Gesetzgebungsverfahren aus technischer Sicht
 - Erforschung des Spannungsfeldes von Freiheit und Sicherheit
 - Technische Lösungen zum Grundrechtsschutz
 - Informatik als gesellschaftliche Aufgabe
- Weitere Informationen
 - <https://svs.informatik.uni-hamburg.de>



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Mit rund 20.000 persönlichen Mitgliedern die größte neutrale Fachgesellschaft für Informatik im deutschsprachigen Raum

■ Mitglieder der GI

- Informatikerinnen und Informatiker aus Forschung und Lehre
- IT-Fachleute aus Verwaltung, Wirtschaft oder Industrie
- Lehrkräfte, die an einer Schule Informatik unterrichten
- Auszubildende und Studierende

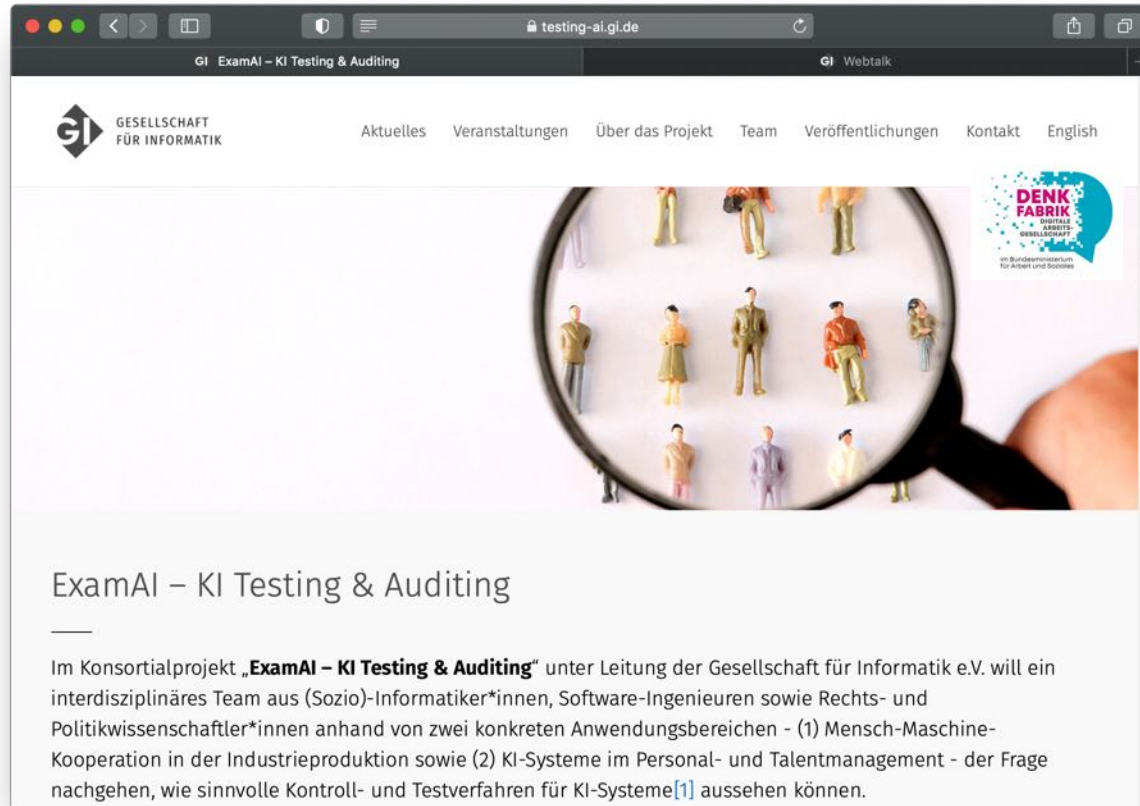
GESELLSCHAFT
FÜR INFORMATIK



- 14 Fachbereiche
- 150 Fachgruppen
- 30 Regionalgruppen

Die GI hat ein großes Netzwerk von Partnern. Beispiele:





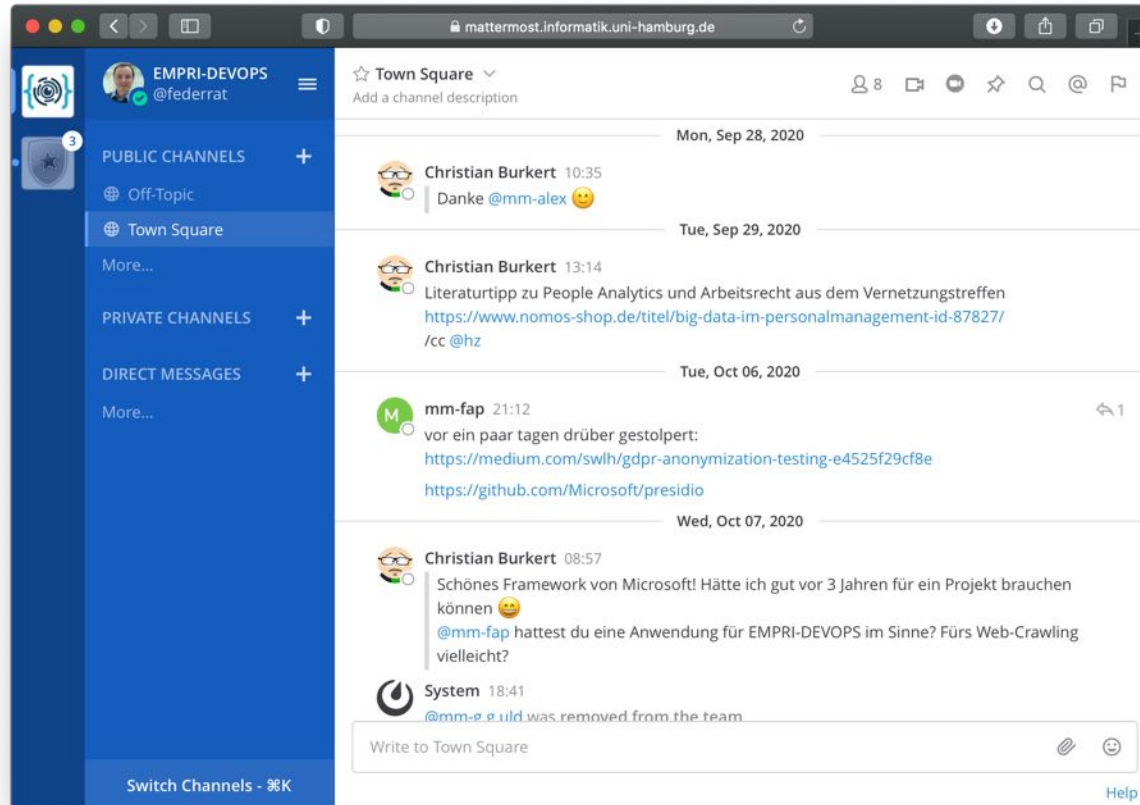


Beschäftigte sind heute während der Nutzung von IT-Systemen nahezu lückenlos überwachbar

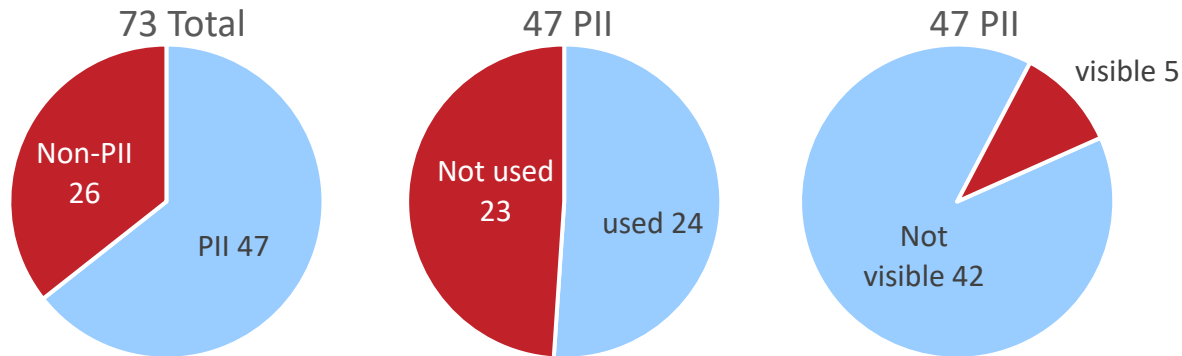
- **Dienstbetreiber (z.B. Betriebs-, Cloudsysteme, App-Anbieter)**
 - kennen teilweise die Inhalte
 - kennen die Metadaten von Beschäftigten
 - Zeit, Ort, IP-Adresse der Dienstnutzungen
 - Risiken beziehen sich jeweils »nur« auf das jeweilige Ökosystem (Microsoft, Apple, SAP, ...)
- **Arbeitgeber (oder zumindest dessen technisch-administratives Personal)**
 - lückenlose, plattformübergreifende Überwachung aller Inhalte und Verkehrsdaten (Metadaten) möglich
 - technische Möglichkeit bedeutet nicht, dass von den verfügbaren Daten tatsächlich Gebrauch gemacht wird



The image shows a screenshot of the EMPRI-DEVOPS website. The browser's address bar displays 'empri-devops.de'. The website's header includes the logo and the text 'EMPRI-DEVOPS', followed by navigation links: 'ÜBERSICHT', 'BLOG', 'VERBUND', 'PUBLIKATIONEN', and 'KONTAKT'. The main content area has a light blue background with the title 'DATENSCHUTZ FÜR DEVOPS' in large white letters. Below the title, it states: 'EMPRI-DEVOPS IST EIN FORSCHUNGSPROJEKT ZUR PRIVATSPHÄRE AM ARBEITSPLATZ VON SOFTWAREENTWICKLERN UND ADMINISTRATOREN (DEVOPS)'. To the right of this text is an illustration of a laptop displaying code. Below the main content area, there is a list of three bullet points, each with a red square icon. The first bullet point is 'Überwachung von Leistung' with a sub-point 'Schwach an Montagen?'. The second is 'Fortschritt' with a sub-point 'Hängt an einer Aufgabe?'. The third is 'Verhaltensweise' with a sub-point 'Arbeitet nach Mitternacht?'. To the right of these bullet points is a cartoon illustration of two desks separated by a vertical line. On the left desk, a person is sitting at a computer. On the right desk, a person is sitting at a computer, and a large magnifying glass is positioned over the desk, symbolizing surveillance or monitoring.



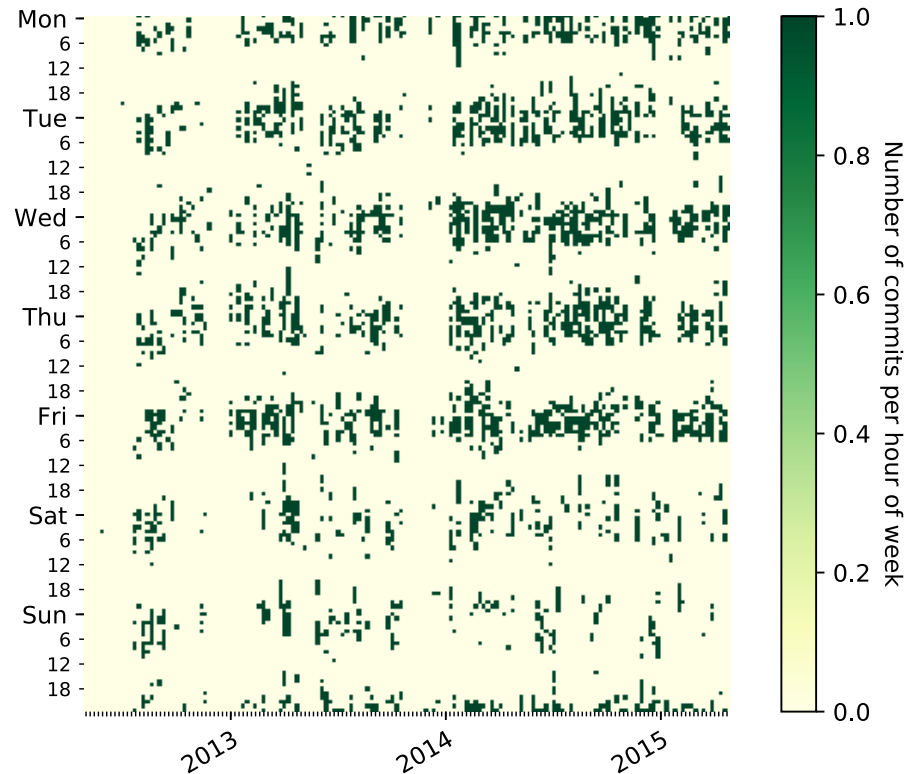
- Semiautomatische Analyse des Kollaborationstools Mattermost
 - Auswertung der Datenfelder mit Personally Identifiable Information (PII)



- Fokus der Studie waren Softwareentwickler und ihre Tools
- These:
 - Methoden (und Ergebnisse?) auf andere Domänen anwendbar

Visualisierung der Aktivität von Softwareentwicklern

■ Entwicklung der Anzahl der Commits eines Softwareentwicklers



■ Identity Disclosure

- Nutzer A hat um 3 Uhr noch einen Text hochgeladen – und zwar über den WLAN-Hotspot eines Nachtclubs!

Es war Alice!

■ Attribute Disclosure

- Bob hat seit heute früh 93 E-Mails verschickt.

Er hat die vorgeschriebene Höchstarbeitszeit nicht eingehalten!

■ Inferential Disclosure

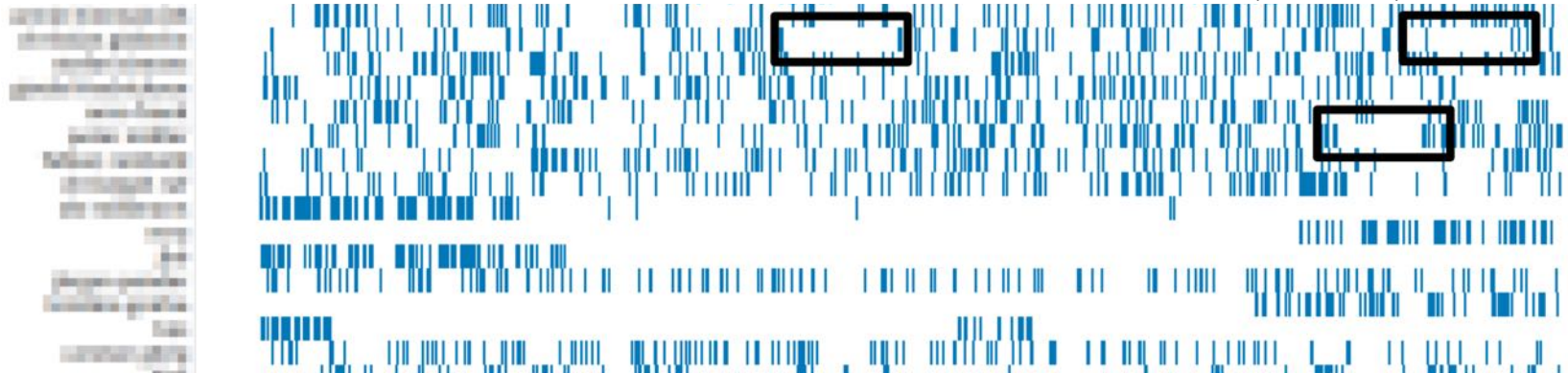
- Carol schreibt oft »finally« in ihren öffentlichen Posts.

Beschäftigte, die häufig finally schreiben, arbeiten weniger gründlich. Wir laden sie nicht zum Vorstellungsgespräch ein.

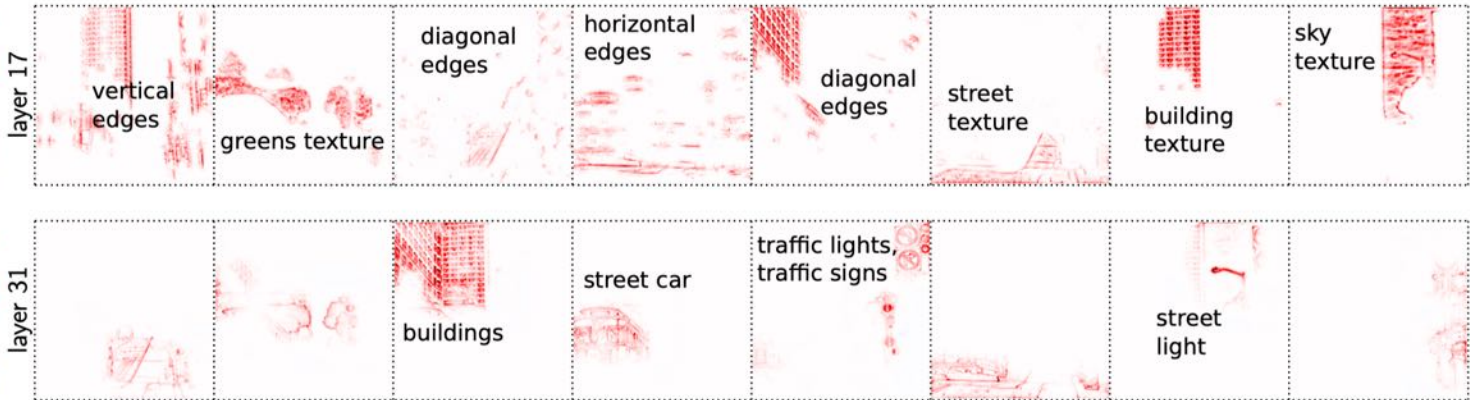
<http://www.dkriesel.com/spiegelmining>

- Analyse der Metadaten von 100.00 Spiegel-Online-Artikeln
 - Spiegel Online veröffentlicht etwa 100 Artikel am Tag
 - davon die Hälfte in den Ressorts Politik, Sport und Panorama
 - kürzeste Texte in den Ressorts Politik, Sport und Panorama
 - längste Texte im Ressort Kultur
 - Texte des Ressorts Kultur gehen morgens später und abends früher Online
 - Urlaubszeiten einiger Spiegel-Autoren aus Metadaten erkennbar

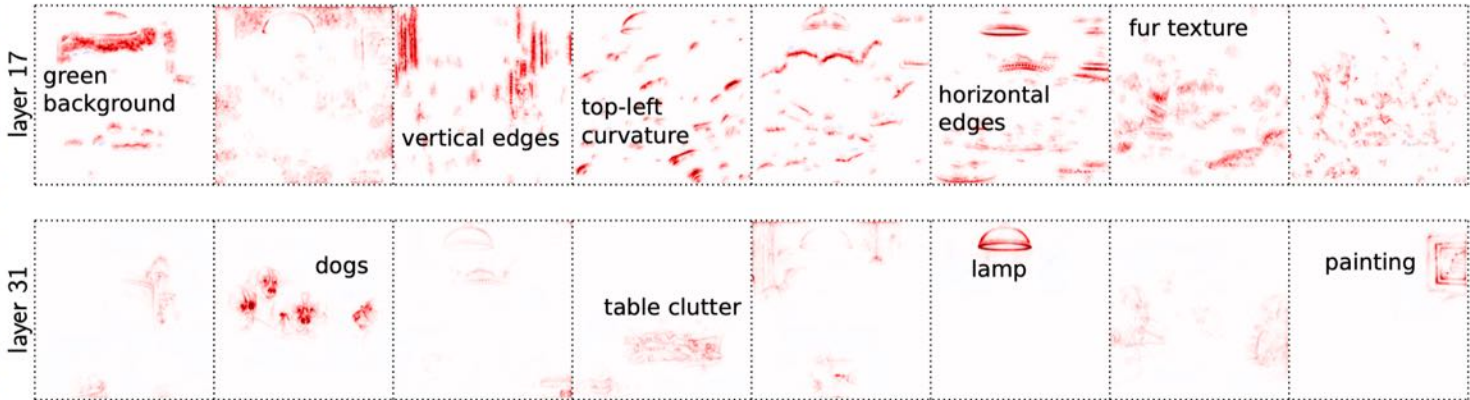
»Wer, wann was, mit wem?«



City and streetcar

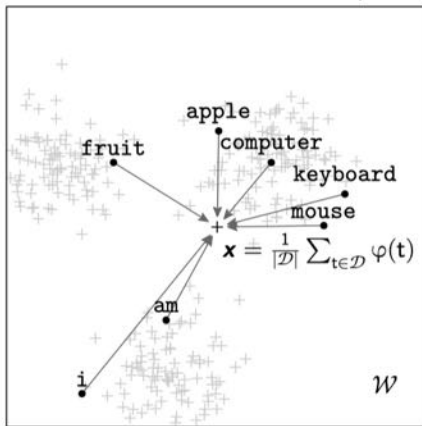


"Poker Game" (Coolidge, 1894)

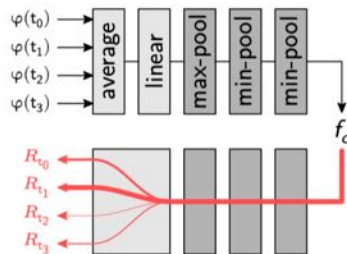


J. Kauffmann, M. Esders, G. Montavon, W. Samek, K. Müller, From Clustering to Cluster Explanations via Neural Networks CoRR, 2019, <https://arxiv.org/abs/1906.07633>

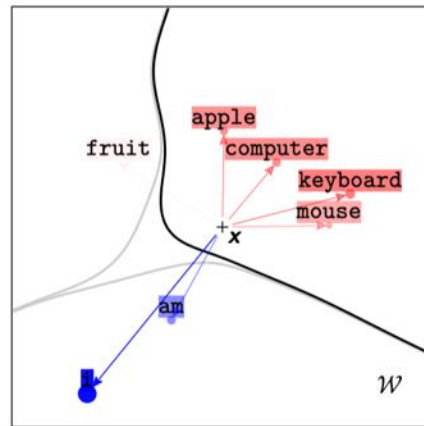
A. Document in Word-Vector Space



B. Network View of Redistribution



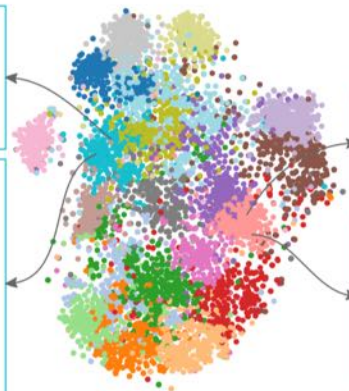
C. Function View of Redistribution



True Labels



Cluster Assignments



D. talk.politics.guns

Even if it were a **capital offense**, the **warrant** was not even an **arrest warrant**, but a **search warrant**. In other words, there was no **evidence of illegal arms**, just enough of a suggestion to get a **judge** to sign a license to search for **illegal evidence**.

E. sci.crypt

You can find the **salient difference** in any number of **5th amendment** related Supreme **Court** opinions. The **Court** limits 5th **amendment** **protections** to what they call "**testimonial**" **evidence**, as opposed to physical **evidence**.

The whole question would hinge on whether a **crypto key** would be considered "**testimonial**" **evidence**. I suppose **arguments** could be made either way, though obviously I would hope it would be **considered testimonial**.

F. rec.motorcycles

I'm not sure on the **older bikes**, but the **Yamaha Virago 535** has spec'd **seat height** of 27.6 in. and the **Honda Shadow** 27.2 in.

G. misc.forsale

For Sale: A Thule **Car rack** with 2 **bike holder accessories**. Comes with **Nissan Pathfinder** brackets but you can buy the appropriate ones for your **car** cheap. Looking for \$100.00 for everything. I live in the Bethesda area. Thanks for your interest.

Algorithmen sind nicht neutral, sondern enthalten immer Wertungen

■ Machine Learning:

- Auswahl des Trainingssets erzeugt stereotype Muster
- verfestigt solche Muster (algorithmische Normalisierung)
- Indeterminismus erzeugt Unschärfe
- macht Algorithmen anpassungsfähiger

Learning Motor Primitives for Robotics

Jens Kober, Jan Peters

Department of Empirical Inference and Machine Learning

Max Planck Institute for Biological Cybernetics

<https://www.ias.informatik.tu-darmstadt.de/Research/LearningMotorPrimitives>

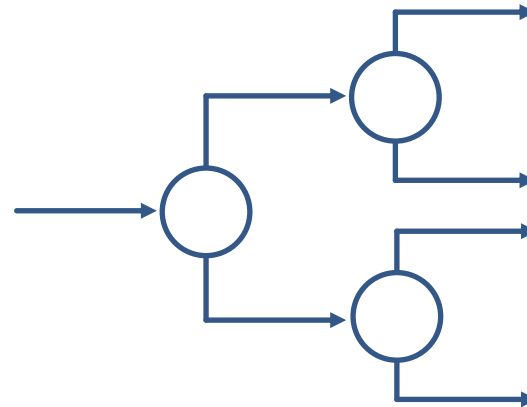
Algorithmen sind nicht neutral, sondern enthalten immer Wertungen

■ Machine Learning:

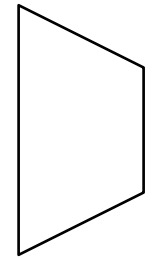
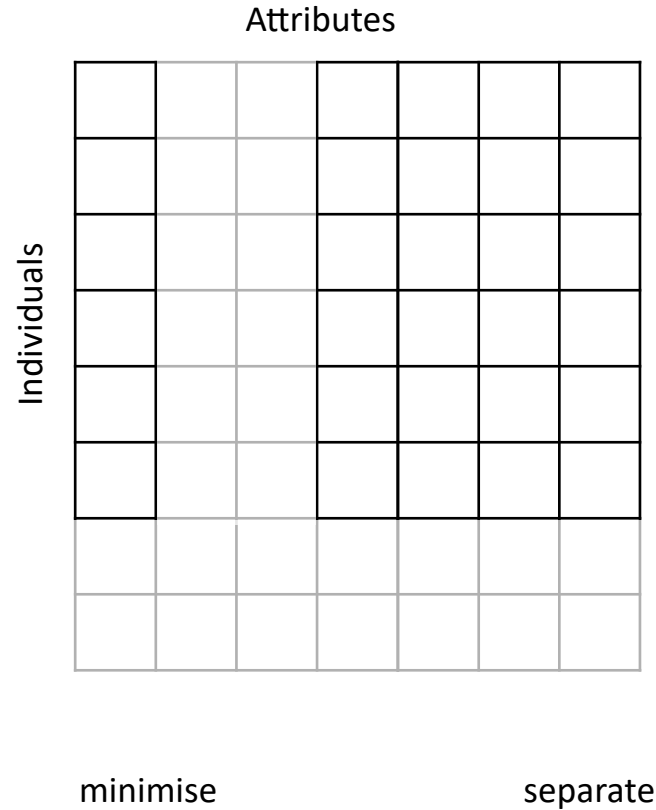
- Auswahl des Trainingssets erzeugt stereotype Muster
- verfestigt solche Muster (algorithmische Normalisierung)
- Indeterminismus erzeugt Unschärfe
- macht Algorithmen anpassungsfähiger

■ Einfache Entscheidungslogiken

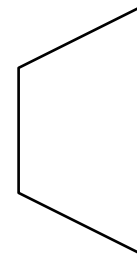
- Entwickler hat expliziten Entscheidungsbaum hinterlegt
- Deterministische Verhaltensweise



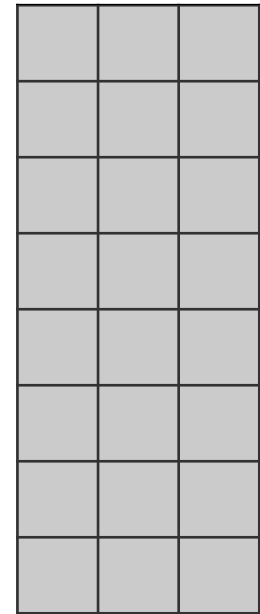
- Verknüpfung scheinbar harmloser Daten
 - Anwendung von Data Mining auf scheinbar harmlose Daten, um Erkenntnisse zu gewinnen
 - Risiko falscher Schlussfolgerungen: »guilt by association«
- Sekundäre Nutzung / Zweckänderung
 - Anwender haben möglicherweise zugestimmt, Ihre Privatsphäre für einen bestimmten Anwendungszweck aufzugeben,
 - aber nicht für weitergehende Zwecke
- Ausschluss / heimliche Datenverwendung
 - Die betroffene Personen wissen nicht, welche Informationen für welchen Zweck verwendet werden.



aggregate



perturbate



hide

■ Technisch

- **Minimise**: Nur notwendige Daten speichern und verarbeiten
- **Separate**: Daten verteilt verarbeiten und speichern
- **Aggregate**: Daten auf das notwendige Maß zusammenfassen
- **Perturbate**: Daten durch zufällige Störungen ungenau machen
- **Hide**: Daten nicht in offener Form speichern

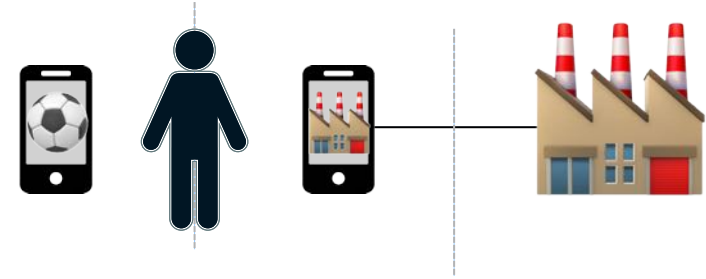
■ Organisatorisch

- **Enforce**: Durchsetzung einer Datenschutz-Policy (access control)
- **Inform**: Betroffene über Datenverwendung informieren (P3P)
- **Control**: Eingriffsmöglichkeit der Betroffenen (informed consent)
- **Demonstrate**: Überprüfbarkeit (privacy management, logging)

Bring Your Own Device (BYOD) erfordert auf dem Endgerät ...

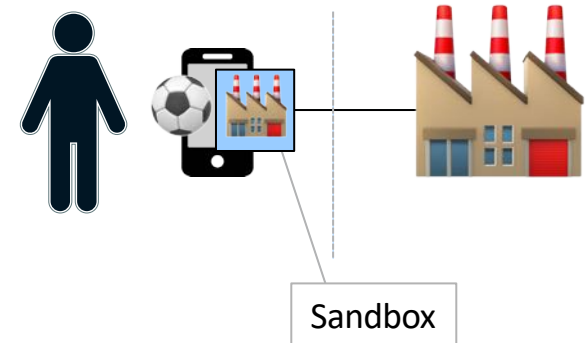
■ entweder: strikte Trennung von Privatem und Beruflichen

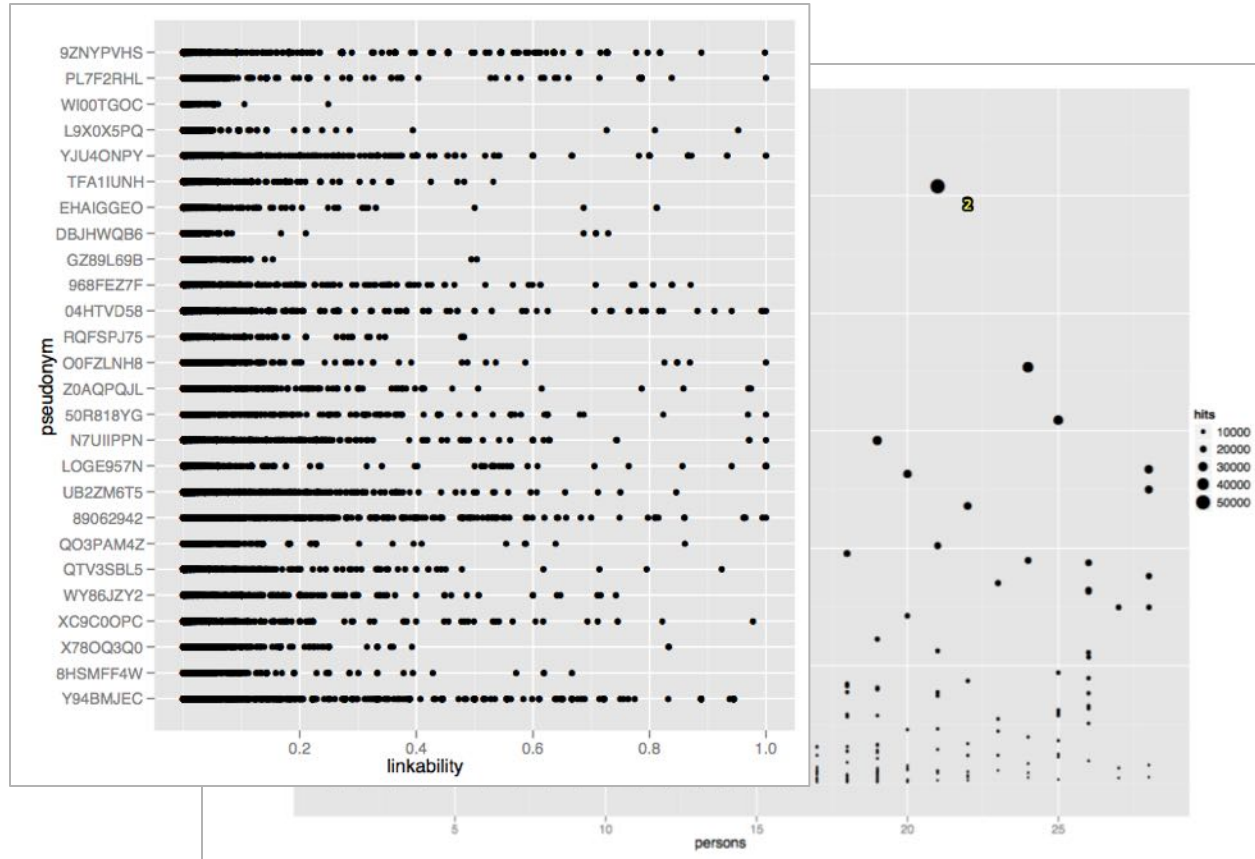
- macht BYOD in vielen Szenarien unmöglich
- Mobile Device Management nur auf beruflichem Gerät
- Analogie: Verbot der privaten Nutzung des Internet am Arbeitsplatz



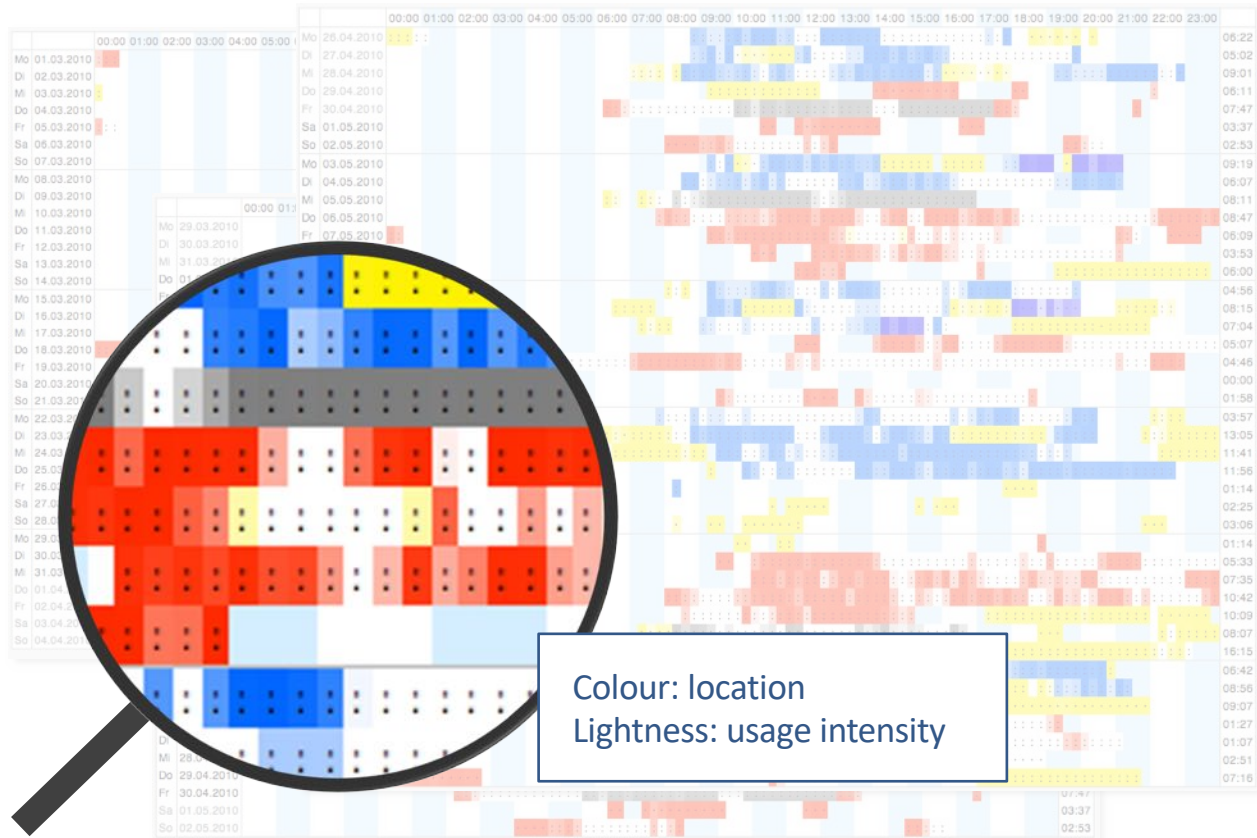
■ oder: Verzicht des Arbeitgebers auf Mobile Device Management und nahezu jegliche Zugriffsmöglichkeit auf das lokale Gerät

- Vorschlag: Verbot der betrieblich veranlassten Steuerung und Kontrolle der privaten Endgeräte während der beruflichen Nutzung
- netzseitige (betriebsinterne) Überwachung bleibt weiterhin technisch möglich

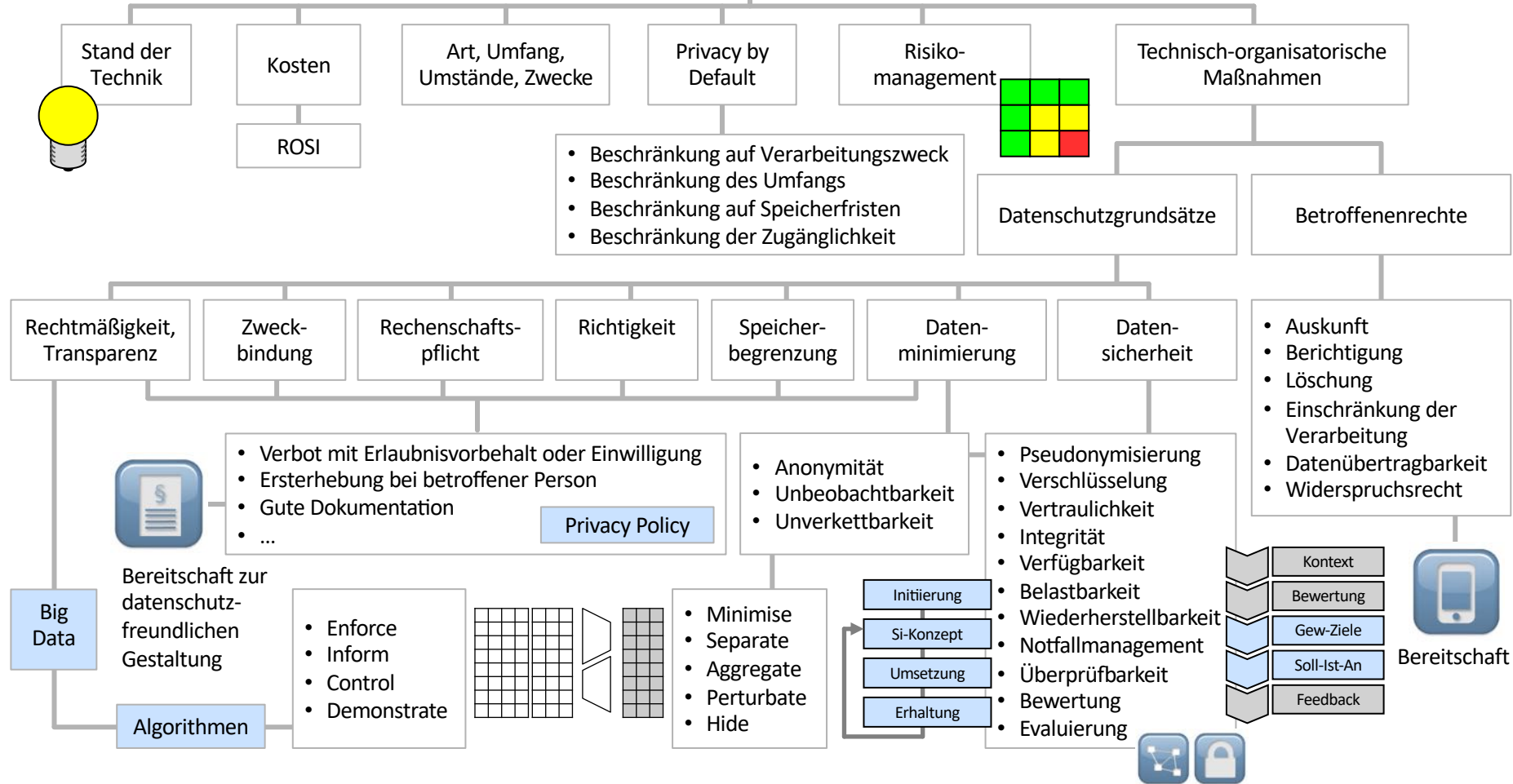




Usage timelines and ip-geo-tagging



Privacy by Design



Regelungsempfehlungen für den Einsatz von IT und KI im Beschäftigtendatenschutz

- Bei erlaubter privater Nutzung von dienstlichen Geräten: strenge Vorgaben für Arbeitgeber
- Instrument der Datenschutz-Folgenabschätzung bei KI-basierten Systemen intensiv nutzen
- Grundsätze der Speicherbegrenzung und Datenminimierung kontrollieren und konsequent durchsetzen
- Anlernen der Systeme durch Daten mit frühzeitiger Minimierung des Personenbezugs (Aggregation, Perturbation) – Pseudonymisierung spielt hier vermutlich eher eine untergeordnete Rolle
- Mitbestimmungsregeln der Personalvertretung beim (unterstützenden) Einsatz von automatisierten Entscheidungssystemen
- Konsequentes Verbot vollautomatisierter Entscheidungssysteme im Personalwesen, ggf. Offenlegungs- und Dokumentationspflicht der Tools des Arbeitgebers
- Generell: Die Vorgaben der DSGVO sollten streng zugunsten der Beschäftigten ausgelegt werden.





Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DEPARTMENT OF INFORMATICS
SECURITY AND PRIVACY

[HOME](#) [COURSES](#) [THESES](#) [RESEARCH](#) [PEOPLE](#) [SERVICE](#)



SECURITY AND PRIVACY

Foto: UHH/Denstorf

[UHH](#) → [MIN-Fakultät](#) → [Fachbereich Informatik](#) → [Einrichtungen](#) → [Arbeitsbereiche](#) → [Security and Privacy](#) → [Home](#)

WORKING GROUP ON «SECURITY AND PRIVACY»

Security and Privacy

Information systems become more and more important in critical infrastructures, while the Internet has evolved to a critical infrastructure itself. The secure operation of these infrastructures is vital and their failure can have severe impacts up to the loss of human lives.

Security refers to the fact that protection goals are achieved in the presence of malicious attacks and system failures. Typical security goals can be confidentiality, integrity, accountability, and availability. Security and privacy in information systems addresses both technical and organizational aspects, such as building and establishing security concepts and security infrastructures as well as risk analysis and risk management.

Privacy can be a conflicting goal to security, but they can also benefit from each other. Hence, it is necessary to balance both when developing secure information systems.

Prof. Dr. Hannes Federrath
Fachbereich Informatik
Universität Hamburg
Vogt-Kölln-Straße 30
D-22527 Hamburg

Telefon +49 40 42883 2358

federrath@informatik.uni-hamburg.de

<https://svs.informatik.uni-hamburg.de>