

Die Plattformisierung von (Wissens-)Arbeit: Datenintegrations- und -analysenplattformen und die Transformation der Arbeitswelt in der Datengesellschaft

Dr. Simon Egbert

Universität Bielefeld
Fakultät für Soziologie

*ERC-Projekt: The Future of Prediction. Social
Consequences of Algorithmic Forecast in Insurance,
Medicine and Policing*



The Future of Prediction

► Projects People ► Resources Contact

The Social Consequences of Algorithmic Forecast in Insurance, Medicine and Policing

Elena Esposito, **Bielefeld University** / University of Bologna

Social field	Dimension	Challenge	Data and methods
Personalized Insurance	Individualization	<ul style="list-style-type: none"> personalization of risk calculation inversion of information asymmetry proactivity 	<ul style="list-style-type: none"> archives of personalized car insurance policies online offers of collaborative insurance platforms; interviews <i>Open-F@b Call4Ideas</i> for preventive insurance models
Precision Medicine	Generalization	<ul style="list-style-type: none"> n=all vs. n-of-1 non-statistical notion of health right not-to-know 	<ul style="list-style-type: none"> algorithmic semantic analysis of biomedical journals archives HREB requests for approval court cases; debates in medical and bioethical literature
Predictive Policing	Bias	<ul style="list-style-type: none"> distinction prevention and repression intransparency of predictive algorithms 	<ul style="list-style-type: none"> ethnographic research in police departments citizen's petitions and administrative answers

Übersicht

- 1) Einleitung: Plattformen in der Datengesellschaft
- 2) Beispiel: Palantir Technologies
- 3) Datenintegrations- und -analyseplattformen als Plattformen
- 4) Das transformative Potential von Datenintegrations- und analyseplattformen in der Polizeiarbeit
- 5) Fazit

1) Einleitung: Plattformen in der Datengesellschaft

- In einer Gesellschaft wie der heutigen, die (fast) vollständig digitalisiert ist und in der es immer mehr Daten gibt, stehen viele Organisationen und Institutionen vor dem Problem, eine immer größer werdende Menge an Daten nutzen zu müssen, die an immer mehr Orten zu finden ist.
- In diesem Zusammenhang spielen Plattformen in mindestens zweierlei Hinsicht eine bedeutsame Rolle:
 - Zum einen sind sie eine der Hauptquellen der Datenflut. Insbesondere die großen Social-Media-Plattformen wie TikTok und Instagram sind für die stetig wachsende Datenmenge verantwortlich.
 - Auf der anderen Seite sind Plattformen auch prominente Werkzeuge, um die Datenflut sinnvoll zu nutzen bzw. überhaupt handelbar zu machen
 - => Insbesondere Datenintegrations- und -analyseplattformen, wie sie derzeit vor allem von Palantir Technologies genutzt werden, haben sich in unterschiedlichen gesellschaftlichen Kontexten – Polizei, humanitäre Arbeit, Gesundheitswesen, Einwanderung etc. – etabliert.

- Aufgrund der aktuellen Flut solcher Plattformen ist davon ausgehen, dass solche Plattformen die Zukunft der Arbeit mit Daten in vielen gesellschaftlichen Bereichen mitbestimmen und sowohl die dortige Wissensgenerierung als auch die Entscheidungsprozesse stark prägen werden.
- Aufgrund ihres zentralen Ziels, Daten zu aggregieren und assoziativ verarbeitbar zu machen, und aufgrund ihrer modularen Architektur sind solche Plattformen nämlich ideal geeignet, um in viele verschiedene datenbezogene Arbeiten integriert zu werden.
- Kernziel dieser Plattformen ist es dabei, datenbankübergreifende Analysen zu ermöglichen, indem die wechselseitige Verknüpfung von heterogenen Datenbeständen forciert wird. => Data Warehouse wird simuliert, wo es kein Data Warehouse gibt.
- Vor diesem Hintergrund ist anzunehmen, dass Plattformen dieser Art in Zukunft eine gewichtige Rolle spielen werden – nicht nur aber vor allem auch in der Transformation der Arbeitswelt.

2) Beispiel: Palantir Technologies

- Das bekannteste Beispiel für eine Datenanalyse- und -integrationsplattform, stammt von dem in den USA ansässigen Unternehmen Palantir Technologies.
- Das 2003 gegründete und nach den sehenden Steinen (*palantíri*) aus J. R. R. Tolkiens Roman Herr der Ringe benannte Unternehmen ist ein digitales Technologieunternehmen, das gemeinhin als bemerkenswert geheimnisvoll und in der Öffentlichkeit umstritten gilt, nicht zuletzt aufgrund seines bewusst schattenhaften Auftretens einerseits und des Investors Peter Thiel andererseits.
- Zudem: Eine Erstinvestition kam neben der von Thiels Firma Founders Fond auch von In-Q-Tel, dem Investmentarm des US-Auslandsgeheimdienstes CIA.
- Aus ihrem Börsenprospekt geht hervor, dass Palantir im Jahr 2020 125 Kunden hatte (Palantir, 2020b). In Europa arbeitet Palantir im Sicherheitsbereich für Europol, den französischen Inlandsgeheimdienst und für die dänischen Polizeikräfte. Zudem mit Ferrari, Airbus, dem World Food Programme und dem National Health Service (NHS) im Vereinten Königreich.

- Das Geschäftsmodell von Palantir besteht nach offiziellen Angaben nicht aus dem Sammeln und Verkaufen von (personenbezogenen) Daten – wie bei vielen anderen digitalen Plattformen:

„Palantir is not a ‘data broker’ or ‘data aggregator.’ Unlike many tech companies, our business model is not based on the monetisation of personal data. We do not collect, store, or sell personal data. We don’t use personal data to train proprietary AI or machine learning models to share or resell to other customers.” (Palantir, 2020)

- Stattdessen bietet Palantir hochgradig modularisierbare Plattformen für die Integration und Analyse bereits vorhandener Daten:

„We build digital infrastructure for data-driven operations and decision-making. Our products serve as the connective tissue between an organisation’s data, its analytics capabilities, and operational execution. Palantir’s platforms tie these together by bringing the right data to the people who need it, allowing them to take data-driven decisions, conduct sophisticated analytics, and refine operations through feed-back.“ (Palantir, 2020)

Our Offerings

Gotham

The operating system for
defense & resilience.

DEFENSE

INTELLIGENCE

DISASTER RESPONSE

Foundry

The operating system for the
modern enterprise.

MANUFACTURING

ENERGY

FINANCE

AVIATION

SHIPPING

HEALTHCARE

PHARMA

LOGISTICS

RETAIL

& MANY MORE

Apollo

The operating system for
continuous delivery.

ON-PREMISE DATA CENTERS

CLASSIFIED NETWORKS

EMBEDDED EDGE DEVICES

MULTI-CLOUD & HYBRID CLOUDS

- Die grundlegende Funktionsweise der Datenintegrations- und -analyseplattformen von Palantir ist eigentlich recht simpel: Sie bieten eine Benutzeroberfläche, über die Kunden auf ihre Daten – die an verschiedene Orten liegen können – zentral zugreifen und diese analysieren können.
- Die Besonderheit ist dabei, dass die Palantir-Plattformen in der Lage sind, Daten aus verschiedenen strukturierten und unstrukturierten „Datensilos“ in einer einheitlichen Umgebung oder einem Datenökosystem zu kombinieren, was die Interoperabilität der Daten(quellen) verbessern soll.
- Der grundlegende Ansatz der Palantir-Plattformen ist daher die *Desilosierung*: Getrennte Datensilos, die nicht gemeinsam analysiert werden können, sollen aufgebrochen und miteinander verknüpft werden um Zusammenhangsanalysen möglich zu machen (Palantir, 2020).

“Palantir is a data integration, analysis, and visualization platform that helps in-fuse legacy information source with new value by unlocking previously inaccessible workflows and applications. Palantir fuses disparate legacy data systems (with their varied formats, data types, object models, access control restrictions, and retention rules) and maps those disparate data sets into a unified view with a harmonized and flexible data model. Once all data is integrated, Palantir exposes it to powerful analytical tools that enable (...) analysts to search authorized data, generate rich analytical products, and discover new insights in a common environment. In addition to providing a powerful suite of analytical tools, Palantir is designed to be flexible and configurable, allowing organizations to operate responsibly in complex and evolving legal and ethical regimes.” (Palantir, n.d.-a, p. 3)

3) Datenintegrations- und analyseplattformen als Plattformen

- Obwohl es in den letzten Jahren ein bemerkenswertes Wachstum der Literatur zu Plattformstudien gibt (z.B. Gillespie, 2010; Helmond, 2015; van Dijck et al. 2018; Gorwa, 2019; Dolata & Schrape 2022; Kirchner & Ametowobla, 2023), wurden bisher kaum Studien speziell zu Datenintegrations- und Analyseplattformen durchgeführt.
- Palantir verwendet selbst ausdrücklich den Plattformbegriff. Entsprechende Hinweise finden sich an vielen Stellen, zumeist ohne gesonderte Begründung (P. Cipierre & J. Hiesserich, 2022; Palantir, n.d.; Payne et al., 2008; Wojtas & Cipierre, 2023).
- In einer Veröffentlichung von Palantir-Mitarbeitern heißt es jedoch, dass man bei Palantir folgendes teile: „the belief in the need for analytic platforms, not products. Openness and flexibility are key to being able to cope with the demands of a broad range of analytic tasks.” (Payne et al., 2008, S. 202)

=> Dies deutet darauf hin, dass Palantir den Begriff der Plattform für seine eigenen Produkte verwendet, um deren offenes und flexibles Design zu betonen.

- Nennenswerte Ausnahmen in der Analyse von Palantir Plattformen stammen von Brayne (2017; 2021), Munn (2017; 2018) und Iliadis & Acker (2022)
- Diese fokussieren insbesondere auf den Überwachungsimpetus, der diesen Plattformen zu eigen ist: indem versucht wird, so viele Daten wie möglich zu sammeln, laufen die Plattformen Gefahr, gewichtige Datenschutzrisiken zu implizieren.
- Dieser Fokus auf Überwachung ist zwar nachvollziehbar, aber nicht in der Lage, zu erklären, wie die Nutzung von Datenintegrations- und -analyseplattformen die Arbeit von Menschen verändert, insbesondere mit Blick auf ihre Prozesse der Wissensgenerierung und Entscheidungsfindung. Gerade in diesen Prozessen haben die Plattformen ihre größten transformatives Potential.

=> Datenintegrations- und -analyseplattformen sind mehr als „surveillance platforms“
(Iliadis & Acker 2022)

4) Das transformative Potential von Datenintegrations- und -analyseplattformen in der Polizeiarbeit

- Seit 2017 stetige „Plattformisierung“ der Polizeiarbeit in Deutschland:
 - 2017: hessenDATA in Hessen
 - 2020: DAR (Datenbankübergreifende Analyse- und Recherche) in NRW
 - 2023: VeRA (Verfahrensübergreifende Recherche und Analyse) in Bayern
 - ???: Bundes-VeRA in weiteren Bundesländern
- Alle Plattformen basieren auf der Gotham-Plattform von Palantir

Home - PPTM_Demo

Erstellung Bearbeiten Home Einstellungen Anwendungen Plugins Hilfe

Home Browser Graph Karte Objektexplorer Kooperation Picture Cards Table

- Ein Update von Gotham auf die Version 3.20.7 wurde durchgeführt.
- Die Massensuche wurde nun integriert und ist unter „Keyword Search“ zu finden. Der BulkSearch-Helfer ist dadurch nicht mehr nötig (Details sind den Multiplikatoren/Trainern/Support-Teams bekannt).
- Das Problem, dass einzelne Eigenschaften aus Convor in hessenDATA noch zu finden waren, wurde gelöst.

Bekannte Fehler:

- es liegen aktuell keine bekannten Fehler vor.

Viele Grüße, Euer hessenDATA-Team!

Nach Schlüsselwort suchen Erweiterte Hilfe

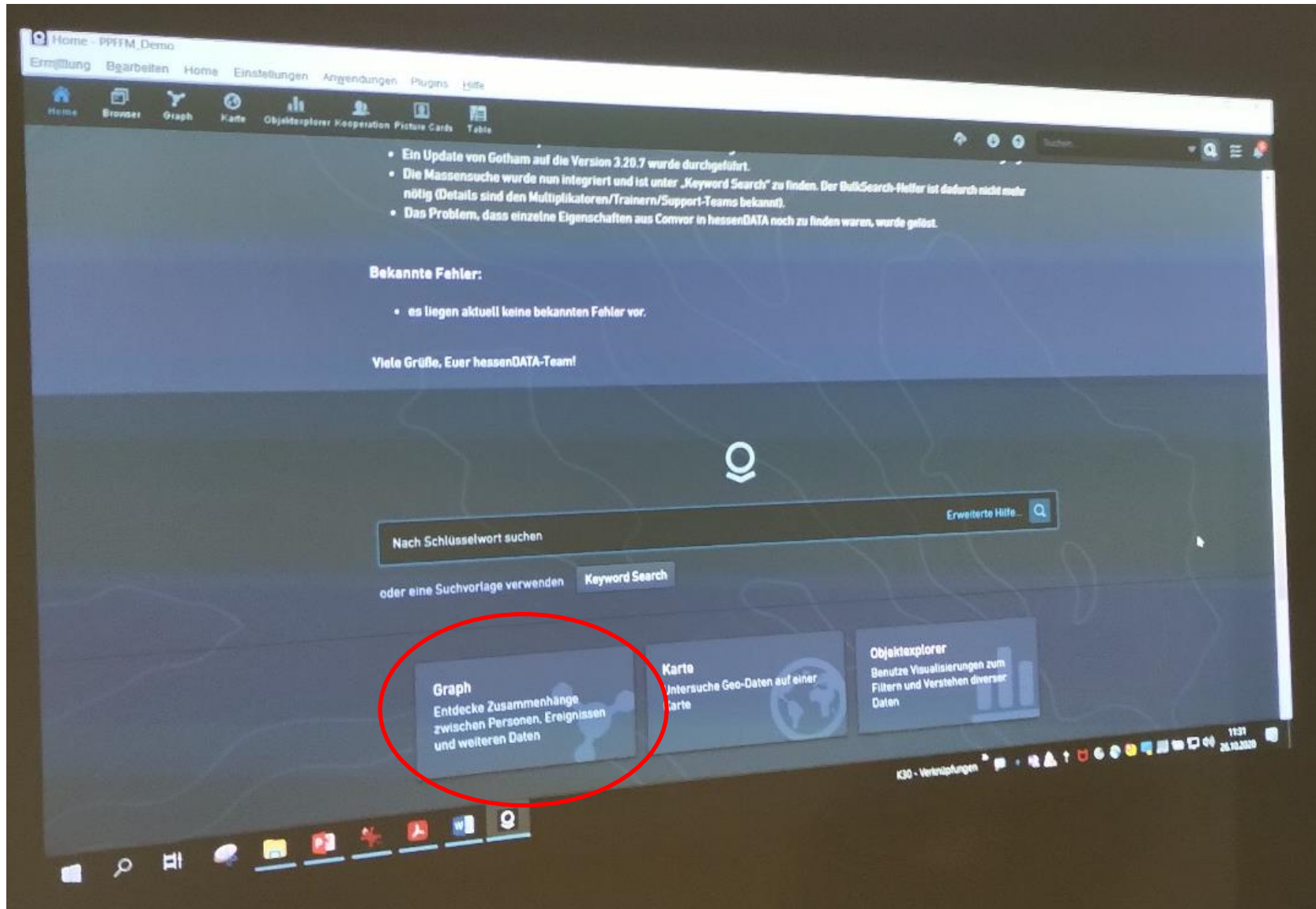
oder eine Suchvorlage verwenden **Keyword Search**

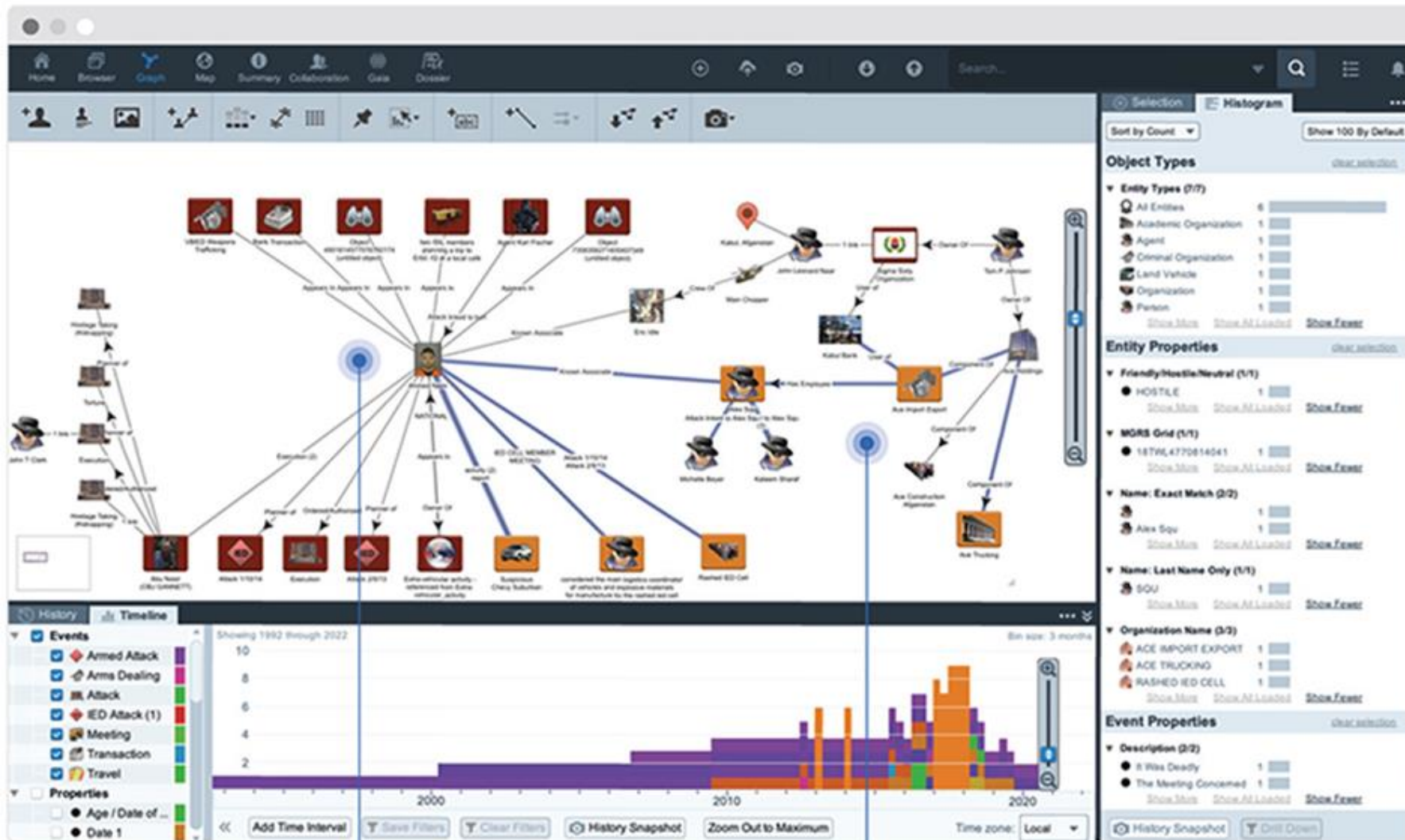
Graph
Entdecke Zusammenhänge zwischen Personen, Ereignissen und weiteren Daten

Karte
Untersuche Geo-Daten auf einer Karte

Objektexplorer
Benutze Visualisierungen zum Filtern und Verstehen diverser Daten

K30 - Verknüpfungen 11:31 25.10.2020





Investigators can map out key pieces of evidence associated with an entity of interest, such as phone records associated with a suspicious phone number.

Analysts can explore relationships between different parts of computing infrastructure in order to analyze and understand cyber threats.

- Durch Applikationen wie „Graph“ sorgen Datenintegrations- und –analyseplattformen für einen epistemischen Fokus der Anwender:innen auf Zusammenhänge zwischen Daten
- eine assoziative Wissenslogik wird reproduziert, die sich insb. Auf soziale Netzwerkanalyse konzentriert

=> Personen, die mit einer gefährlichen/verdächtigen Person in Kontakt standen, sind per einem Verdacht ausgesetzt (guilt by association/Kontaktschuld)

- Verdachtskonstruktionsprozess wird auf den Kopf gestellt: Statt auf Grund eines Verdachts begründen zu können, Daten einer Person näher zu analysieren, wird durch die Datenanalyse selbst der Verdacht konstruiert und es zeigt sich erst hinterher, ob die Analyse gerechtfertigt war.
- Durch den Fokus auf Zusammenhänge, werden ferner Mythen der Big Data-Befürworter:innen aktualisiert (z.B. Anderson 2007): je mehr Daten man zur Verfügung hat, desto besser, weil mit (korrelativ fundierten) Informationen zu rechnen ist.

5) Fazit

- Datenintegrations- und -analyseplattformen verbreiten sich derzeit schnell und scheinen in Zukunft ein dominantes Werkzeug in vielen Gesellschaftsbereichen zu werden, mit denen Organisationen unterschiedlicher Art die stetig zunehmende Datenflut handhabbar gestalten wollen.
- Marktführer ist derzeit die Firma Palantir Technologies mit ihren Plattformen Apollo, Gotham und Foundry.
- Die Grundfunktion dieser Plattformen ist Desilosierung: das Aufbrechen von Datensilos und die Verknüpfung von Datenbeständen, die in unterschiedlichen Formaten und an unterschiedlichen Orten vorliegen.
- In der Polizeiarbeit zeigt sich, dass solche Plattformen eine Assoziationslogik verstärken, die die Logik von Wissens-/Verdachtsgenerierung verändern: Kontaktschuld und Umkehrung Verdachtslogik.



The Future of Prediction

Vielen Dank für Ihre Aufmerksamkeit!

simon.egbert@uni-bielefeld

This project is funded by the European Research Council

PREDICT (Advanced Research Project) no. 833749