Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Vorlesung

Datenschutz und Privatheit in vernetzten Informationssystemen

Kapitel 2: Digitale Identitäten

Erik Buchmann
buchmann@ipd.uka.de

Wiederholung: Prinzipien des Datenschutzes

- Jeder Bürger soll selbst bestimmen können, und
- Jeder Bürger soll wissen,
  - wer was wann und unter welchen Bedingungen
    - über ihn weiß.
    - über ihn in Erfahrung bringen darf.

- Ausnahmen nur auf gesetzlicher Basis
  - wenn das Interesse Dritter bzw. der Allgemeinheit schwerer wiegt als die Schutzinteressen des Betroffenen

Klassische Datenschutz-Schwerpunkte

- Schutz des Bürgers vor dem Staat
  - Recht auf freie Meinungsäußerung wird bedeutungslos, wenn Regierung den Sprecher im nachhinein identifizieren (und abstrafen) kann
    → **Datenschutz wichtig für Demokratie**

- Schutz des Bürgers vor privaten Unternehmen
  - Sind die individuellen Vorlieben und Absichten des Käufers bekannt, wird perfekte Preisdifferenzierung und Manipulation des Käufers möglich
    → **Datenschutz ist Kundenschutz**

Neue Schwerpunkte der letzten Jahre

- Schutz des Bürgers vor dem Dienstanbieter
  - Verknüpfung und Mining von personenbezogenen Daten, die als Nebenwirkung moderner Dienste anfallen, erlaubt unerwartete Rückschlüsse
    (Beispiel: Google erhebt keine Daten über Nutzer, sondern Nutzer geben diese über Suchbegriffe preis.)
    → **Verbleib persönlicher Daten nachvollziehen**

- Schutz des Bürgers vor dem Bürger
  - Werden in Online-Communities private Details und Beziehungen öffentlich preisgegeben, lassen sich Persönlichkeitsprofile von Unbeteiligten erstellen.
    → **Kontrolle über persönliche Daten behalten**

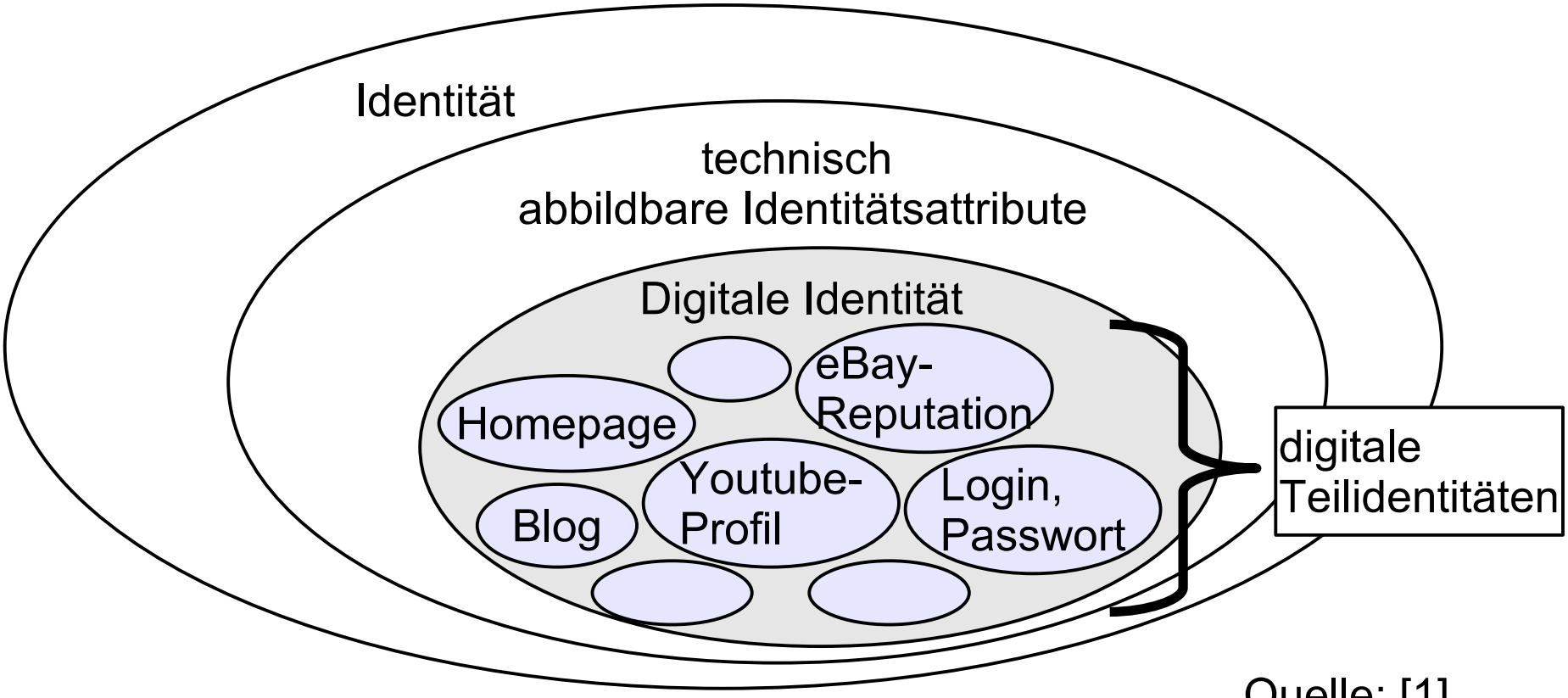
Die Digitale Identität

Die Digitale Identität

- Definition: *"Jede mögliche Form von technisch abgebildeten Daten, die zu einer Person gehören" [1]*
  - Daten zur eindeutigen Authentifizierung, z.B. Adresse, Name, biometrische Daten
  - Daten zur pseudonymen Identifizierung, z.B. Login, Passwörter, Nicknames, Foren-Namen
  - Persönliche Merkmale, z.B. Vorlieben, Hobbies, Religion, Lebensumfeld
  - nicht unbedingt *von jedem* einer Person zuordnbar
    - Beispiel: IP-Adresse ist Teil der digitalen Identität, aber nur vom Internet-Provider zuordnbar

Übersicht: Identität im Netz

- Digitale Teilidentität: Untermenge der digitalen Identität, die eine Untermenge der abbildbaren Attribute sind
  - viele separate **digitale Teilidentitäten** möglich



Quelle: [1]

Unterschiedliche digitale Teilidentitäten

- Datenspuren im täglichen Leben
  - Blogs, Soziale-Netzwerk-Portale, Internet-Communities
  - Finanzamt, Wählerlisten, Anträge bei Behörden
  - Verkehr mit Wirtschaftsunternehmen, Einkäufe im Supermarkt oder Web-Shops

- Nicht alle Datenspuren werden wissentlich hinterlassen
  - Was weiß der Supermarkt (Rabatt-System) oder der Mobilfunkprovider (Positionsdaten vom Handy)?

    → durch **Verkettung von Teilidentitäten** Aufbau umfassender Persönlichkeitsprofile möglich!

Eigenschaften von Identitätsattributen

Weniger relevant für Privatsphäre	Potentiell gefährlich für die Privatheit
anonym	eindeutig identifizierend
nicht wiedererkennbar	wiedererkennbar
ändert sich im Verborgenen über die Zeit	unveränderlich
leicht änderbar	nicht änderbar
weitergebbar / übertragbar	nicht weitergebbar/übertragbar
flüchtig	langfristig gespeichert
nur einmal verwendet	häufig wiederverwendet
Authentizität unklar	authentisch / bestätigt durch Dritte
Zugriff anderer nicht möglich bzw. kontrollierbar	Zugriff anderer möglich oder intransparent
ermöglicht keinen direkten Kontakt	ermöglicht unmittelbaren Kontakt
unauffällig/geht in der Masse unter	unnormal oder herausragend
für wenige Teile des eigenen Lebens relevant / als trivial empfunden	betrifft zentrale Bereiche des täglichen Lebens / besonders sensibel
keine zusätzlichen Informationen enthaltend	enthält für Weitergabe nicht abtrennbare Zusatzinformationen

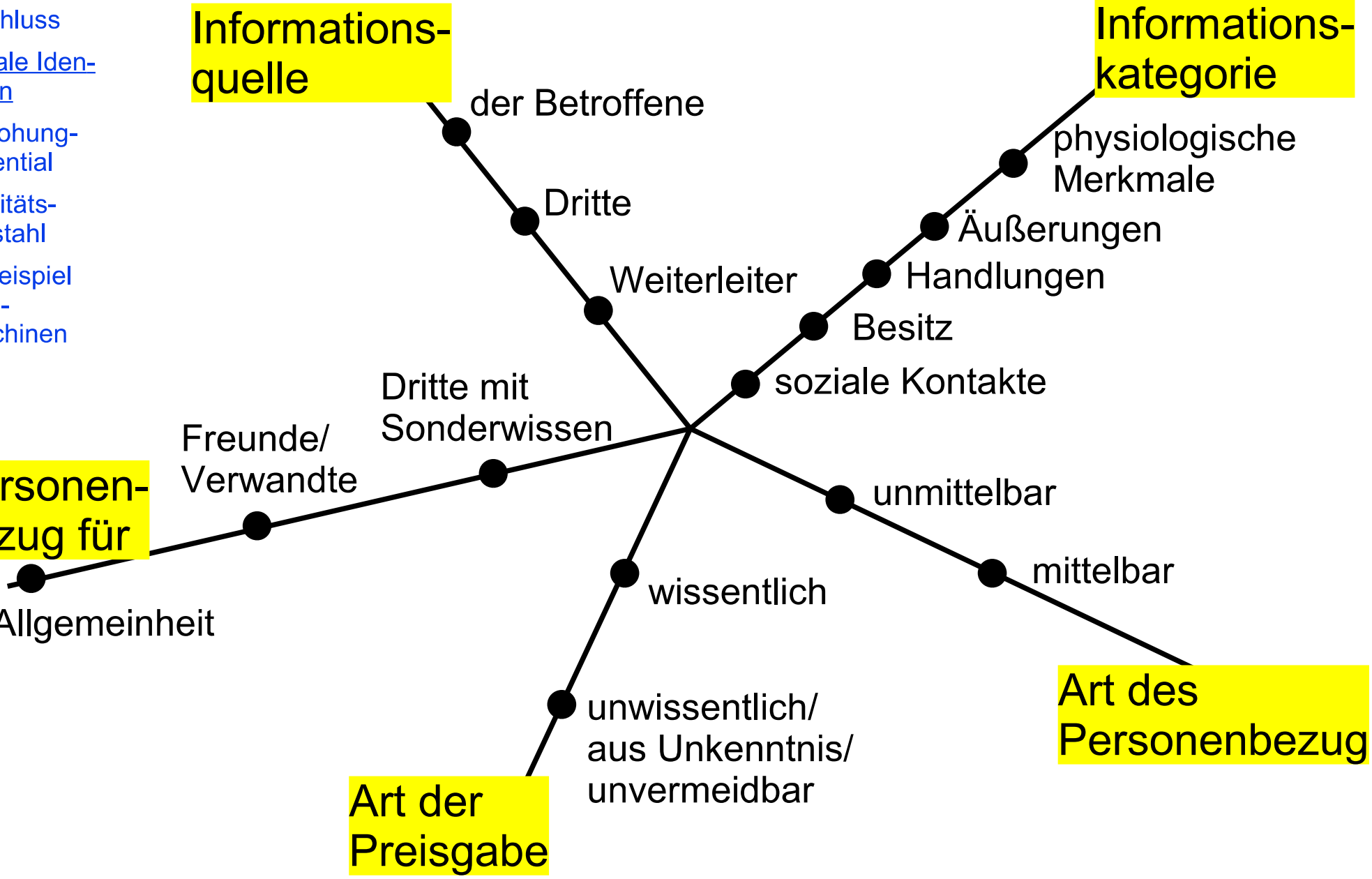
Quelle: [1]

Beispiel

- Biometrische Merkmale (Fingerabdruck, Gesicht)
  - stabil über die Zeit, kaum änderbar, nicht übertragbar, langfristig speicherbar (Reisepass), häufig wiederverwendet (Passkontrolle), authentisch, andere können darauf zugreifen (sofern nicht verschleiert)
  - **potentiell gefährlich für die Privatheit**

- selbstbestimmte Identitätsmerkmale (Login, Passwort)
  - anonym oder pseudonym, leicht änderbar, übertragbar an Dritte, Zugriff anderer nicht möglich, oft nur für unwichtige Teile des pers. Lebens relevant
  - **weniger gefährlich für die Privatheit**

Persönliche Daten: 5 relevante Dimensionen



Informationsquellen, unmittelbarer Bezug zum Betroffenen

	Informationsquelle: der Betroffene selbst	Informationsquelle: Dritte
Physiologische oder genetische Merkmale	z.B. gemessen von Sensoren: Größe, Gewicht, Augenfarbe, DNA, Fingerabdruck	z.B. Untersuchungen von Verwandten mit ähnlichen biometrischen Merkmalen
Äußerungen	z.B. Einträge in Web-Formulare, eigene Homepage, Profile in Online-Communities	z.B. Äußerungen über den Betroffenen, Bewertung in einem Reputationssystem
Handlungen	z.B. etwas kaufen, an einer Veranstaltung teilnehmen	z.B. im Namen anderer Personen tätig sein
Soziale Kontakte	z.B. Adressbuch-Einträge im Mailclient, Skype-Kontaktliste	z.B. veröffentlichte Kontakte in Online-Communities
Besitz	z.B. Kleinanzeigen, eBay-Angebote	z.B. im Auftrag des Betroffenen einkaufen

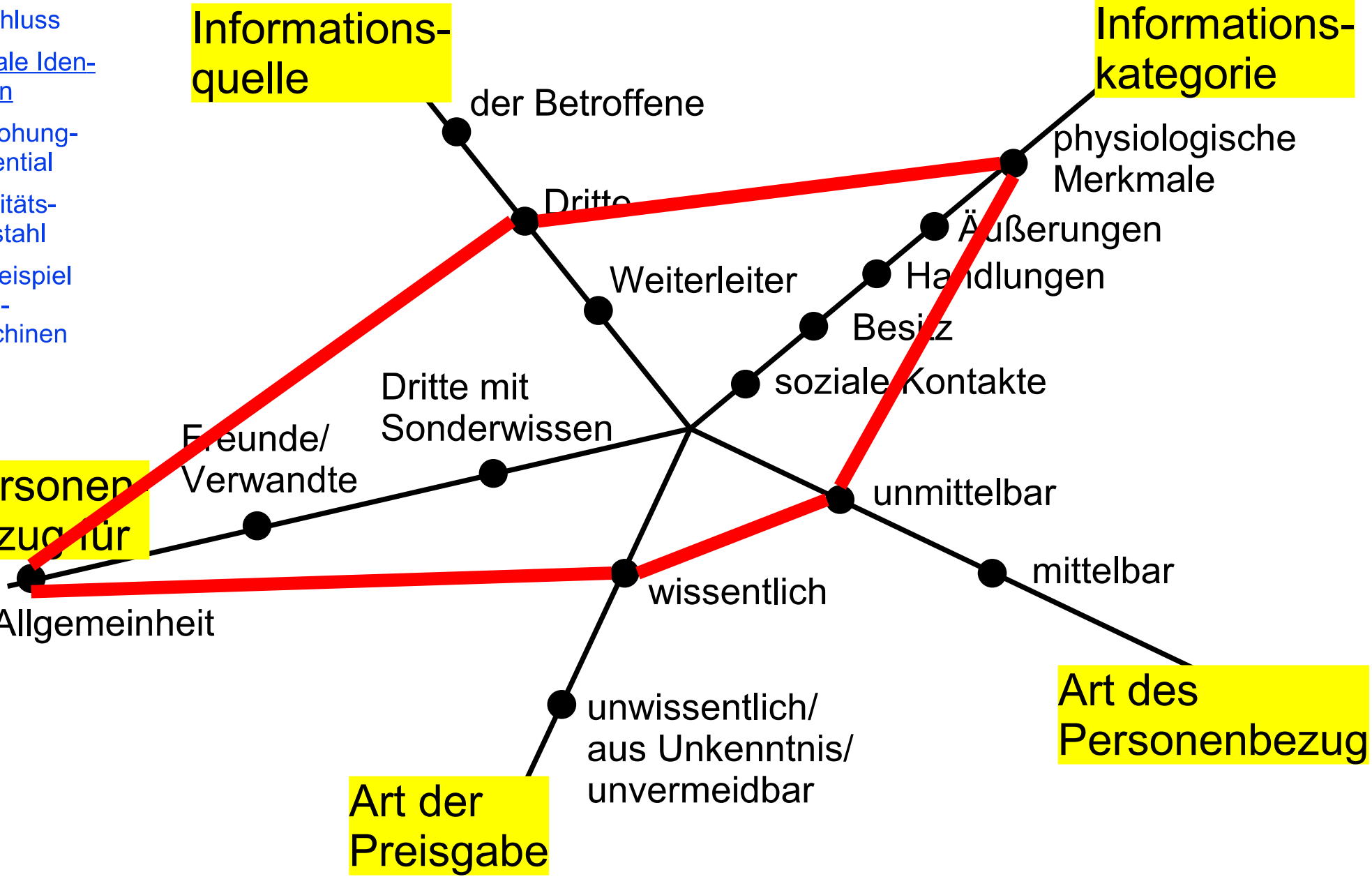
Quelle: [1], erweitert

Informationsquellen, mittelbarer Bezug zum Betroffenen

	Informationsquelle: Dritte
Äußerungen	z.B. Äußerungen über eine Personengruppe, der der Betroffene angehört
Handlungen	z.B. über Data Mining, Collaborative Filtering auf die Absichten des Betroffenen schließen
Indirekte Kontakte	z.B. Freund-meines-Freundes-Kontakte in Online-Communities
Besitz	z.B. Collaborative Filtering: Kundengruppe X kauft häufig Gegenstand Y

Quelle: [1], erweitert

Beispiel: Namentlich annotiertes Bild auf Flickr



Bedrohungen durch die digitale Identität

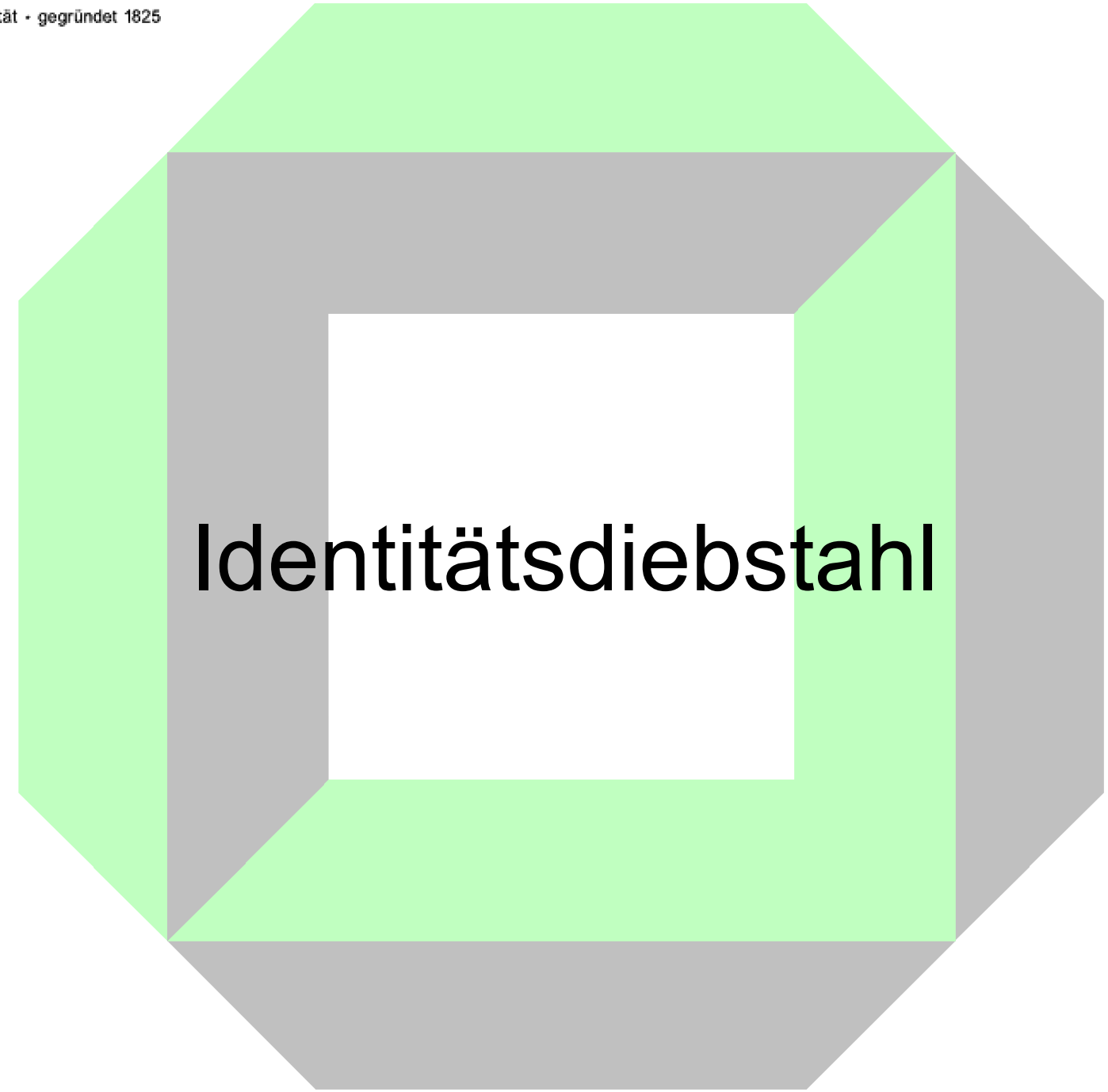
Probleme mit digitalen (Teil-)Identitäten

- Fehlentscheidungen durch Falschinformationen
- Datenmißbrauch
- Zweitverwertung
- Langfristige Aufbewahrung
- Verknüpfbarkeit
- Öffentliche Zugänglichmachung

(Anm.: Liste ohne Anspruch auf Vollständigkeit)

Verknüpfbarkeit

- Zusammenführen von “harmlosen” digitalen Teilidentitäten aus unterschiedlichen Quellen
  - öffentliche Quellen: Telefonbuch, Schufa-Daten, Handelsregister, Liegenschaftsbuch, etc.
  - nichtöffentlich: Daten aus dem Geschäftsbetrieb, Informationen von Behörden, Banken
- Anreichern von eigenen Daten mit Zusatzinformationen
- Suche nach diskriminierenden Merkmalskombinationen in verlinkten Daten → Rasterfahndung
- Problem
  - Aufbau von komplexen Persönlichkeitsprofilen
  - Fehler in den Daten können zu falschen Schlüssen führen


Identitätsdiebstahl

Mißbrauch personenbezogener Daten

- Personenbezogene Daten werden ohne Wissen oder Zustimmung des Betroffenen zu Zwecken gesammelt, gespeichert, verarbeitet oder übermittelt, die
  - den Interessen des Betroffenen zuwiderlaufen, z.B.
    - **Identitätsdiebstahl**
    - Stalking, Mobbing
    - Kreditbetrug
  - aber nicht gesetzlich legitimiert sind, z.B.
    - Steuerfahndung
    - polizeiliche Ermittlungen
- Für den Betroffenen nicht zu verhindern, da ohne Wissen und Zustimmung erfolgt

Identitätsdiebstahl USA 2008

- Zahlen 2008: Data Breach Stats (*"Einbruchstatistik"*)

	# of Breaches	*# of Consumer Records*
Banking/Credit/Financial	78	18,731,947
Business	240	5,886,960
Educational	131	806,142
Government/Military	110	2,954,373
Medical/Healthcare	97	7,311,833
total:	**656**	**35,691,255**

- Quelle: http://www.idtheftcenter.org, Breach Database *(Anm.: Zahlen beschreiben nur Verlust von pers. Daten, nicht die Mißbrauchsfälle)*

Identitätsdiebstahl USA 2009

- Zahlen 2009: Data Breach Stats (*"Einbruchstatistik"*)

	# of Breaches	*# of Consumer Records*
Banking/Credit/Financial	57	8,364
Business	208	132,402,177
Educational	78	803,667
Government/Military	90	79,470,963
Medical/Healthcare	65	10,461,818
total:	**498**	**223,146,989**

- Quelle: http://www.idtheftcenter.org, Breach Database *(Anm.: Zahlen beschreiben nur Verlust von pers. Daten, nicht die Mißbrauchsfälle)*

Account Takeover

- Übernahme einer bestehenden digitalen Identität
  - *Pishing*, gefälschte EMails erfragen Kontodaten, eBay-Konten, Kreditkartennummern, *www.meinebank.de.pisher.org*
  - *Pharming*, Webbrowser wird durch DNS-Spoofing o.ä. auf manipulierte Webseiten umgeleitet, die eBay- oder Banken-Webseiten gleichen
  - *Malware* auf dem Rechner protokolliert Anmeldeinformationen
  - *Social Engineering*,
    beliebtes Passwort: Name der Freundin,
    beliebte PIN: Geburtsdatum des Kindes

Zusammenfassung

Zusammenfassung

- digitale Identitäten als Untermenge aller technisch abbildbaren Attribute mit Personenbezug
- digitale (Teil-)Identitäten sind grundsätzlich bedrohlich, so harmlos sie auch zunächst scheinen mögen
- Identitätsdiebstahl als derzeit größte Bedrohung durch den unbedachten Umgang mit pers. Informationen
- Verkettung digitaler Identitäten führt zu Informationsanreicherung und Profilbildung
- Aktuelles Fallbeispiel: Verkettung in Suchmaschinen