

Digitale Spuren im Internet als Rohstoff erster Güte

Noah Walle



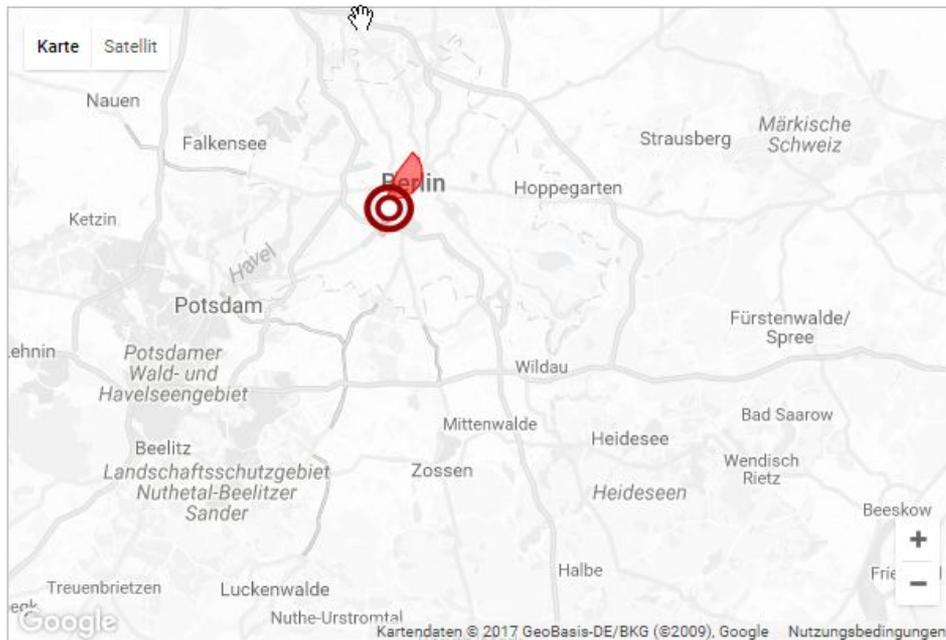
Spuren

- Hinweisgebende Hinterlassenschaften
- Durch Fortbewegung bewirkte Markierung auf dem Boden

→ Digitale Spuren sind vor allem Metadaten

Eigenschaften von Metadaten

- Einfach zu analysieren
- Produktion von Metadaten ist unvermeidbar
- Man kann von Metadaten auf Inhalte schließen
- Aggregation von Metadaten zur Gewinnung neuer Erkenntnisse



Montag, 31. August 2009



Malte Spitz spricht bei den Erlangerer Grünen zum Thema "Internetzensur ist nicht akzeptabel" (Quelle: [Parteiwebsite](#))



6 eingehende Anrufe
21 ausgehende Anrufe
Gesamtdauer: 1h 16min 8s



34 eingehende Nachrichten
29 ausgehende Nachrichten

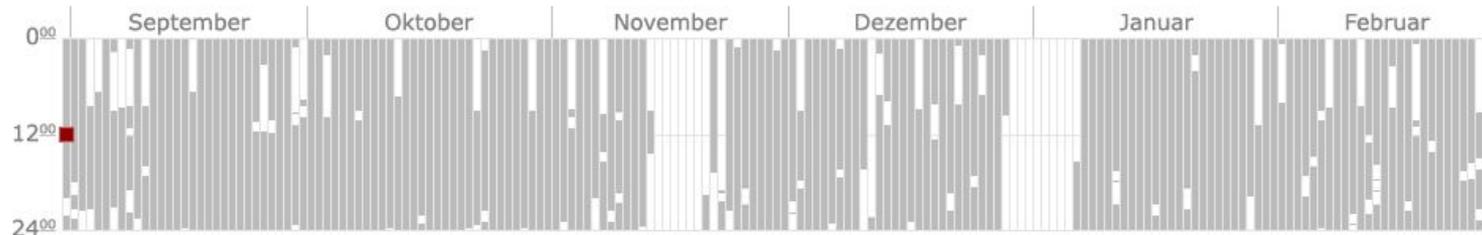


Dauer der Verbindung mit dem Internet: 21h 17min 25s



Wann hielt sich Malte Spitz im gewählten Kartenausschnitt auf ?

[Download Datensatz](#)



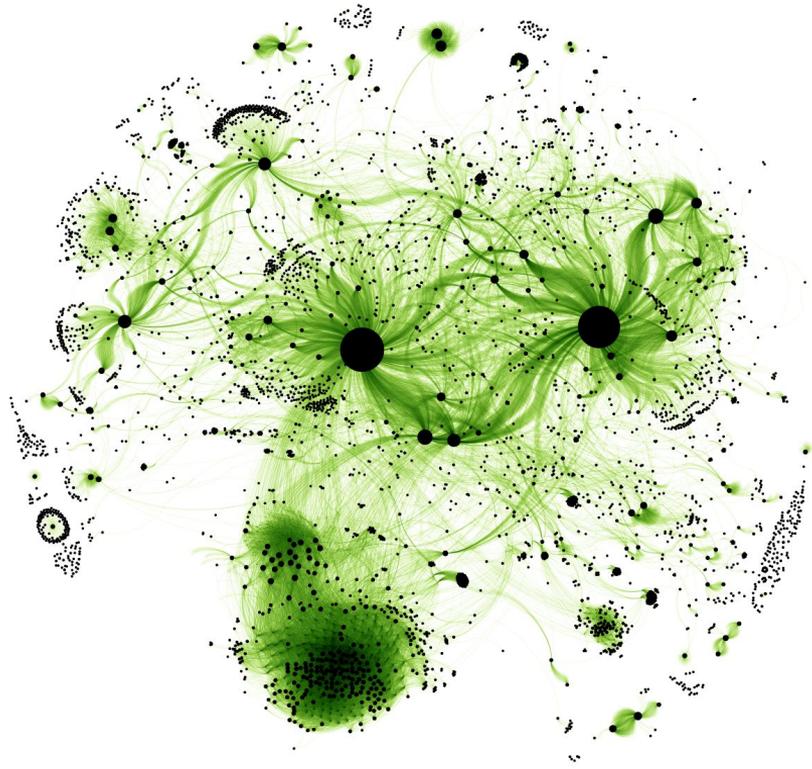
Eigenschaften von Metadaten

- Einfach zu analysieren
- Produktion von Metadaten ist unvermeidbar
- Man kann von Metadaten auf Inhalte schließen
- Aggregation von Metadaten zur Gewinnung neuer Erkenntnisse

Soziales Netzwerk anhand von E-Mail-Kommunikation

Schwarze Punkte: 4.000 E-Mail-Adressen

Verbindungslinien: 22.000 E-Mails



Rohstoff Metadaten

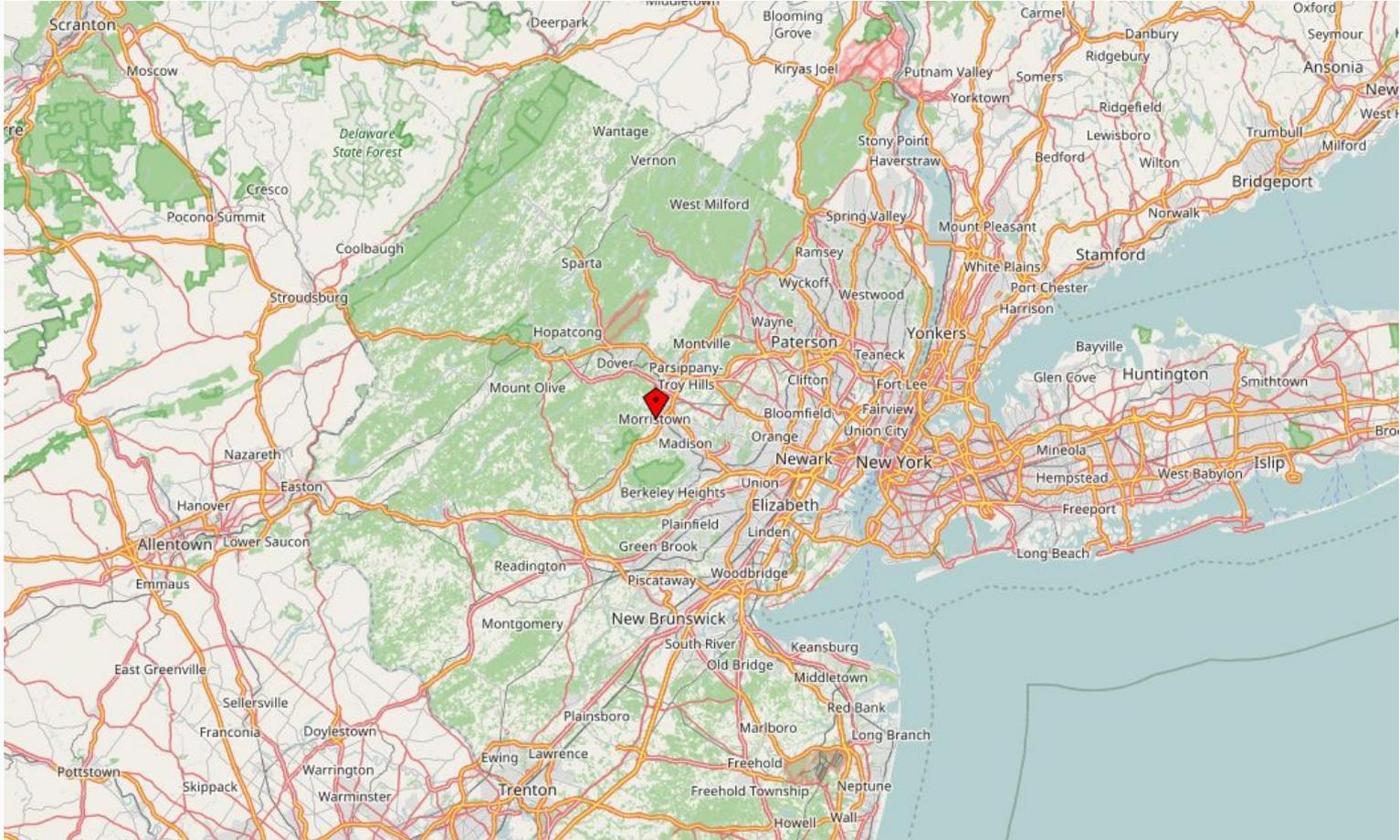
- Wie wird der Rohstoff Metadaten verarbeitet? - Ein Beispiel
- Studie der Forschungsabteilung von AT&T:
 - “Clustering Anonymized Mobile Call Detail Records to Find Usage Groups”

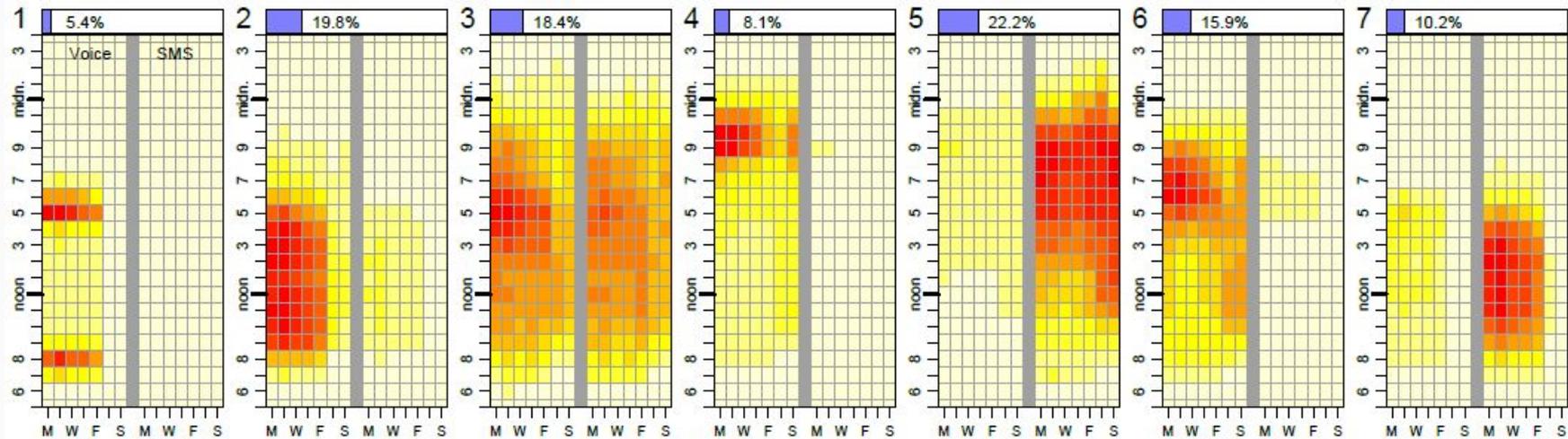


AT&T

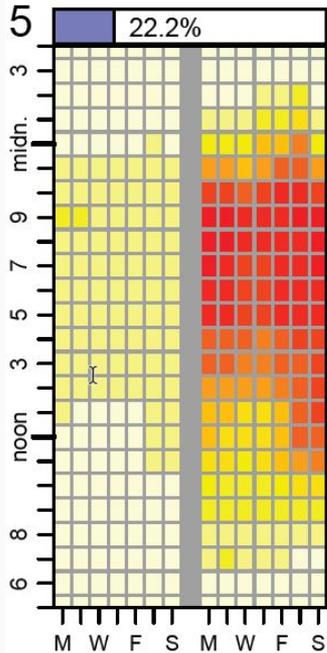
“Clustering Anonymized Mobile Call Detail Records to Find Usage Groups”

- Call Detail Records (CDRs) enthalten Zeiten jedes Anrufs und jeder SMS sowie Standort des Telefons
- Mittels eines Clustering-Algorithmus werden die CDRs auf Muster untersucht
- Nutzer, welche dieselben Muster aufweisen werden zu sozialen Gruppen zusammengefasst
- Aufgrund bestimmter Mustereigenschaften kann identifiziert werden um welche sozialen Gruppen es sich handelt



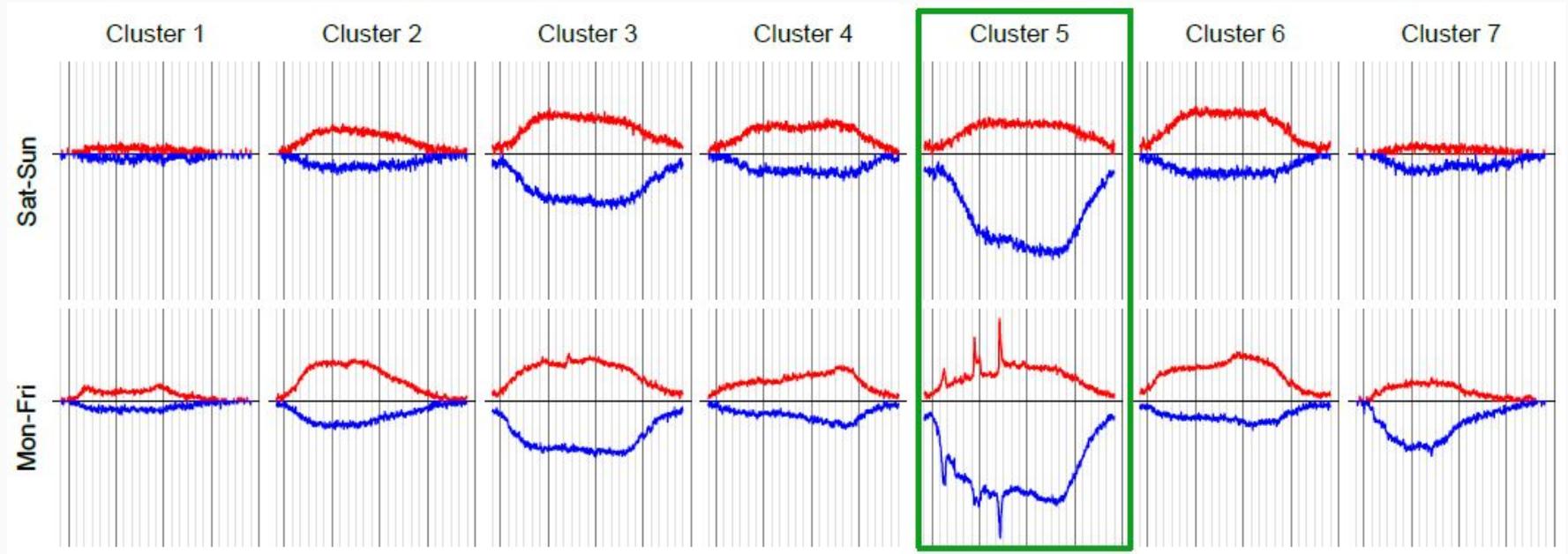


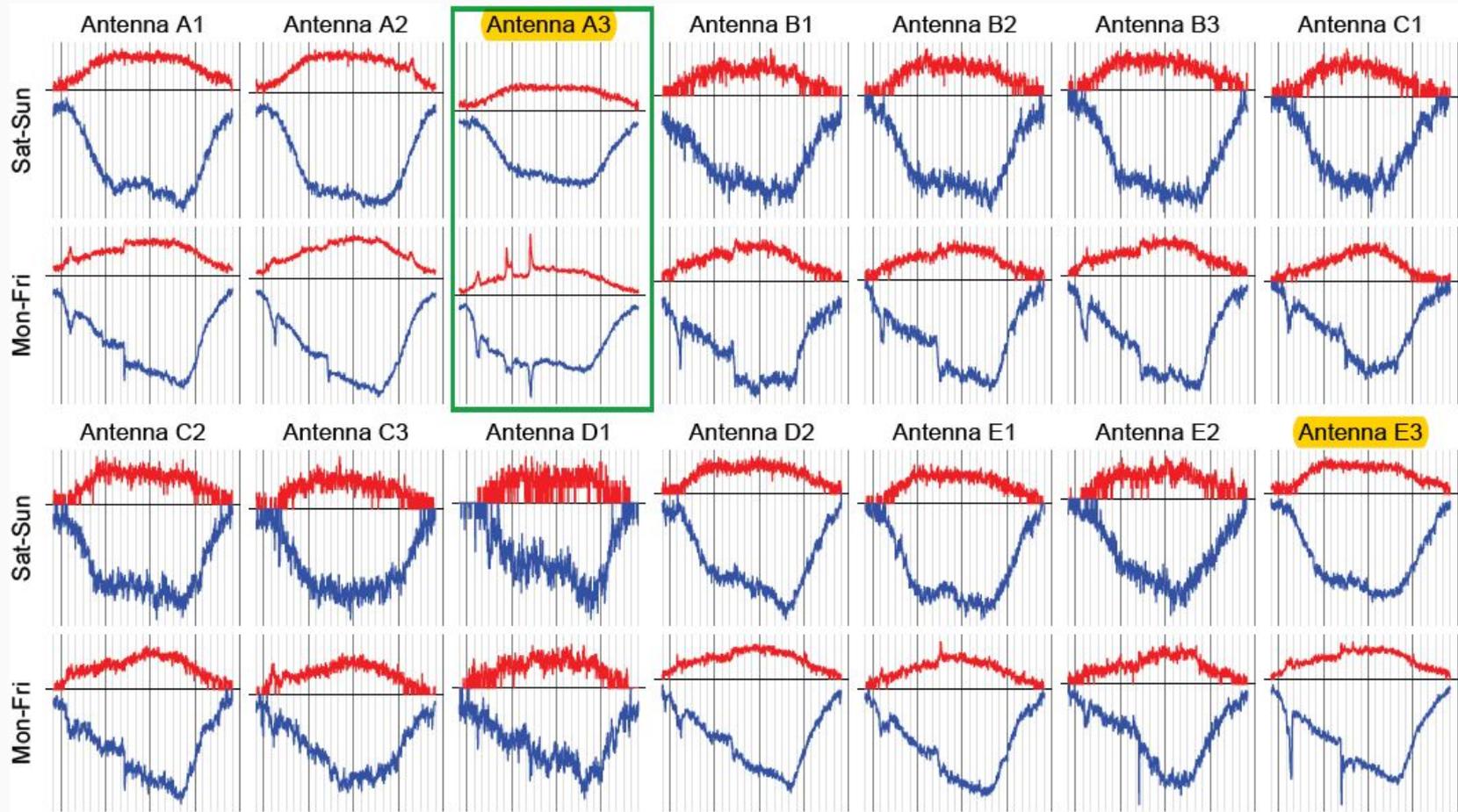
Interpretation der Cluster



- Hohe SMS-Nutzung zwischen 15 Uhr - 22 Uhr innerhalb der Woche
 - Wochenende ebenfalls hohe SMS-Nutzung, allerdings schon früher und auch länger
 - Telefonate in denselben Zeiten, aber deutlich weniger
- Vermutung: High school / Middle school Schüler

Clusteraktivität an der High School Antenne



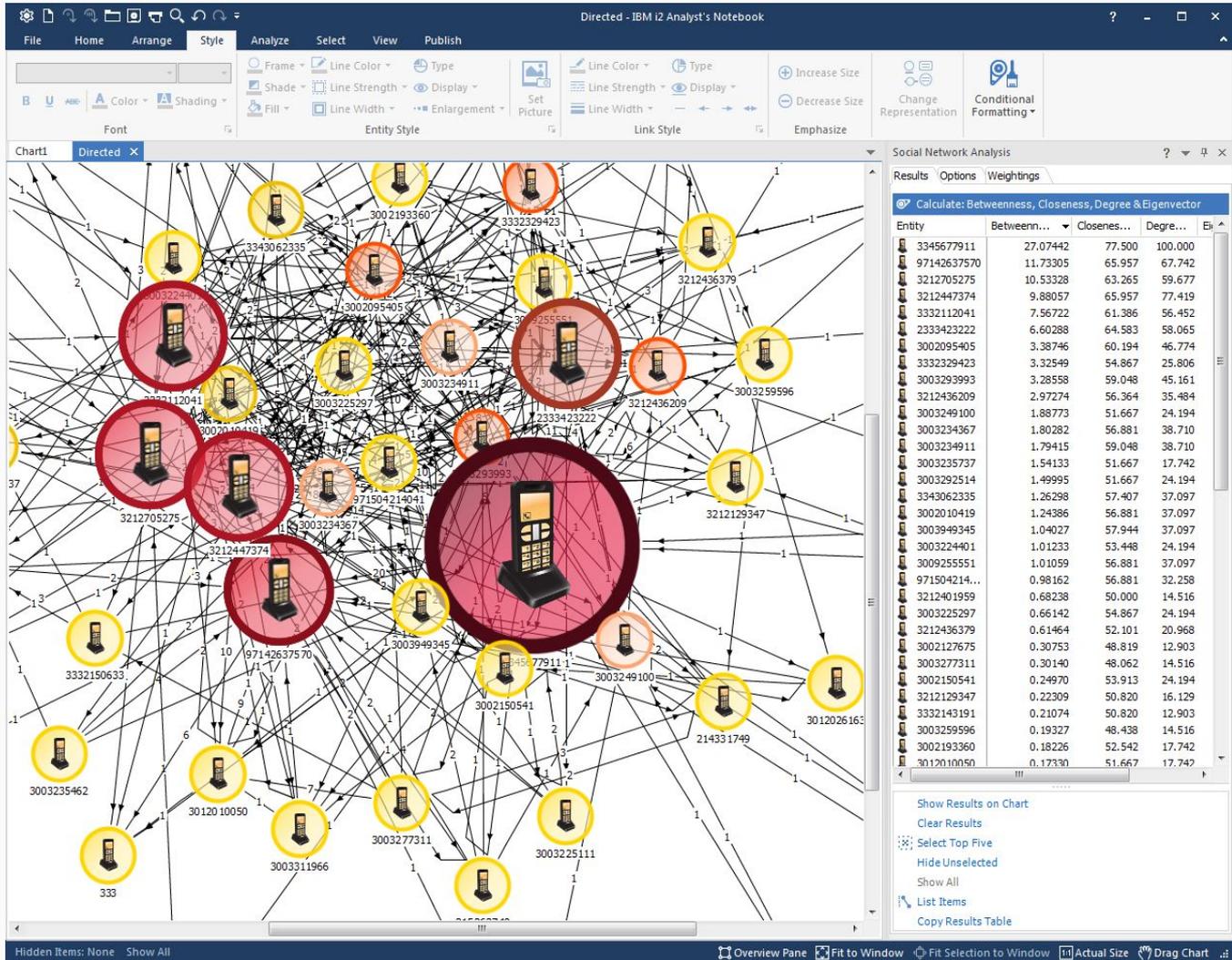


Nutzung der Erkenntnisse aus Metadaten

- Entwicklung von Anwendungen und Services unter Berücksichtigung von Nutzungsmustern aus Mobilfunkdaten
- Sammeln von Wissen über Kunden
- Planung von Marketingstrategien, Einsatz maßgeschneiderter Services
- Verbesserung der Kundenbindung, gezielte Werbung
- Anpassung von Online-Diensten an die psychografischen Eigenschaften der Kunden

Software

- IBM i2 Analyst's Notebook
 - Visuelles Analysetool zur Gewinnung von neuen Erkenntnissen aus aggregierten Daten
 - Zusammentragen von Daten aus unterschiedlichen Quellen
 - Beinhaltet zahlreiche Visualisierungen



Esri Example - IBM i2 Analyst's Notebook

File Home Arrange Style Analyze Select View Publish

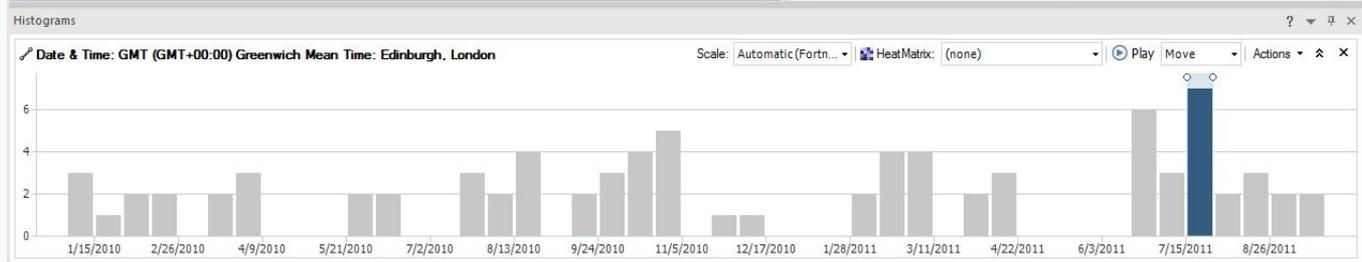
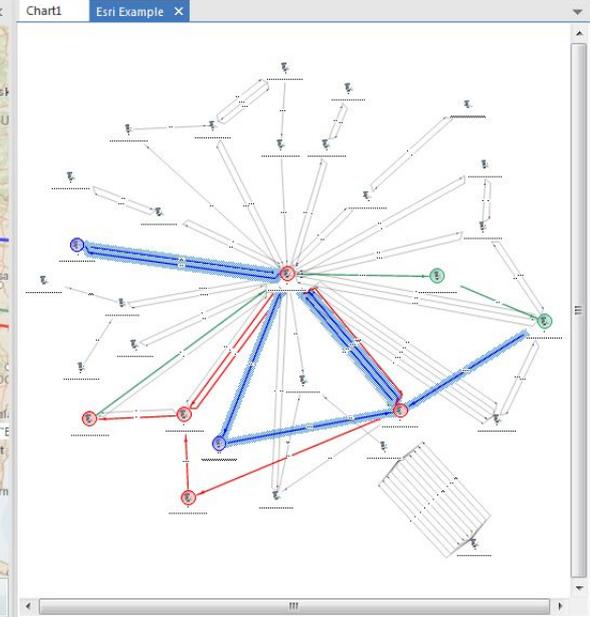
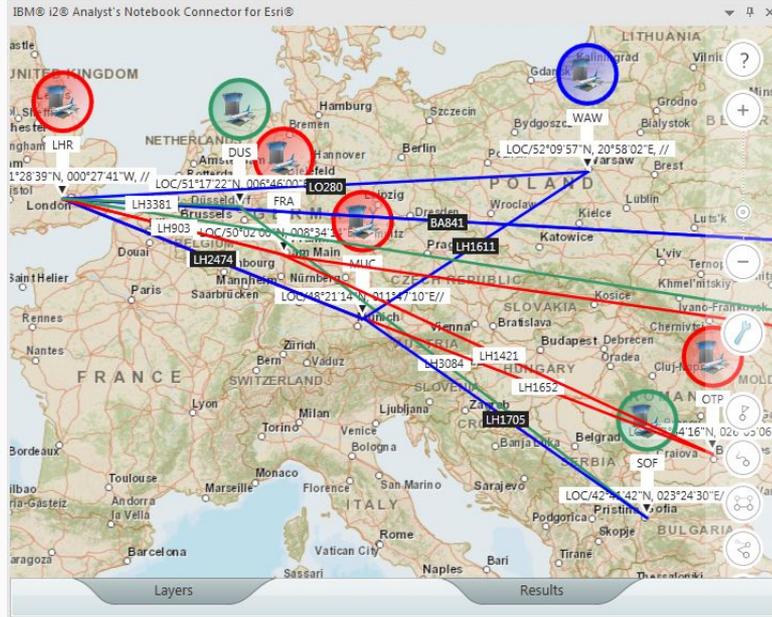
Hide Selected Hide Unselected Show and Hide Show and Hide Items

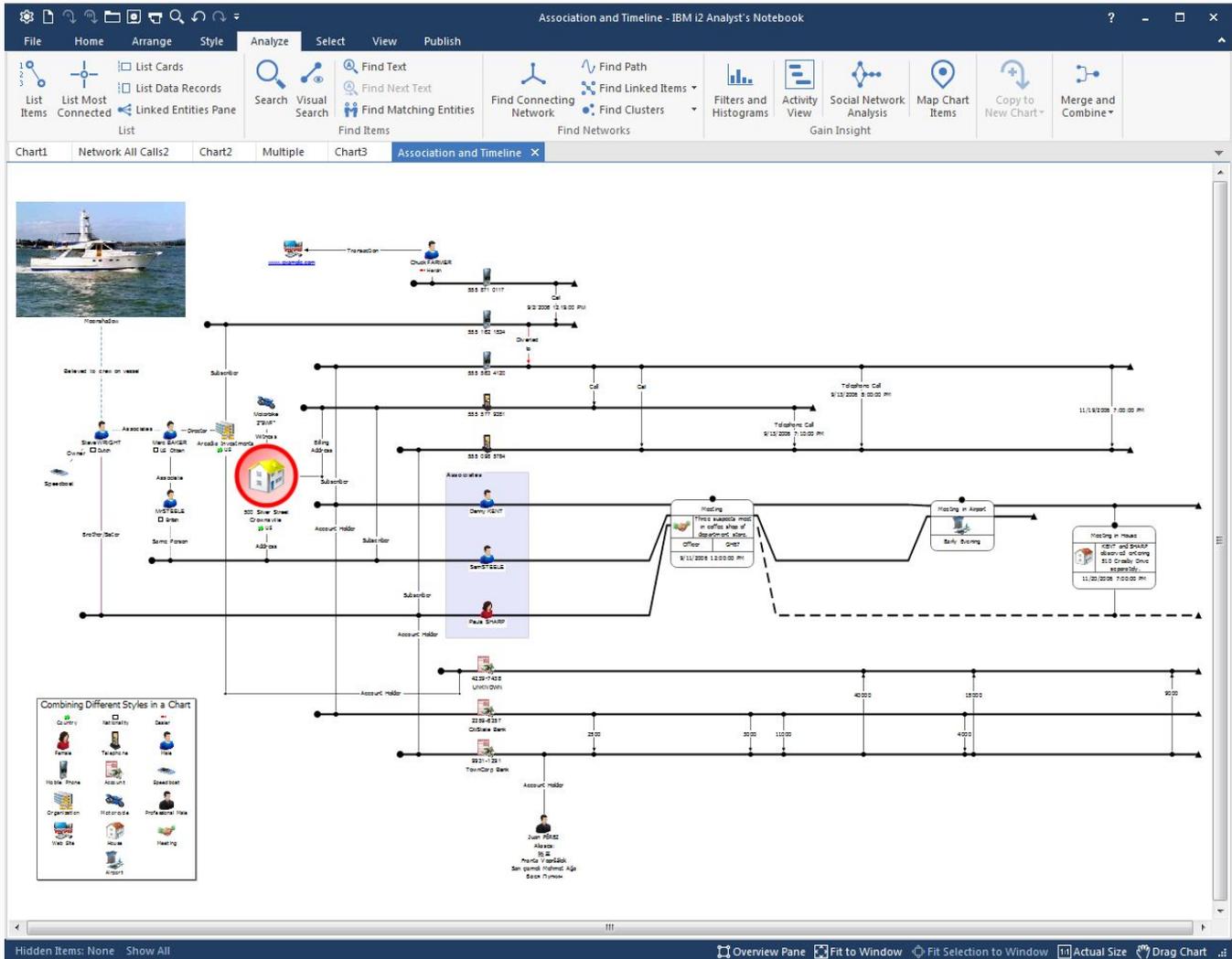
Reveal Show All Fit to Window Fit Selection to Window Actual Size Drag Chart Zoom In Zoom Out Zoom to Area Overview Pane

Page Boundaries Gridlines Time Bar

Infotips More Panes View Multiple Split New Window Full Screen

Show Window

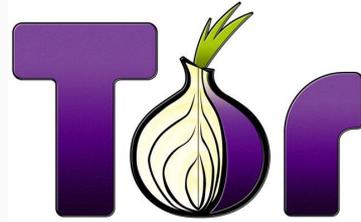




Wie kann man sich schützen?

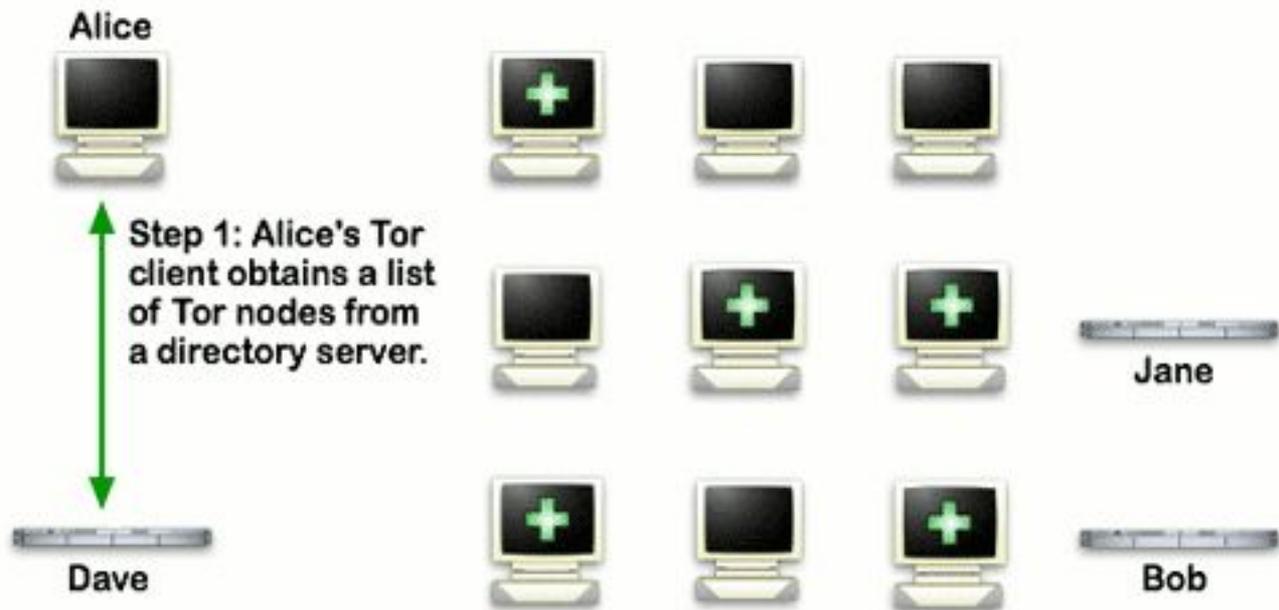
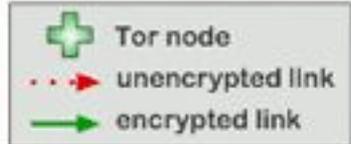
- Verzicht auf Facebook, Google, ...
- Anonymer Modus des Browsers
- Nutzung von Werbeblockern
- Verzicht auf Smartphone, Nutzung von Prepaid-SIMs

Tor

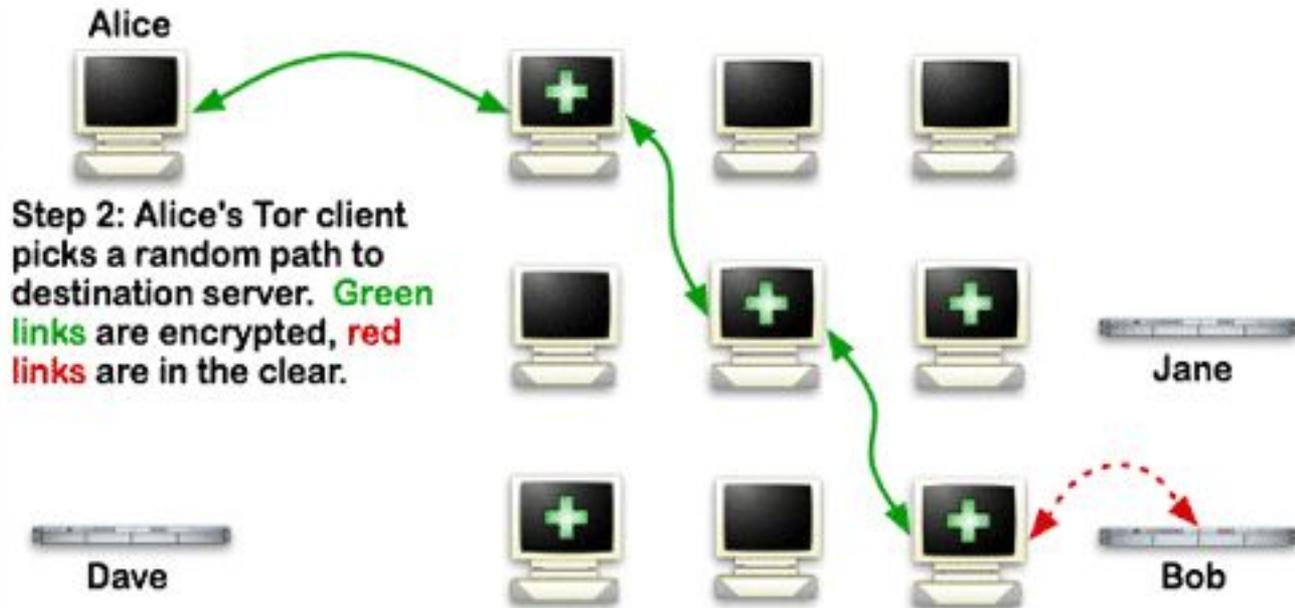
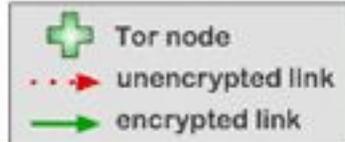


- Nutzung des Tor-Netzwerks
- Statt direkter Verbindung, Verbindung über Knotenpunkte
- Stand 2017: 7.300 Tor-Knoten mit Datenübertragungsrate von 72 Gbit/s

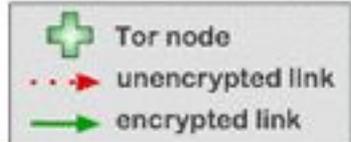
How Tor Works: 1



How Tor Works: 2



How Tor Works: 3



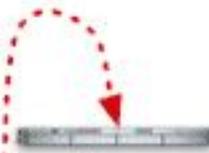
Alice



Step 3: If at a later time, the user visits another site, Alice's tor client selects a second random path. Again, **green links** are encrypted, **red links** are in the clear.



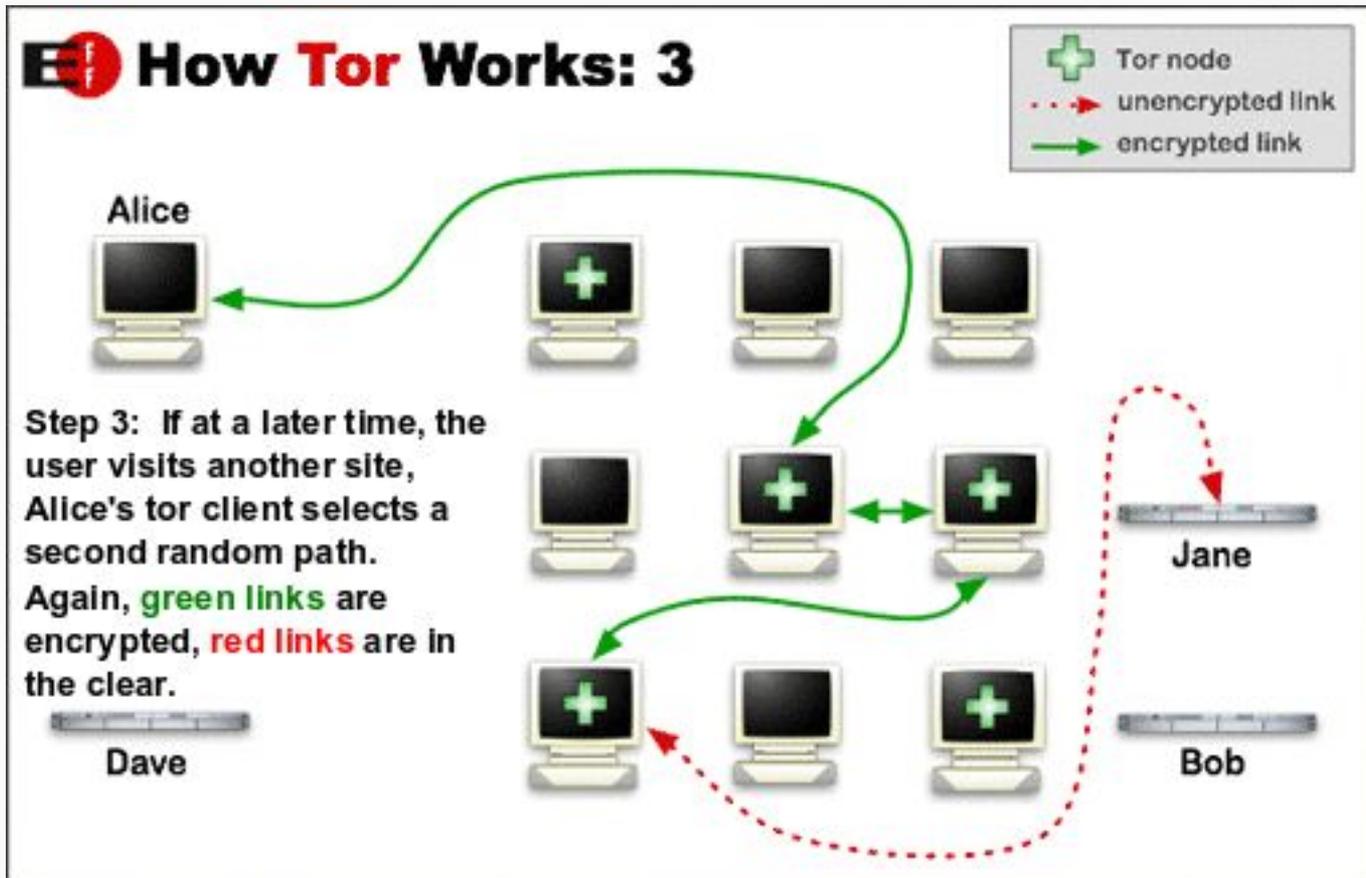
Dave



Jane



Bob



Quellen

- <https://de.wiktionary.org/w/index.php?title=Spur&oldid=5509457>
- <https://www.boell.de/de/2014/07/22/was-sind-eigentlich-metadaten>
- <http://www.zeit.de/datenschutz/malte-spitz-vorratsdaten>
- <https://netzpolitik.org/2012/vorratsdatenspeicherung-visualisiert-was-verbindungsdaten-alles-verraten>
- http://www.research.att.com/export/sites/att_labs/techdocs/TD_100397.pdf
- <https://de.wikipedia.org/w/index.php?title=AT%26T&oldid=166095022>
- http://www.ic.unicamp.br/~oliveira/doc/CHI2011-WIP_Towards-a-psychographic-user-model-from-mobile-phone-usage.pdf
- <https://www.ibm.com/de-en/marketplace/analysts-notebook>
- <https://netzpolitik.org/2014/sicherheitszusammenarbeit-mit-aegypten-und-tunesien-verschlossene-augen-vor-menschenrechtsverletzungen/>
- <http://www.zeit.de/digital/datenschutz/2014-05/daten-leben-sparsamkeit>
- <https://www.torproject.org/about/overview.html.en>