



Software-based data collection - Theorie

HIIG Methodenworkshop, 30. Mai 2013.

Dr. Carolin Gerlitz

University of Amsterdam/Goldsmiths, University of London

Digital Research Methods

- Spezielle Perspektive auf software-basierte Datenerhebung: Digital Research Methods (Rogers 2013).

digital methods initiative
dmi home

Course

- The Link
- The Website
- The Engine
- The Spheres
- The Webs
- Post-demographics
- Networked Content

Digital Methods

- Summer 2011
- About
- FAQ
- Course
- Tools
- Projects by Theme
- Research Protocols
- Summer School
- Winter School
- Papers and Publications
- Blog

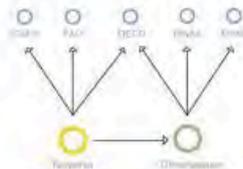
Course | Tools | Projects | About | FAQ

[The Link](#) | [The Website](#) | [The Engine](#) | [The Spheres](#) | [The Webs](#)

Welcome to the Digital Methods course, which is a focus on how to study digital media. It consists of seven units with digital research protocols, dedicated to how *else* links, Websites, engines and other digital environments are opposed to importing standard methods from the social sciences. Short literature reviews are followed by practical projects.

Course Overview

[Unit 1: The Link](#)



There are at least three dominant approaches to studying hyperlinks, hypertext theory (Landow, 1994), small world and path theory (Watts, 1999), and associational sociology (Park and Thelwall, 2003). To literary theorists of hypertext, sets of hyperlinks form a multitude of distinct pathways through text. The surfer, or clicking text navigator, may be said to author a story by choosing routes (multiple clicks) through the text (Elmer, 2001). Thus the story told through link navigation is of interest. For small world theorists, the links that form paths show distance between actors. Social network analysts use pathway thought, and zoom in on how the ties, uni-directional or bi-directional, position actors (Krebs, 2004). There is a special vocabulary that has been developed to characterize an actor's position, especially an actor's centrality, within a network. For example, an actor is 'highly between' if there is a high probability that other actors must pass through him to reach each other. To associational sociologists, as least as it's described here, links matter for a different reason. As with social network analysis, the interest is in actor positioning, but not necessarily in terms of distance from one another, or the means by which an actor may be reached through 'networking.' Rather, ties are reputational

“ontological distinction between the natively digital and the digitized, that is, the objects, content, devices and environments that are ‘born’ in the new medium, as opposed to those that have ‘migrated’ to it.”
(Rogers 2009, 1)

Digital Research Methods

- Wie können Medien und ihre Eigenschaften für die Forschung genutzt werden bzw. welche Probleme stellen sie dar?
- Wieviel Daten sind genug?
- Einsichten in Medien- oder Themendynamik?



Schedule

Intro & Kontext

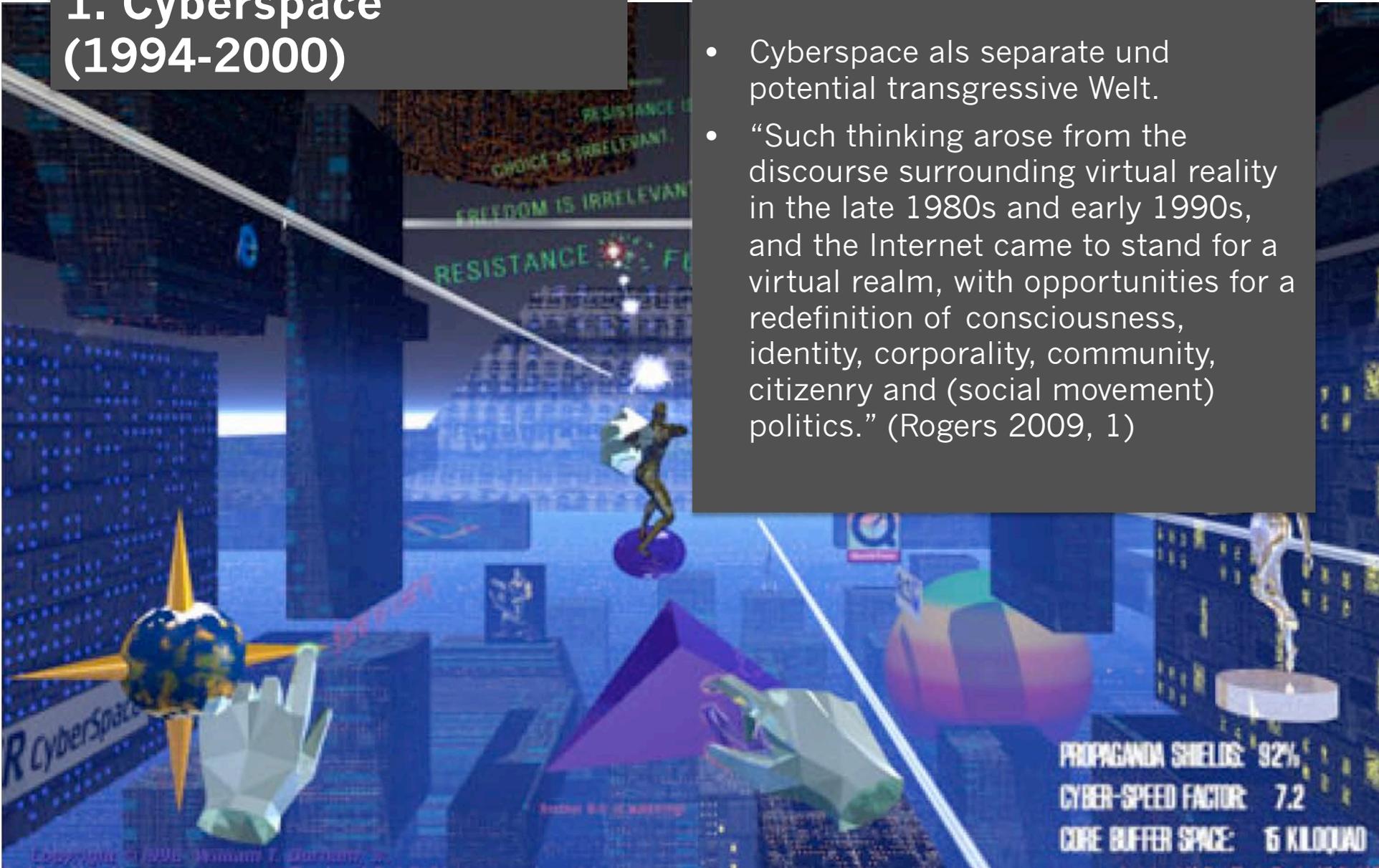
1. Aneignung/Repurposing
2. Big & Small Data
3. Medien vs. Gesellschaft

Intro & Kontext



1. Cyberspace (1994-2000)

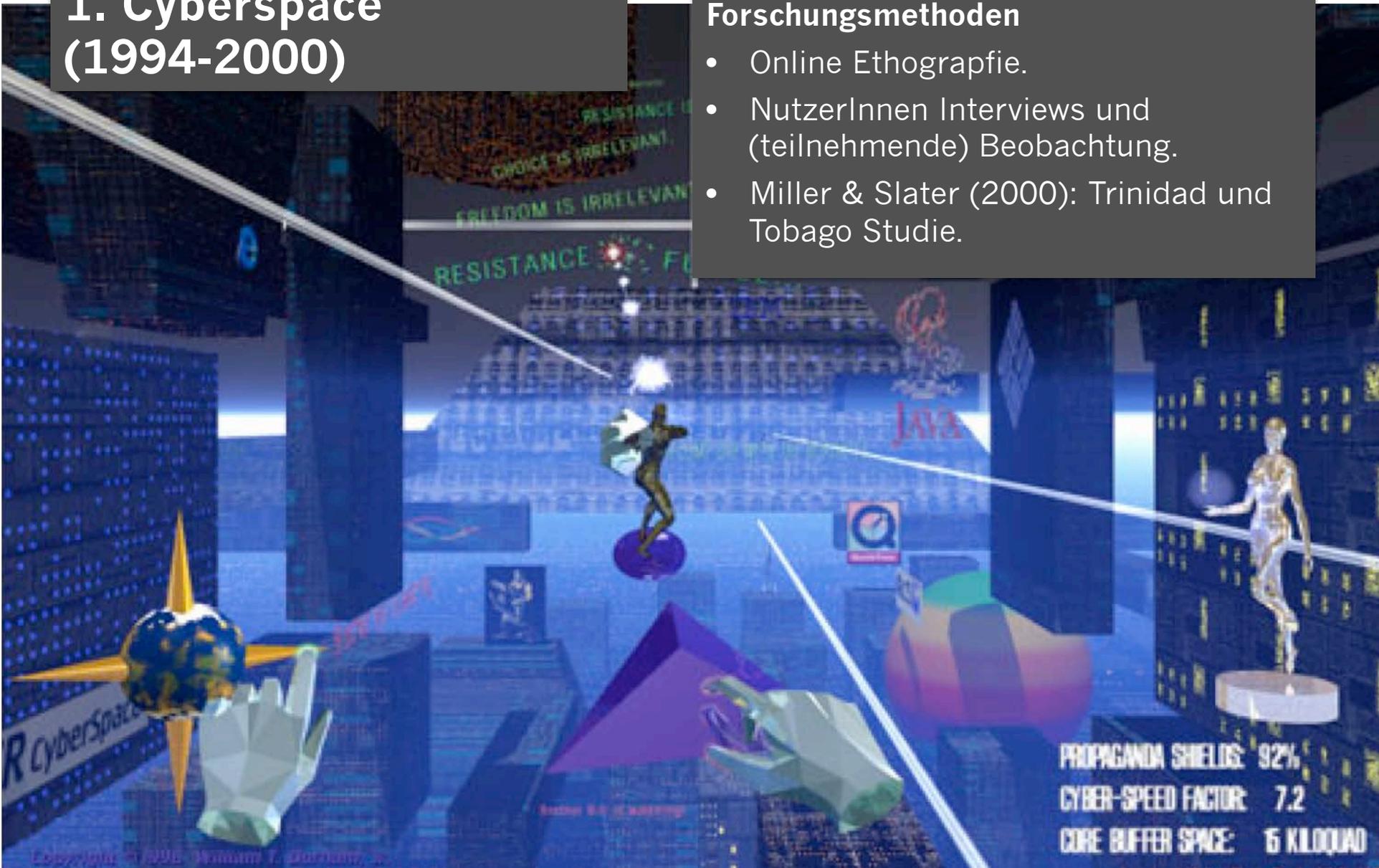
- Cyberspace als separate und potential transgressive Welt.
- “Such thinking arose from the discourse surrounding virtual reality in the late 1980s and early 1990s, and the Internet came to stand for a virtual realm, with opportunities for a redefinition of consciousness, identity, corporality, community, citizenry and (social movement) politics.” (Rogers 2009, 1)



1. Cyberspace (1994-2000)

Forschungsmethoden

- Online Ethnografie.
- NutzerInnen Interviews und (teilnehmende) Beobachtung.
- Miller & Slater (2000): Trinidad und Tobago Studie.

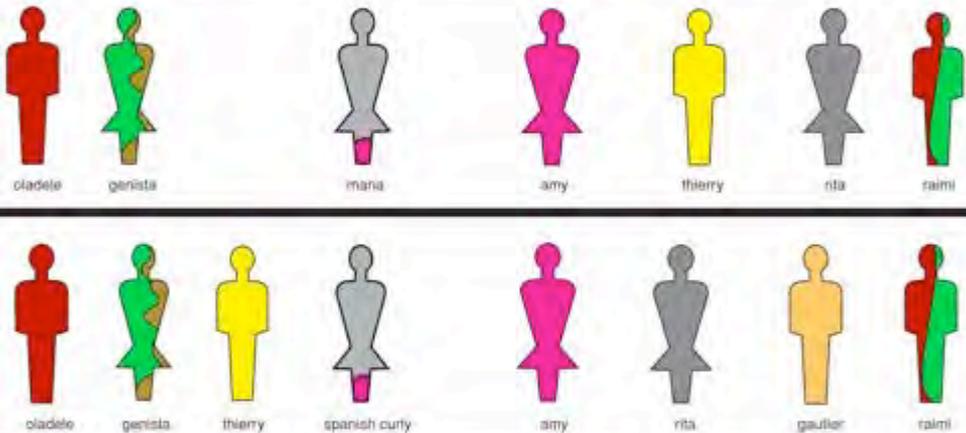


2. Virtual Society (2000-2007)

Netnography

Doing Ethnographic Research Online

Robert V. Kozinets



- **Online und offline Welten sind verknüpft**, nicht getrennt.
- **Methodenmigration**: Wie können bestehende soziologische oder ethnologische Methoden online genutzt werden (Hine 2004, Kozinets 1997)?



3. Digital Methods (from 2007)

Google Autocomplete

Input

Input words to retrieve autocomplete for, one per line

Select one or more languages:

- Armenian
- Belarusian
- Bulgarian
- Catalan
- Chinese (Simplified)
- Chinese (Traditional)
- Croatian
- Czech
- Danish
- Dutch

Select one or more countries:

- Monaco
- Mongolia
- Montserrat
- Morocco
- Mozambique
- Myanmar [Burma]
- Namibia
- Nauru
- Nepal
- Netherlands

Autocomplete

- Online medien sind mit dem offline Leben **verbunden, aber auch spezifisch.**
- Fokus auf Spezifität von Medien und Methoden.
- Keine Methodenmigration sondern **medien-spezifische Methoden** (Rogers 2009).
- Langfristige Frage: Sind digitale Forschungsmethoden heute tatsächlich rein spezifisch?

Medium-Spezifizität

- Tools auf Grundlage von Devices & Plattformen.
- Tools auf Grundlage von Links, Tags, Kommentaren, Meta-Daten, Likes etc.

digital methods initiative
wiki

Course

- The Link
- The Website
- The Engine
- The Spheres
- The Webs
- Post-demographics
- Networked Content

Digital Methods

- Summer 2011
- About
- FAQ
- Course
- Tools
- Projects by Theme
- Research Protocols
- Summer School
- Winter School
- Papers and Publications
- Blog

DMI Tools

Media Analysis: Media Monitoring | Mapping | Clouding
Data Treatment: Data Collection | Data Analysis | Informatics
Natively Digital: The Link | The URL | The Tag | The Domain
Device Centric: Google | Google Images | Google News
IssueCrawler | Twitter | Facebook
Spherical: Web Sphere | News Sphere | Blogosphere |

2 Ziele

- 1. Nutze Medium-Spezifizität um Medienformen zu verstehen.
- 2. Nutze Medium-Spezifizität um soziale Fragen/Themen zu verstehen.

Actor Profiler



[Launch tool](#) [Instructions & Scenarios of Use](#)

The Actor Profiler works in tandem with the Issue Crawler. It calculates the top ten actors on an Issue Crawler map, and profiles each of them. The profile consists of each actor's inlinks a...

Censorship Explorer

IRAN	response code for request
www.collegehumor.com/	Block
www.dailymotion.com/	OK
www.gayhealth.com/	OK
www.ancientbookaccess.net/	OK
www.webhosting.com/	OK
www.supernews.com/	OK
madure.digit	OK
news.bbc.co.uk/	OK
	HTTP/1.0 503 Service Unavailable

[Launch tool](#) [Instructions & Scenarios of Use](#)

Check whether a URL is censored in a particular country by using proxies located around the world.

Compare Lists

http://www.	http://www.
http://www.	http://www.

[Launch tool](#) [Instructions & Scenarios of Use](#)

Compare two lists of URLs for their commonalities and differences.

Convert Issuecrawler to Navicrawler



[Launch tool](#) [Instructions & Scenarios of Use](#)

Convert an Issuecrawler XML file into the WXSf format of the Navicrawler file

Medium-Spezifität

- **McLuhan** (1964): Medien sind spezifisch im Ansprechen von Sinnen.
- **Williams** (1974): Medien haben spezifische Formen.
- **Hayles** (2004): Medien haben eine spezielle Materialität.
- **Esposito** (2004): De- & Rekombination von medien-spezifischen Elementen in neue Beziehungen.

facebook DEVELOPERS Documentation Support blog Apps Search

Social Plugins

Core Concepts > Social Plugins

Social plugins let you see what your friends have liked, commented on or shared.

- Like Button**
The Like button lets users share pages from your site with their friends.
- Send Button**
The Send Button allows your users to easily send content to their friends.
- Comments**
The Comments plugin lets users comment on your content.
- Activity Feed**
The Activity Feed plugin shows users what their friends have liked, commented on or shared.

Your Like Button plugin code:

Iframe

```
<iframe src="http://www.facebook.com/plugins/like.php?layout=button_count&
amp;show_faces=true&width=300&action=like&font=verdana&
amp;colorscheme=light" scrolling="no" frameborder="0" allowTransparency="true"
style="border:none; overflow:hidden; width:300px; height:px"></iframe>
```

XFBML

```
<fb:like layout="button_count" show_faces="true" width="300" action="like"
font="verdana" colorscheme="light"></fb:like>
```

XFBML is more flexible than iframes, but requires you use the JavaScript SDK.

Done

Aneignung/Repurposing



Aneignung I

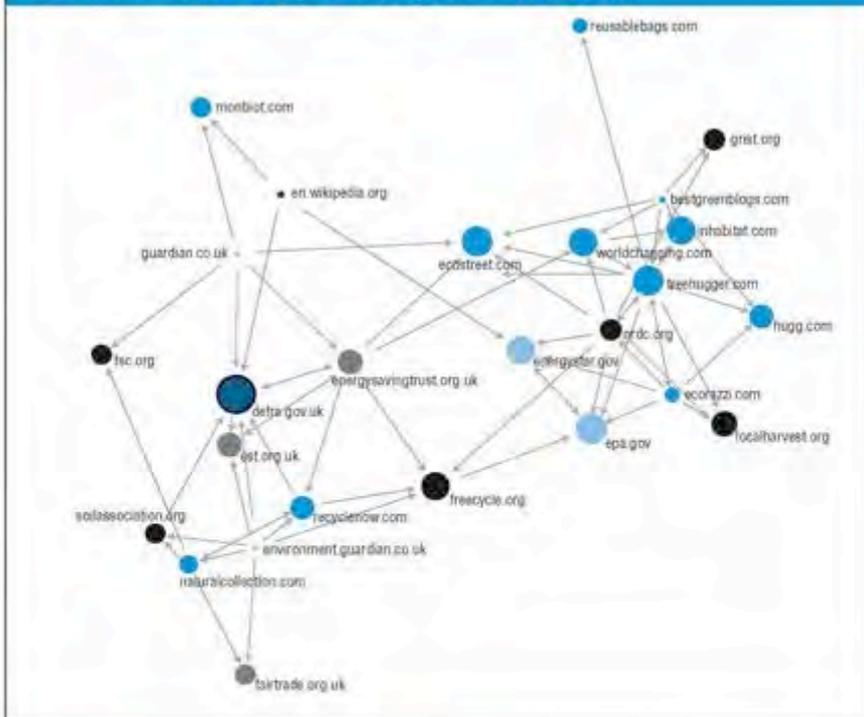


- Online Medien ermöglichen Interaktion durch **vordefinierte Handlungsoptionen**.
- **Grammatik der Handlung** (Agre 1994) produzieren standardisierte Handlungen sowie **Datenformate**.
- Facebook: Like, teilen, kommentieren.
- Twitter: Tweets, @replies, retweets, favs...
- Ausgangspunkt für Erhebung von vorstrukturierten Daten.



Aneignung I

Figure 1. Issue network disclosed by green home blogs. Issuercrawler, March 2008



Map details

Author: Noortje Marres
Email: marres@delta.nl
Crawl start: 4 Mar 2008 - 05:10
Crawl end: 4 Mar 2008 - 08:05
Privilege starting points: off
Analysis mode: site
Iteration: 2
Depth: 2
Node count: 25

Map generated from Issuercrawler.net by the Geocomp.org Foundation, Amsterdam.

Legend

● (.com)
● (.gov.uk)
● (.org)
● (.org.uk)
● (.gov)
● (.co.uk)

- Beispiel: Issuercrawler untersucht Linkbeziehungen zwischen Webseiten.
- Ausgangspunkt: Link als Assoziations- & Beziehungsmarker.
- Etablierung von Website-Netzwerken (Rogers 2002, Marres 2007).

Aneignung II

- Medien haben **eingebaute Sortier- und Analysefähigkeiten**.
- Google sortiert Suchergebnisse anhand von „Relevanz“ oder „Aktualität“.
- Twitter erlaubt Suche von Hashtags, @replies, Stickworten.
- Beruhen auf (unbekannten) Algorithmen wie Page oder Edgerank.



EdgeRank

$$\sum U_e W_e D_e$$

Affinity Score between viewing user & edge creator

Weight for this edge type (create, comment, like, tag)

Time decay factor based on how long ago the edge was created

Analytische Fähigkeiten von Medien

1. Forschung versucht Algorithmen zu verstehen.
2. Forschung nutzt die analytischen Fähigkeiten von Algorithmen.

- Nicht Entschlüsselung von Blackbox-Algorithmen sondern Aneignung.
- Digital Methods Initiative nutzt online Such- und Sortierungsmechanismen als Grundlage von Tools.
- **„Repurposing dominant devices“** (Rogers 2009).

Repurposing Tools

& Sce

does

Tools auf Grundlage von bestehenden online Devices wie Suchmaschinen, APIs, Archiven, Browser Plugins.

Wikipedia Entry Check



retrieved <http://en.wiki>
greenpeace not found
weather found 18 times
polar bear not found
sun found 25 times
ice found 62 times

[Launch tool](#)

[Instructions & Scenarios of Use](#)

This tool checks if the issues exist as an organization is mentioned on that page.

Wikipedia History Flow Companion



36,620 bytes	(Doc: Link FA -'s')
36,936 bytes	(Reverted to revision 2906)
36,806 bytes	(robot Adding: el:Κλασσικ...

[Launch tool](#)

[Instructions & Scenarios of Use](#)

This script allows you to specify the range of Wikipedia revisions for use with IBM's History Flow visualization. Normally, History flow only allows you to input a maximum of 100 edits.

Wikipedia Network Analysis



[Launch tool](#)

[Instructions & Scenarios of Use](#)

Find a hyperlink network around a Wikipedia topic (see [here](#) for more information, and <http://thepoliticsofsystems.net/2011/02/21/visualizing-wikipedia-on-data-vi...>

Yahoo Inlink Scraper



```
http://www.nytimes.com/2011/02/21/nyregion/...
```

[Launch tool](#)

[Instructions & Scenarios of Use](#)

Retrieves all inlinks to a site, according to Yahoo Site Explorer.

YouTube Video Discovery



<http://youtube.com/w>
<http://youtube.com/w>
<http://youtube.com/w>
<http://youtube.com/w>

[Launch tool](#)

[Instructions & Scenarios of Use](#)

Analyzes a results file from the Google Scraper to discover, count, and rank YouTube and Google Video links in the descriptions. (It may also be used for Ikbis and other video compilati...

Google Scraper

Past Jobs | Output

Google Scraper

The GoogleScraper queries Google and makes the results available for further analysis. In the top text box, place URLs. In the bottom text box, place key words. Google will be asked if each keyword occurs in each URL.

Results are displayed as a tag cloud and an html table. They also are written to a text file.

Currently the GoogleScraper only works in Firefox with the [DMI toolbar](#) installed.

Method & Visualization - How To's

Source clouds

Show the partisanship or commitment of sources to issues. The cloud displays sources, each resized according to the number of mentions of a particular issue.

Issue clouds

Show the issue commitment or partisanship of a single source or multiple sources. The cloud displays issues, each resized according to the number of mentions by one or more sources.

Enter URLs:

(Or harvest URLs from a body of text. See [harvester functionality](#).)

Only query discrete sites (recommended).

Checking this option chops all URLs to their host name and then removes duplicate hosts from the URL input box. For example, if you enter <http://www.un.org> and <http://www.un.org/issues/m-child.html>, when this option is checked only one query is performed for the site <http://www.un.org>.

Enter key words, one per line:

(You can perform normal [Google-style queries](#). Further considerations are listed [on our wiki](#).)

Options:

Number of results per query (max 1000)

Google Scraper

- Einstellen von Zeitintervallen.
- Wahl zwischen globaler (.com) oder lokalen Google Versionen.
- Suchbegriffposition.

Options:

Number of results per query (max 1000)

Name your result file

Tag cloud for 'site:' queries generate do NOT generate

Advanced options

Google Local version

Language

Region

Last updated

Or select a custom daterange:

Terms appearing

File type

Usage rights

Google Scraper

- Entwickelt für die Analyse von **Kontroversen und die Resonanz von Themen** in verschiedenen Webseiten.
- “Lippmannian Device”: Parteilichkeit und Bias von Quellen und Webseiten.

digital methods initiative

wiki

Digital Methods

- Summer School 2013
- About
- People
- FAQ
- Tools
- Projects by Theme
- Research Protocols
- Summer School
- Winter School
- Papers and Publications
- Blog

Course

- The Link
- The Website
- The Engine
- The Spheres
- The Webs
- Post-demographics
- Networked Content

Googlescraper (Lippman

google

Launch tool

Batch queries Google. Query the resonance of a particular term, or a series of terms, in a set of Websites.

berkeley.edu (191) (200)
etcgroup.org (126) (200)
jcvl.org (19) (200)
nature.com (200)
nytimes.com (200)
sciencemag.org (200)

Instructions

The Googlescraper (also known as the [Lippmannian Device](#)) queries Google and makes the results available for further analysis. In the top text box, place the [source set](#), in this case a list of URLs. In the bottom text box, place key words. Google will be asked if each keyword occurs in each URL. Results are displayed as a tag cloud and an html table. They also are written to a text file which you can access at the bottom or through previous results.

Harvester feature: In the top box, you may also place a combination of URLs and text, and the URLs will be fetched out of the text and queried for the key words placed in the bottom box. Detailed instructions of use and use cases are [available](#).

Sample project

The Googlescraper can be used for a number of specific research projects, including [censorship research](#), and [source distance research](#). The most common use of the tool is researching the presence as well as the ranking of particular sources within Google engine results. A sample project is this tag cloud, which visually presents unique hosts from top 100 URLs returned from the query of "synthetic biology." The hosts are sized by occurrences of "Venter" on each site. The method for this project: 1) Search google for "synthetic biology". 2) Paste top 100 results in the top box, and enter Venter in the bottom box.

Case: Klimawandel Skeptiker

Persons

- S. Fred Singer
- Robert Balling
- Sallie Baliunas
- Patrick Michaels
- Richard Lindzen
- Steven Milloy
- Timothy Ball
- Paul Driessen
- Willie Soon
- Sherwood B. Idso
- Frederick Seitz

Organisations

- American Enterprise Institute
- American Legislative Exchange Council
- Center for Science and Public Policy
- Committee for a Constructive Tomorrow
- Competitive Enterprise Institute
- Frontiers of Freedom
- Marshall Institute
- Heartland Institute
- Tech Central Station

- **Forschungsfrage:** Wie prominent sind Klimawandel Skeptiker im Web und auf relevanten Seiten zum Thema Klimawandel?
- **Schritt 1:** Liste relevanter Skeptiker durch Triangulation bestehender Listen.

Case: Klimawandel Skeptiker

Google Scraper

The GoogleScraper queries Google and makes the results available for further analysis. In the top text box, place URLs. In the bottom text box, place key words. Google will be asked if each keyword occurs in each URL.

Results are displayed as a tag cloud and an html table. They also are written to a text file.

Currently the GoogleScraper only works in Firefox with the [DMI toolbar](#) installed.

Method & Visualization - How To's

Source clouds

Show the partisanship or commitment of sources to issues. The cloud displays sources, each resized according to the number of mentions of a particular issue.

Issue clouds

Show the issue commitment or partisanship of a single source or multiple sources. The cloud displays issues, each resized according to the number of mentions by one or more sources.

- **Schritt 2:** Top 100 Google Ergebnisse für "Climate Change".
- **Schritt 3:** Ergebnisse in URL Liste des Google Scrapers einfügen und Skeptiker Namen als Suchbegriffe nutzen.
- **Schritt 4:** Häufigkeiten von Skeptikern in Suchergebnisse als Tagcloud visualisieren.

The screenshot shows the GoogleScraper web interface. It has a light beige background. At the top left, there is a section titled "Enter URLs:" with a subtext "(Or harvest URLs from a body of text. See [here](#))". Below this is a large empty text box. To the right of this box is a checkbox labeled "Only query discrete sites (recommended)." with a checked mark. Below the checkbox is a paragraph of text: "Checking this option chops all URLs to their host name and then removes duplicate hosts from the URL input box. For example, if you enter <http://www.un.org> and <http://www.un.org/issues/m-child.html>, when this option is checked only one query is performed for the site <http://www.un.org>." Below this is another section titled "Enter key words, one per line:" with a subtext "(You can perform normal [Google-style queries](#). Further considerations are listed [on our wiki](#).)". Below this is a large text box containing the text "climate change". At the bottom left, there is a section titled "Options:" with a label "Number of results per query (max 1000)" and a text input field containing the number "100".

Climate Change Sceptics on the Web (Sallie Baliunas)

Research Question_To what extent are climate change 'skeptics' present in the climate change spaces on the Web?

Findings_There is distance between the skeptics and the top of the search engine returns.

Case: Klimawandel Skeptiker

epa.gov (0) bbc.co.uk (0) defra.gov.uk (0) unep.org (0) bom.gov.au (0) ipcc.ch (0) pewclimate.org (0)
david Suzuki.org (0) panda.org (0) mfe.govt.nz (0) ec.gc.ca (0) exploratorium.edu (0) climatechange.com.au (0)
greenpeace.org (0) climatechallenge.gov.uk (0) guardian.co.uk (0) iisd.org (0) g8.gov.uk (0) campaigncc.org (0)
foe.co.uk (0) state.gov (0) scidev.net (0) eea.europa.eu (0) whoi.edu (0) cbc.ca (0) energy.gov (0)
marshall.org (6) **climateark.org (2)** un.org (0) dar.csiro.au (0) theglobeandmail.com (0)
acfonline.org.au (0) gcric.org (0) nature.com (0) grida.no (0) nature.org (0) ecokids.ca (0) royalsoc.ac.uk (0)
climatechangecentral.com (0) iea.org (0) ecn.ac.uk (0) ecy.wa.gov (0) worldwildlife.org (0)

realclimate.org (55)

faqs.org (0) metoffice.gov.uk (0) open2.net (0) scienceagogo.com (0) eldis.org (0) ft.com (0) who.int (0)
climatecrisis.net (0) Itscotland.org.uk (0) abc.net.au (0) climatechange.ca.gov (0) envirolink.org (0) mofa.go.jp (0)
sourcewatch.org (0) iucn.org (0) dfat.gov.au (0) ncdc.noaa.gov (0) climatescience.gov (0)
climatechangecollege.org (0) ciel.org (0) ucar.edu (0)

Source_google.com

Query_ "Sallie Baliunas"

Method_Search for query "Sallie Baliunas" in top 100. Organized in order.

Tools_Google Scraper and Tag Cloud Generator

Date_30 July 2007

Product_of the Digital Methods Initiative, dmi.media studies.nl. Analysis_by Bram Nijhof, Richard Rogers and Laura van der Vlies. Design_Annie Helmond.



CC BY-NC-SA

Climate Change Sceptics on the Web (Timothy Ball)

Research Question To what extent are climate change 'skeptics' present in the climate change spaces on the Web?

Findings There is distance between the skeptics and the top of the search engine returns.

Case: Klimawandel Skeptiker

epa.gov (0) bbc.co.uk (0) defra.gov.uk (0) unep.org (0) bom.gov.au (0) ipcc.ch (0) pewclimate.org (0)
davidsuzuki.org (0) panda.org (0) mfe.govt.nz (0) ec.gc.ca (0) exploratorium.edu (0) climatechange.com.au (0)
greenpeace.org (0) climatechallenge.gov.uk (0) guardian.co.uk (0) iisd.org (0) g8.gov.uk (0)
campaigncc.org (6) foe.co.uk (0) state.gov (0) scidev.net (0) eea.europa.eu (0) whoi.edu (0) cbc.ca (0)
energy.gov (0) **marshall.org (2)** climateark.org (0) un.org (0) dar.csiro.au (0) theglobeandmail.com (0)
acfonline.org.au (0) gcric.org (0) nature.com (0) grida.no (0) nature.org (0) ecokids.ca (0) royalsoc.ac.uk (0)
climatechangecentral.com (0) iea.org (0) ecn.ac.uk (0) ecy.wa.gov (0) worldwildlife.org (0) realclimate.org (0)
faqs.org (0) metoffice.gov.uk (0) open2.net (0) scienceagogo.com (0) eldis.org (0) ft.com (0) who.int (0)
climatecrisis.net (0) ltscotland.org.uk (0) abc.net.au (0) climatechange.ca.gov (0) envirolink.org (0) mofa.go.jp (0)
sourcewatch.org (0) iucn.org (0) dfat.gov.au (0) ncdc.noaa.gov (0) climatescience.gov (0)
climatechangecollege.org (0) ciel.org (0) ucar.edu (0)

Source_google.com
Query_ "Timothy Ball"
Method_Search for query "Timothy Ball" in top 100. Organized in order.
Tools_Google Scraper and Tag Cloud Generator
Date_30 July 2007

Product_of the Digital Methods Initiative,
dmi.mediaudies.nl. Analysis_by Bram
Nijhof, Richard Rogers and Laura van der
Vlies. Design_Ane Helmond.



Climate Change Sceptics on the Web (Paul Driessen)

Research Question To what extent are climate change 'skeptics' present in the climate change spaces on the Web?

Findings There is distance between the skeptics and the top of the search engine returns.

Case: Klimawandel Skeptiker

epa.gov (0) bbc.co.uk (0) defra.gov.uk (0) unep.org (0) bom.gov.au (0) ipcc.ch (0) pewclimate.org (0)
davidsuzuki.org (0) panda.org (0) mfe.govt.nz (0) ec.gc.ca (0) exploratorium.edu (0) climatechange.com.au (0)
greenpeace.org (0) climatechallenge.gov.uk (0) guardian.co.uk (0) iisd.org (0) g8.gov.uk (0) **campaigncc.org (1)**
foe.co.uk (0) state.gov (0) scidev.net (0) eea.europa.eu (0) whoi.edu (0) cbc.ca (0) energy.gov (0)
marshall.org (0) climateark.org (0) un.org (0) dar.csiro.au (0) theglobeandmail.com (0) acfonline.org.au (0)
gcrio.org (0) nature.com (0) grida.no (0) nature.org (0) ecokids.ca (0) royalsoc.ac.uk (0) climatechangecentral.com (0)
iea.org (0) ecn.ac.uk (0) ecy.wa.gov (0) worldwildlife.org (0) realclimate.org (0) faqs.org (0) metoffice.gov.uk (0)
open2.net (0) scienceagogo.com (0) eldis.org (0) ft.com (0) who.int (0) climatecrisis.net (0) ltscotland.org.uk (0)
abc.net.au (0) climatechange.ca.gov (0) envirolink.org (0) mofa.go.jp (0)

sourcewatch.org (28)

iucn.org (0) dfat.gov.au (0)

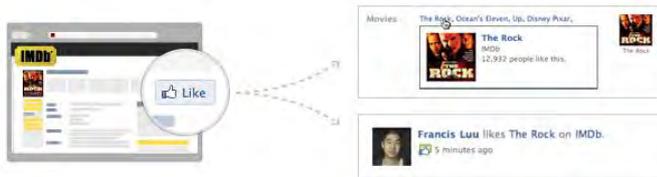
ncdc.noaa.gov (0) climatescience.gov (0) climatechangecollege.org (0) ciel.org (0) ucar.edu (0)

Source_google.com
Query_ "Paul Driessen"
Method_Search for query "Paul Driessen" in top 100. Organized in order.
Tools_Google Scraper and Tag Cloud Generator
Date_30 July 2007

Product_of the Digital Methods Initiative,
dmi.mediaudies.nl. Analysis_by Bram
Nijhof, Richard Rogers and Laura van der
Vlies. Design_Anne Helmond.



Case 2: Tracker Tracker



Tony Adam, Todd Mintz and 31 others like this. Unlike



Visiting New York City, NY 411 people like this. Be the first of your friends.

→ TripAdvisor best fares
Los Angeles to New York: only \$259
[See more flights](#)

Empire State Building Observatory Deck - Downtown View Dec 2009 (Jordanski)

All traveler photos · All traveler videos

Top-rated hotels All 420 hotels

1. Casablanca Hotel 1,200 reviews

Get travel advice from your friends

Ask Friends Replies

- Yoav Bergman
Current city is New York City
[Ask Yoav](#)
- Carlin Brody
Hometown is New York City
[Ask Carlin](#)
- Michael Marantz
Hometown is New York City
[Ask Michael](#)
- Lindsey Disinger
[Can advise about New York City](#)

Post to your Wall or send a message.
[Ask a question](#)

How does TripAdvisor know this?

- Soziale Plattformen expandieren ins Web durch **Social Plugins**.
- Facebook Like Buttons & Twitter Buttons beinhalten Cookies und können Browserverläufe dokumentieren.
- **Forschungsfrage:** Wie präsent sind Social Media Plattformen durch ihre Plugins und Widgets im externen Web?

facebook.com 4 cookies

www.facebook.com 8 cookies

Name	Content	Domain	Path	Send For	Accessible to Script	Created	Expires
_e_YpyF_3							
act							
c_user	c_user	.facebook.com	/	Any kind of connection	Yes	Tuesday, May 10, 2011 9:27:33 AM	Thursday, June 9, 2011 9:27:32 AM
e							
presence							
sct							
wd							
XS							

[Remove](#)

- **Ghostery** ist ein Sicherheitsplugin für den Browser, das Tracker anzeigt & blocken kann.
- Identifiziert über 1000 **Tracking Services**.
- Kategorisiert in Cookies, Ad-Services, Analytics, Beacons & Widgets.
- Wie können wir uns die analytischen Fähigkeiten von Ghostery aneignen?

Repurposing Ghostery

The screenshot displays a web browser window with the Ghostery extension active. The extension's interface is overlaid on the page, showing a list of detected tracking services. The background page is a news article from the New York Times, with a blue banner for a 4-week subscription offer and an 'OPINION' section header.

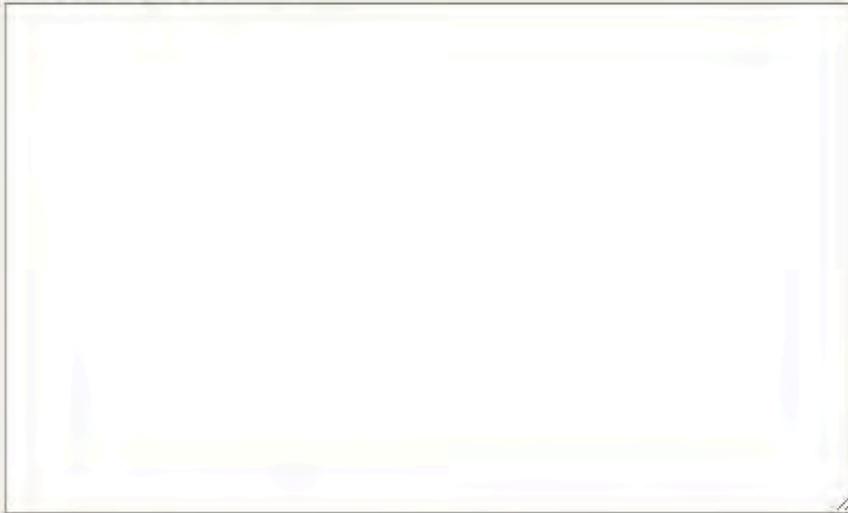
Service Name	More Info Link
Brightcove	more info
CheckM8	more info
Facebook Connect	more info
Google AdSense	more info
New York Times	more info

- **Tracker Tracker Tool** erlaubt systematische Nachfrage welche Tracker aktiv sind.
- **Input:** bis zu 1000 URLs.
- **Output:** Tracker Info zu allen URLs sowie Netzwerkvisualisierung von Tracker-Präsenz.

Tracker Tracker

Input

Enter URLs (1 per line):



Only look at specified pages Also look at subpages (N+1). max subpages/host:

Track trackers

Tracker Tracker, an Introduction

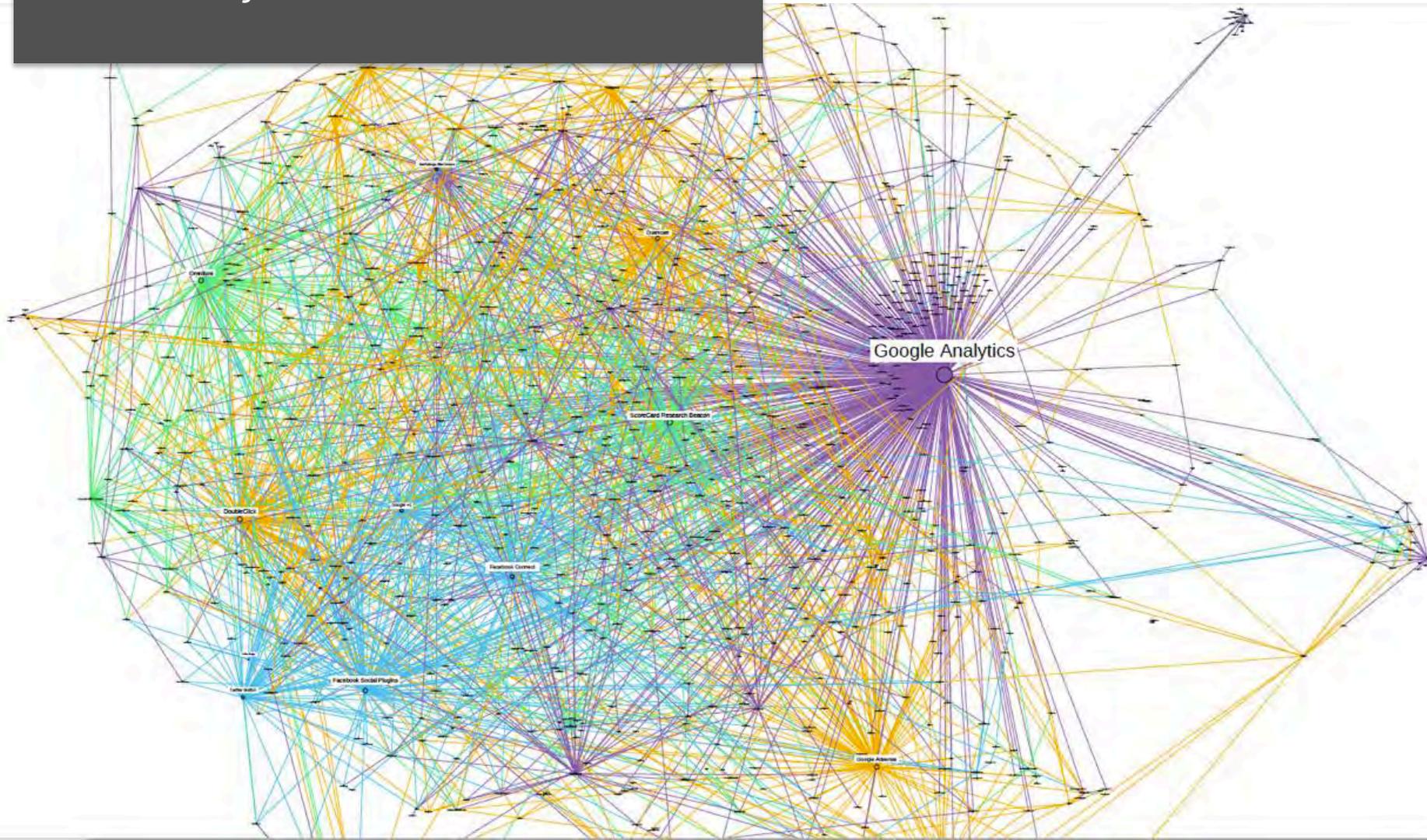
The tool Tracker Tracker detects from a predefined set of 'fingerprints' of web technologies, such as social buttons and trackers, and outputs a list of technologies used.

The tool Tracker Tracker can be used to make (some parts of) the 'cloud' visible. The tool allows for the characterization of a set of websites or pages by detecting a set of 900+ predefined 'fingerprints' of cloud devices, including those that fall under the category of analytics, ad programs, widgets or social plugins, trackers, and privacy. Tracker Tracker may thus be used to gain an overall picture of detectable trackers or for a number of specified analytical purposes, such as social plugin detection, mapping 'power concentrations of the cloud' - mapping the political economy of the cloud, by looking at 'cloud technology'. For more information and example uses see the project page:

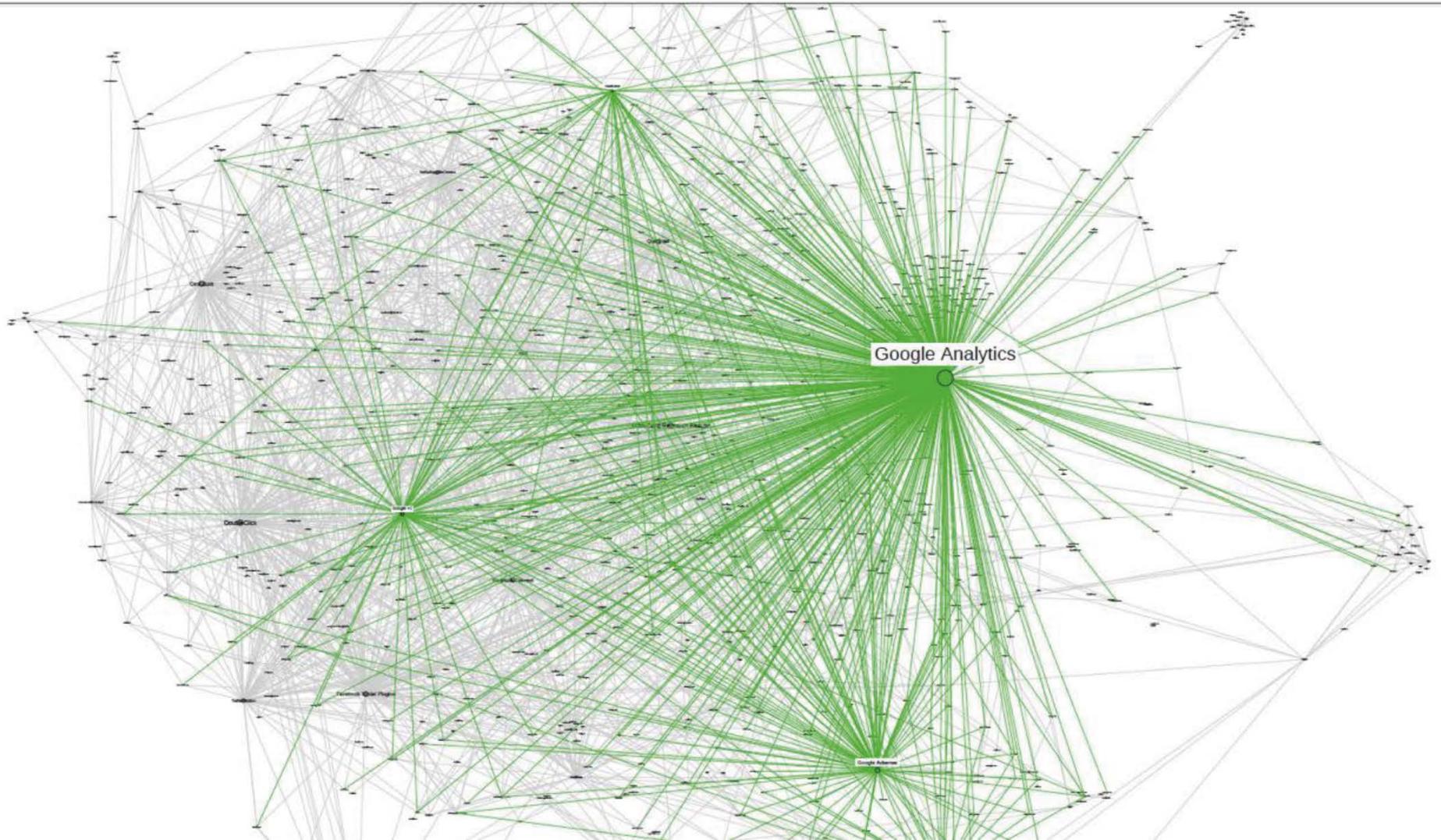
<https://wiki.digitalmethods.net/Dmi/DmiWinterSchool2012TrackingTheTrackers>.

Tracker Tracker, Sample Output ([Toggle](#))

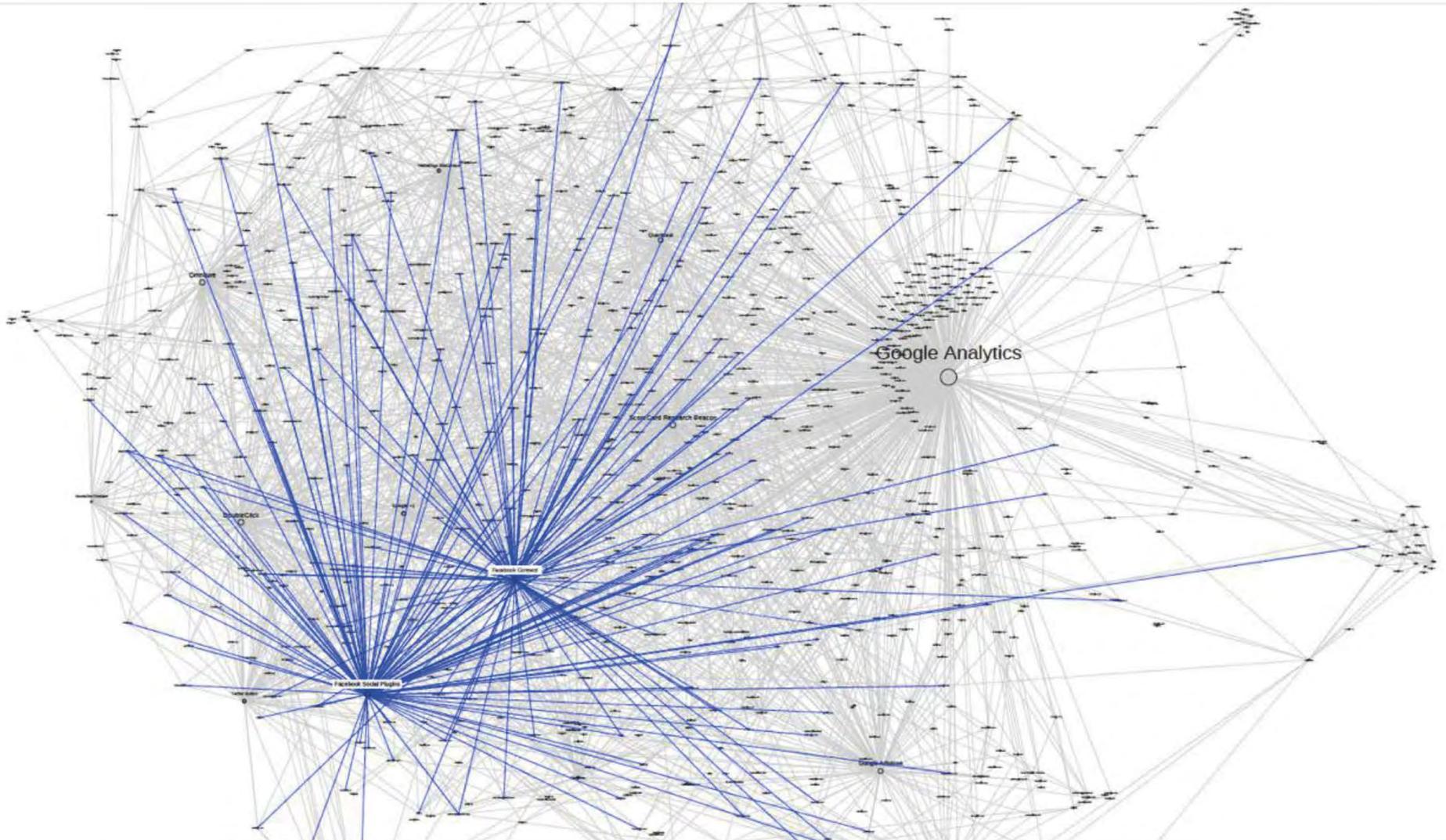
- Tracker in den **Top 1000** Webseiten nach **Alexa** (sortiert nach Seitenaufrufen).
- Blau: Widgets. Lila: Tracker. Ads: Ads.
Grün: Analytics.



Starke Präsenz von Google Tracking Services.



Wachsende Präsenz von Facebook.



Diskussion

1. Was sind die Potentiale und Grenzen von Aneignung/ Repurposing? Müssen Forschende die Algorithmen, die sie nutzen verstehen?
2. Welche anderen Devices, Plattformen oder Datenformate bieten sich zur Aneignung an?



2. Big & Small Data

Big & Small Data

Wachsendes Interesse an Big Data.

- **Burrows & Savage (2007):** Vollständige Datensätze.
- **Manovich (2011):** Daten überschreiten komputative Fähigkeiten.
- **Boyd & Crawford (2011):** Nicht Datenmenge ist entscheidend, sondern Beziehungen zwischen Daten.

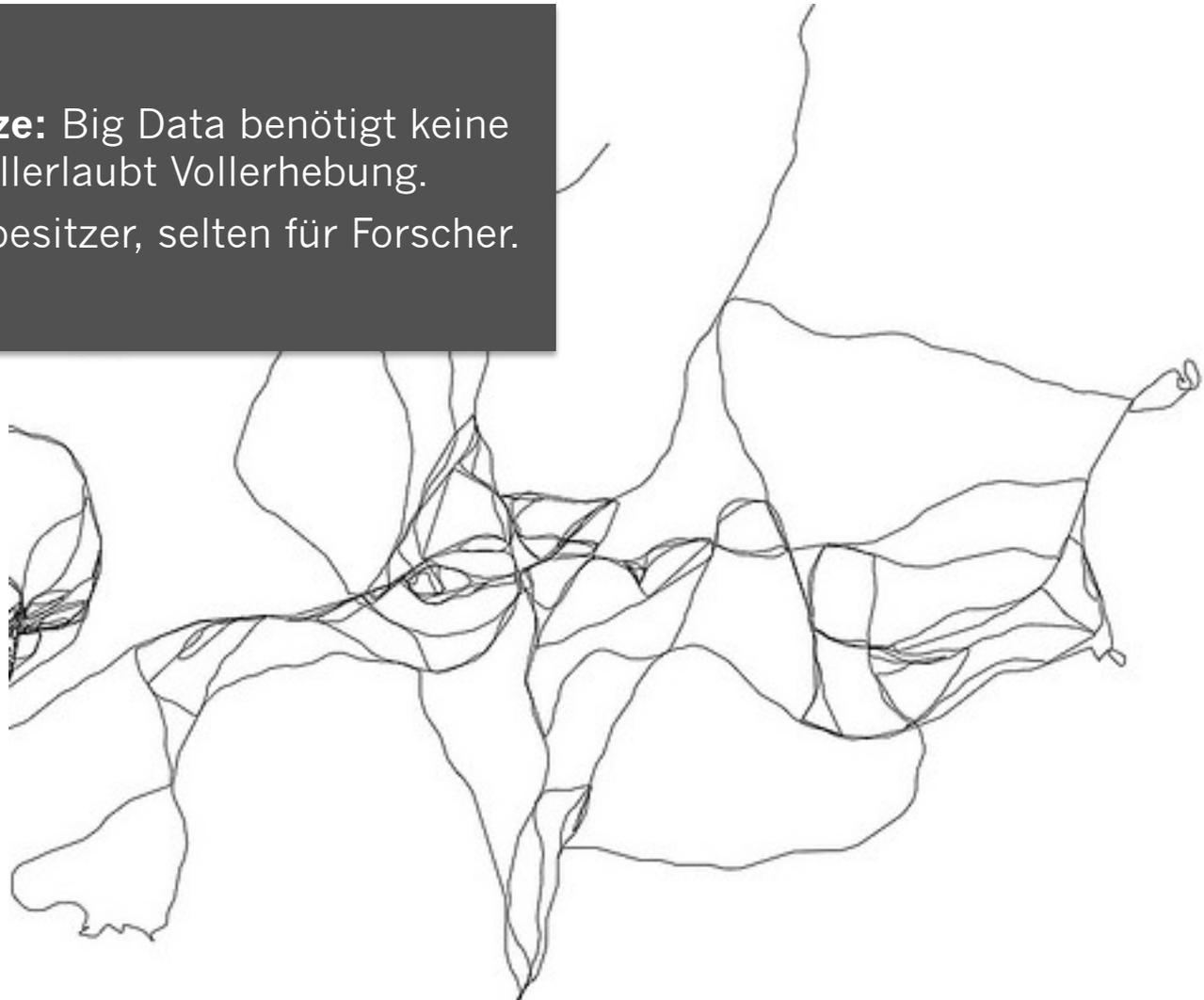


Digital big data

- **1st computational turn:** analoge Daten >> digitale Archive (Manovich 2011).
- **2nd computational turn:** native digitale Daten (Rogers 2009).
- Digitale Medien erlauben Aktion & Aufzeichnung zugleich
- Transaktionale Daten – Nebenprodukte von online Aktivitäten (Burrows & Savage 2007).

Big Data Hoffnungen I

- **Vollständige Datensätze:** Big Data benötigt keine Stichproben sondern allerlaubt Vollerhebung.
- Ja, aber nur für Datenbesitzer, selten für Forscher.



New Digital Divide

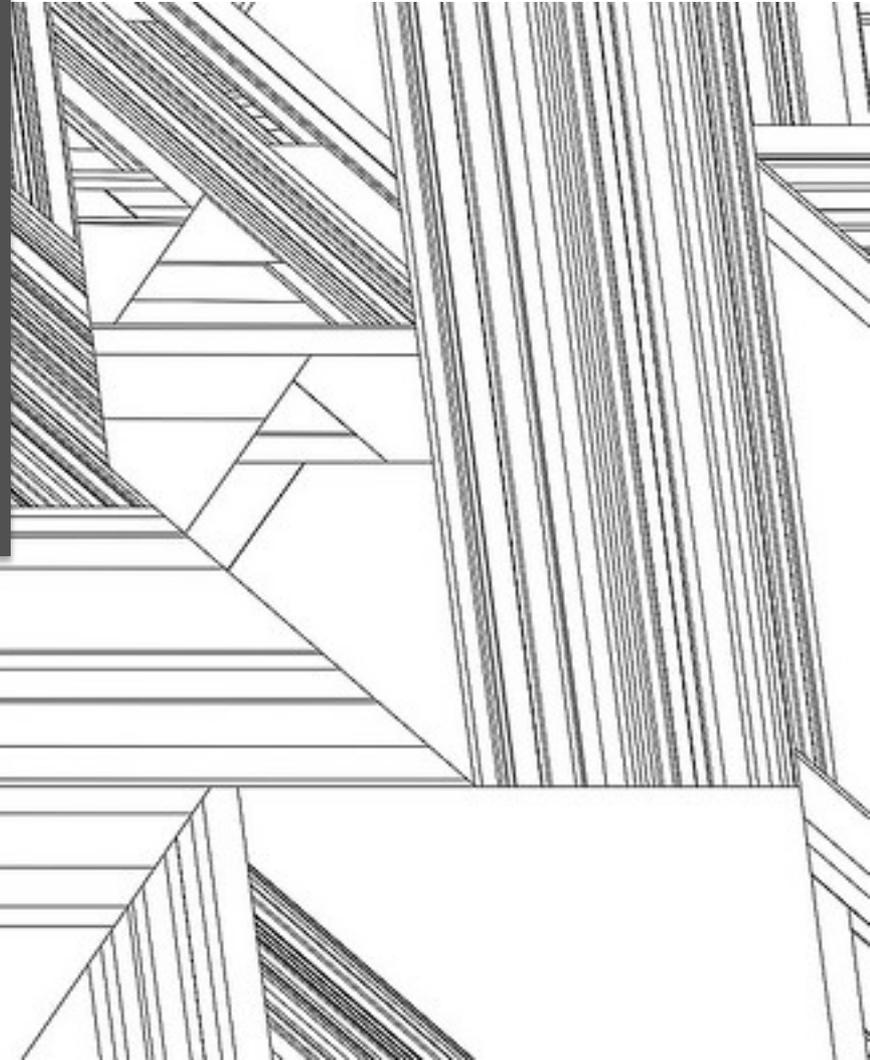
Neue digitale Klassen (Manovich 2011)

1. **Datenproduzenten:** alle online Nutzer.
2. **Datensammler:** wirtschaftliche, akademische oder staatliche Akteure mit Zugang zu bestimmten Daten.
3. **Datenanalysten:** Plattformen und Unternehmen mit Zugang zu kompletten Datensätzen und Analysefähigkeiten.



Small Data

- Nur wenige Akteure haben Zugang zu vollen Datensätzen.
- Akademische Forschung muss sich mit **Fallstudien oder Stichproben** begnügen.
- Stichproben als Demarkierung einzelner Datensätzen.
- **Probabilistische & nicht probalistische.**
- Allerdings: Wenig Reflexion zu Stichprobenverfahren Online.



Stichproben auf Twitter

- **Themenbasiert:** Demarkierung von Themen durch Suchbegriffe oder Hashtags (Bruns & Stieglitz 2012).
- **Schneeball-SP:** Demarkierung von ausgewählten NutzerInnen ausgehend von Listen und Followings (Paßmann et al 2012).
- **Marker-basierte Stichprobe:** Basierend auf Ort, Zeit, Sprache etc.

- Keine klare Relation zur Gesamtsumme aller Tweets sondern definiert durch interne Beziehungen.
- Stichprobenerfassung und Nutzungspraktiken sind verknüpft.

66939					Twitter for iPhone	Pittsburgh, Pa	40.451069	-79.973213
66940					Twitter for Android	HUD fam	53.588634	-1.855814
66941					Twitter for iPhone		24.584719	46.665154
66942					oursquare	Caxias do Sul, RS	-29.740219	-50.006813
66943					Twitter for iPhone	Madrid, Spain	40.479771	-3.73705
66944					Twitter for iPhone		24.812235	-107.358536
66945					Twitter for Android	Naples, Italy	32.455036	-84.989349
66946					Twitter for iPhone		40.602451	-79.566292
66947					Twitter for iPhone		36.812565	-2.578809
66948					Twitter for iPhone	-Lost in Paradise. â*žã™¥	48.592102	2.478296
66949					Twitter for iPhone	Ä*stanbul	41.003941	28.887203
66950	1359058199	39836.8402662037	AndrewGomes	Just got some mugworts a	Twitter for Android	Orlando, Fl	28.568054	-81.218948
66951	1359058199	39836.8402662037	Chrisrauc13008	@MarieFarret16 Bon choi	Twitter for iPad	Paris - France	48.83884	2.281054
66952					Twitter for BlackBerryÄ*		6.264539	-75.589256
66953					oursquare	NiterÄ*i	-22.821402	-43.046257
66954					akaz		40.731323	-73.990089
66955					Twitter for Android		48.368996	2.804719
66956					Twitter for iPhone	Sarnia, Ontario, Canada	42.95956	-82.350525
66957					Twitter for iPhone		48.922161	2.540953
66958					Twitter for Android	PiragÄXismo	39.977386	-0.030591
66959					Twitter for BlackBerryÄ*		-36.052979	-63.56995
66960					Twitter for iPhone	Prefuckingtoria, South Afr	-26.179771	28.051556
66961					Twitter for Windows Phon	Cincinnati	39.159229	-84.463249
66962					Twitter for Android		27.562107	41.682953
66963					Twitter for iPhone	Vers IÄ bas la ..	48.893059	2.445139
66964	1359058211	39836.8404050926	bru_cbn	Que shopping morto! Eu h	Twitter for Android		-9.987362	-67.845589
66965	1359058212	39836.8404166667	maxi_mario22	I tell u.its almost near!!!"	Twitter for BlackBerryÄ*	Offshore	7.48777	4.51759

Zufallsstichprobe

- **Zufallsstichprobe** erhebt Daten ohne Nutzungsweisen zu präferieren.
- Möglichkeit **einer vergleichenden Baseline**: Was ist die relative Bedeutung von Hashtags, Retweets, Sprachen etc?

66939					Twitter for iPhone	Pittsburgh, Pa	40.451069	-79.973213
66940					Twitter for Android	HUD fam	53.588634	-1.855814
66941					Twitter for iPhone		24.584719	46.665154
66942					foursquare	Caxias do Sul, RS	-29.740219	-50.006813
66943					Twitter for iPhone	Madrid, Spain	40.479771	-3.73705
66944	1359058195	39836.8402199074	melanymrz	Ugh!! Mi iphone no agarra	Twitter for iPhone		24.812235	-107.358536
66945	1359058195	39836.8402199074	DanaCoburn	So happy tomorrow is Frid	Twitter for Android	Naples, Italy	32.455036	-84.989349
66946	1359058195	39836.8402199074	J_rose112208	You don't talk shit but yet	Twitter for iPhone		40.602451	-79.566292
66947	1359058196	39836.8402314815	Garci_calabacin	@Ro_lechuguita jop!! Mu	Twitter for iPhone		36.812565	-2.578809
66948	1359058197	39836.8402430556	ManonBsqst	Y a une absence qui date c	Twitter for iPhone	-Lost in Paradise. â*žâ™¥	48.592102	2.478296
66949	1359058198	39836.8402546296	BCanCelik	Sporu iki ay bÄ±raktÄ±m h	Twitter for iPhone	Ä°stanbul	41.003941	28.887203
66950	1359058199	39836.8402662037	AndrewGomes5	Just got some mugwort's a	Twitter for Android	Orlando, Fl	28.568054	-81.218948
66951	1359058199	39836.8402662037	Chrisrauc13008	@MarieFarret16 Bon choi	Twitter for iPad	Paris - France	48.83884	2.281054
66952	1359058200	39836.8402777778	adricano2	Por mi que Cada quien qui	Twitter for BlackBerryÄ°*		6.264539	-75.589256
66953	1359058200	39836.8402777778	limanne_	I'm at Boulevard Shopping	foursquare	NiterÄ°i	-22.821402	-43.046257
66954	1359058200	39836.8402777778	Yakaz_Whisperer	Production Central (212) E	Yakaz		40.731323	-73.990089
66955	1359058200	39836.8402777778	Braychaman	@BenjaDrumEITota mdr	Twitter for Android		48.368996	2.804719
66956	1359058201	39836.8402893519	loafandjug	@theCHIVE @thechivery g	Twitter for iPhone	Sarnia, Ontario, Canada	42.95956	-82.350525
66957	1359058201	39836.8402893519	Inees_Salut	Il ma normal agresser alor	Twitter for iPhone		48.922161	2.540953
66958	1359058202	39836.8403009259	carlostrilles22	Si supierais lo divertido qu	Twitter for Android	PiragÄ°ismo	39.977386	-0.030591
66959	1359058202	39836.8403009259	BrandimarteLu	@Gonzalittus ok para la pi	Twitter for BlackBerryÄ°*		-36.052979	-63.56995
66960	1359058204	39836.8403240741	Just_That0	@Noah_james99 see you	Twitter for iPhone	Prefuckingtoria, South Afr	-26.179771	28.051556
66961	1359058204	39836.8403240741	kmillll	I'm getting back to my wa	Twitter for Windows Phon	Cincinnati	39.159229	-84.463249
66962	1359058206	39836.8403472222	arif_alslman	Ù...Ø³Ù„Ø³Ù„ Ù...ÙfØ³ÙŠÙ	Twitter for Android		27.562107	41.682953
66963	1359058208	39836.8403703704	CaamiBaaby	@ricomani @kurtiss935 @	Twitter for iPhone	Vers IÄ° bas la ..	48.893059	2.445139
66964	1359058211	39836.8404050926	bru_cbn	Que shopping morto! Eu h	Twitter for Android		-9.987362	-67.845589
66965	1359058212	39836.8404166667	maxi_mario22	I tell u.its almost near!!""	Twitter for BlackBerryÄ°*	Offshore	7.48777	4.51759

Using the Twitter Search API

View

[What links here](#)

Updated on Sat, 2012-08-25 10:26

API version 1 API version 1.1

The [Twitter Search API](#) is a dedicated API for running searches against the real-time index of recent Tweets. There are a number of important things to know before using the Search API which are explained below.

Limitations

- The Search API is not complete index of all Tweets, but instead an index of recent Tweets. At the moment that index includes between 6-9 days of Tweets.
- You cannot use the Search API to find Tweets older than about a week.
- Queries can be limited due to complexity. If this happens the Search API will respond with the error:

```
{ "error": "Sorry, your query is too complex. Please reduce complexity and try again." }
```
- Search does not support authentication meaning all queries are made anonymously.
- Search is focused in relevance and not completeness. This means that some Tweets and users may be [missing from search results](#). If you want to match for completeness you should consider using the [Streaming API](#) instead.
- The `near` operator cannot be used by the Search API. Instead you should use the `geoCode` parameter.
- Queries are limited to 1,000 characters in length, including any operators.
- When performing geo-based searches with a radius, only 1,000 distinct subregions will be considered when evaluating the query.

Möglichkeiten einer Zufallsstichprobe

1. REST APIs
2. Search API
3. Streaming APIs (1% Stichprobe)

Zufallsstichprobe

The Streaming APIs

View

[What links here](#)

Updated on Mon, 2012-09-24 13:47

API version 1 API version 1.1

Overview

The set of streaming APIs offered by Twitter give developers low latency access to Twitter's global stream of Tweet data. A proper implementation of a streaming client will be pushed messages indicating Tweets and other events have occurred, without any of the overhead associated with polling a REST endpoint.

Twitter offers several streaming endpoints, each customized to certain use cases.

Public streams

Streams of the public data flowing through Twitter. Suitable for following specific users or topics, and data mining.

User streams

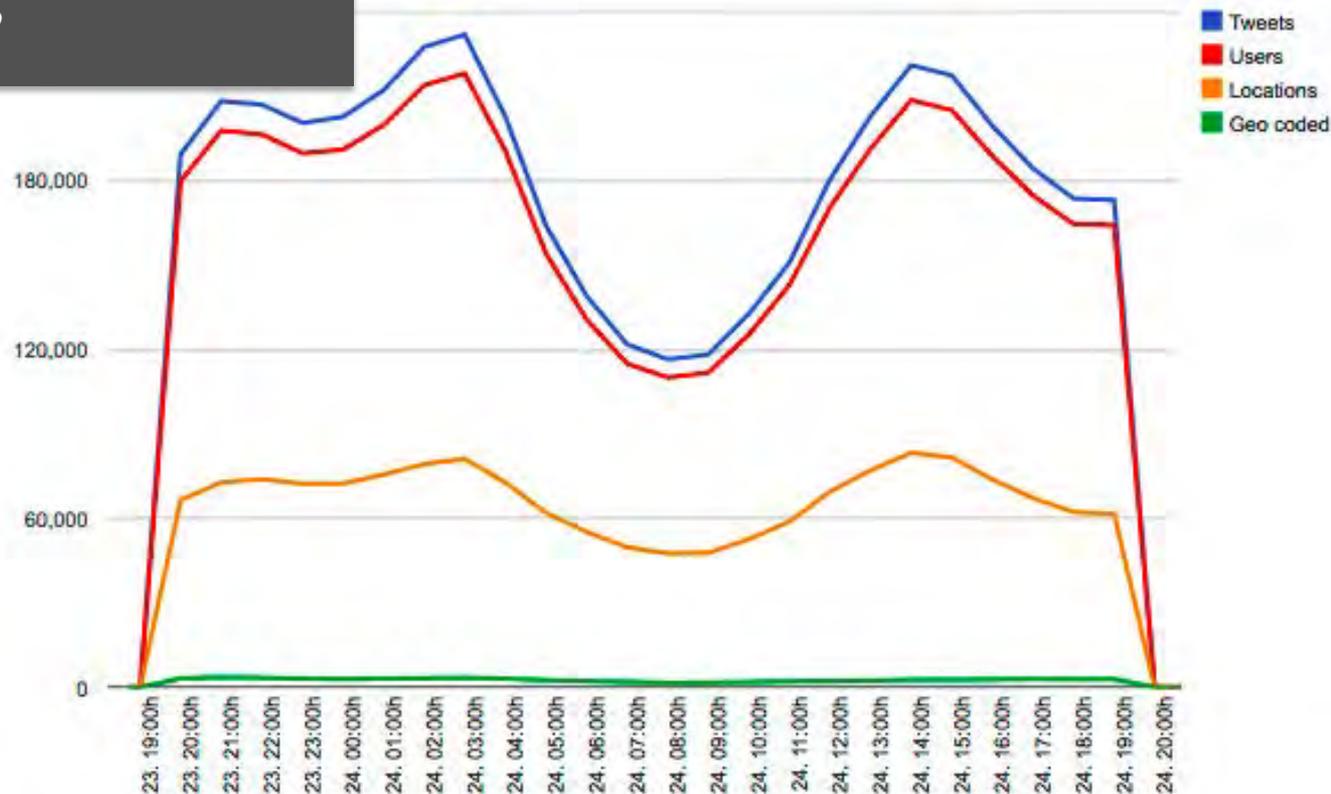
Single-user streams, containing roughly all of the data corresponding with a single user's view of Twitter.

Site streams

The multi-user version of user streams. Site streams are intended for servers which must connect to Twitter on behalf of many users.

Case: Mining One Percent of Twitter

- 24 Stunden 1% Twitterdaten über die Twitter Streaming API.
- 4,376,230 tweets, 3,370,796 accounts
- 60% Englisch.
- Hashtags: 13,18%
- URL: 11.7%
- @mention: 57,2%
- Direkter @reply 46,8%



- Nur 11% aller Tweets erhalten URLs.
- Prädominanz von Social Media, mittlere Bedeutung von Nachrichten.

URLs

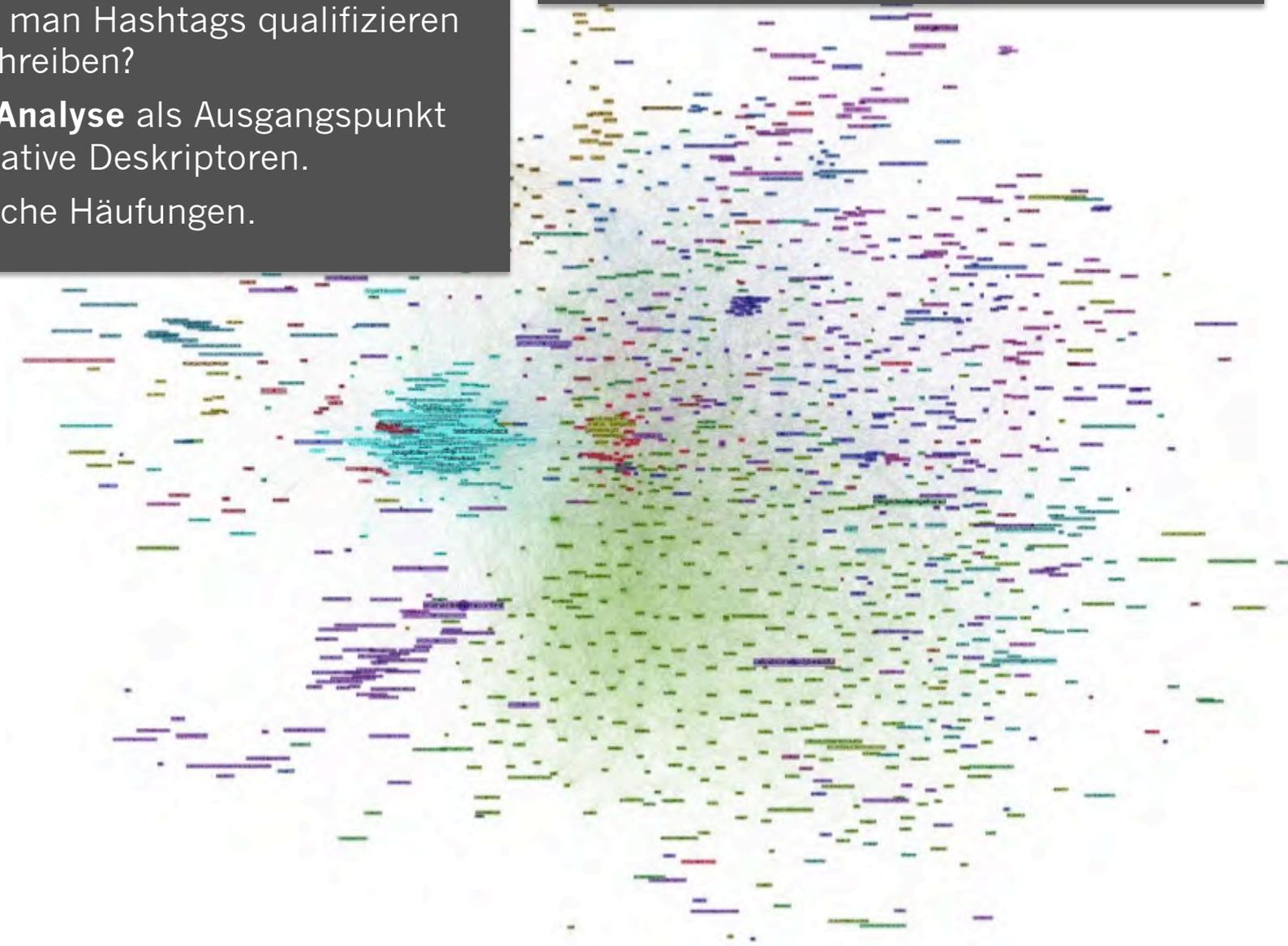
YOUTUBE.COM (38217) FACEBOOK.COM (37627)

INSTAGRAM.COM (35175) ASK.FM (30054) TWITPIC.COM (14176)

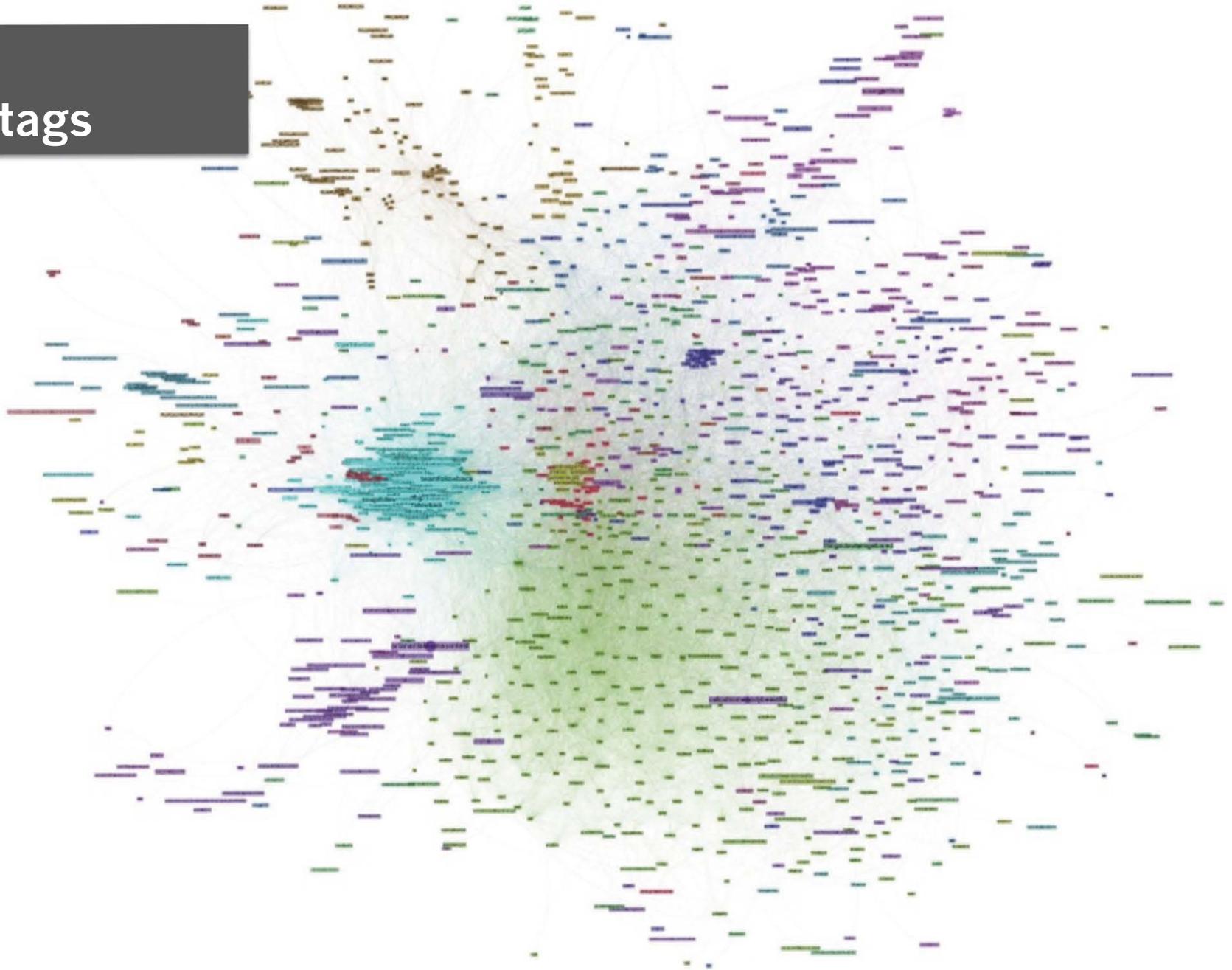
FOURSQUARE.COM (6242) UNFOLLOWERS.ME (6162) TWITLONGER.COM (4281) TWITTER.YFROG.COM (4171) AMAZON.CO.JP (3931) M.TMI.ME (3927) TWITTASCOPE.COM (3648) TWITTER.COM (3524) GOAL.COM (2937) APPS.FACEBOOK.COM (2826) ITUNES.APPLE.COM (2816) INFO-ZERO.JP (2785) TWITCAM.LIVESTREAM.COM (2605) PICS.LOCKERZ.COM (2600) 25.MEDIA.TUMBLR.COM (2521) AMEBLO.JP (2504) AMAZON.COM (2077) WEHEARTIT.COM (2068) 24.MEDIA.TUMBLR.COM (1989) SOUNDCLLOUD.COM (1866) P.TWIPPLE.JP (1757) SHINDANMAKER.COM (1640) **LATIMES.COM (1605)** M.VK.COM (1598) TRIBEZ-GAME.COM (1587) ADF.LY (1558) **BBC.CO.UK (1481)** NICOVIDE0.JP (1424) **REUTERS.COM (1385)** ITEM.RAKUTEN.CO.JP (1353) TWITCOM.COM.BR (1316) **WASHINGTONPOST.COM (1219)** JUSTUNFOLLOW.COM (1209) TWITCASTING.TV (1112) REKKACOPY.COM (1096) URANAITTER.COM (1082) PATH.COM (1000) **ABCNEWS.GO.COM (977)** TMI.ME (970) PAPER.LI (966) GETGLUE.COM (885) PLAY.GOOGLE.COM (884) NEWS.DETIK.COM (860) USTREAM.TV (822) MEDIA.TUMBLR.COM (819) BLOG.LIVEDOOR.JP (805) OW.LY (783) **HUFFINGTONPOST.COM (774)** MASHABLE.COM (774) LIVE.NICOVIDE0.JP (770) ETSY.COM (750) FLLWRS.COM (728) VIA.ME (722) GAME-INSIGHT.COM (722) M.YOUTUBE.COM (716) **GUARDIAN.CO.UK (688)** GUNGHO.JP (639) REVERBNATION.COM (632) **NEWS.YANDEX.RU (617)** **G1.GLOBO.COM (611)** M.FACEBOOK.COM (591) NOTFOLLOW.ME (591) EDITION.CNN.COM (581) STARDOLL.COM (572) SHORTWEB.US (553) MATOME.NAVER.JP (535) PHOTOZOU.JP (531) PBS.TWIMG.COM (526) NUBEE.COM (478) SPORT.DETIK.COM (473) EBAY.COM (463) Q.GS (452) TUITUTIL.NET (413) CDN.KEEK.COM (412) ESPN.GO.COM (410) **HEADLINES.YAHOO.CO.JP (406)** NETWORKEDBLOGS.COM (402) **NEWS.YAHOO.COM (385)** VIMEO.COM (384) I.IMGUR.COM (380) FLICKR.COM (376) KEEK.COM (365) HELIUM.COM (358) OPEN.SPOTIFY.COM (353) MIRRORSOFLBION.COM (350) ALLKPOP.COM (349) REGIONAL.KOMPAS.COM (348) BOLA.NET (344) BLOG.NAVER.COM (322) TECHCRUNCH.COM (318) **FORBES.COM (315)** **DAILYMAIL.CO.UK (310)** CUTT.US (309) DATPIFF.COM (300)

- 844,602 Hashtags, davon 227,029 individuelle.
- Wie kann man Hashtags qualifizieren und beschreiben?
- **Co-Wort Analyse** als Ausgangspunkt für alternative Deskriptoren.
- Thematische Häufungen.

Hashtags

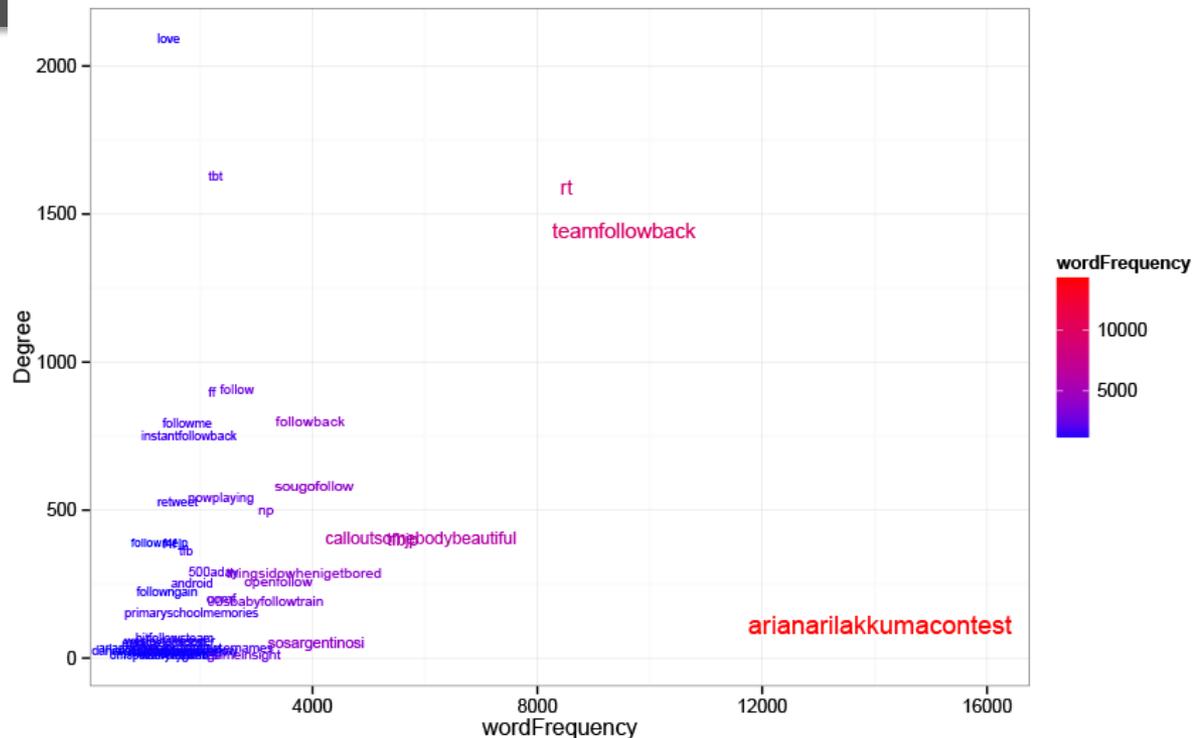


Hashtags



- Dichte und Frequenz der Verknüpfung: **Degree range.**
- **Kombi-Hashtags** ohne Themenbindung: #love, #me, #lol, #instagram, #follow.
- **Themenbindung:** #arianarikkumacontest, #thingsidowhenigetbored.

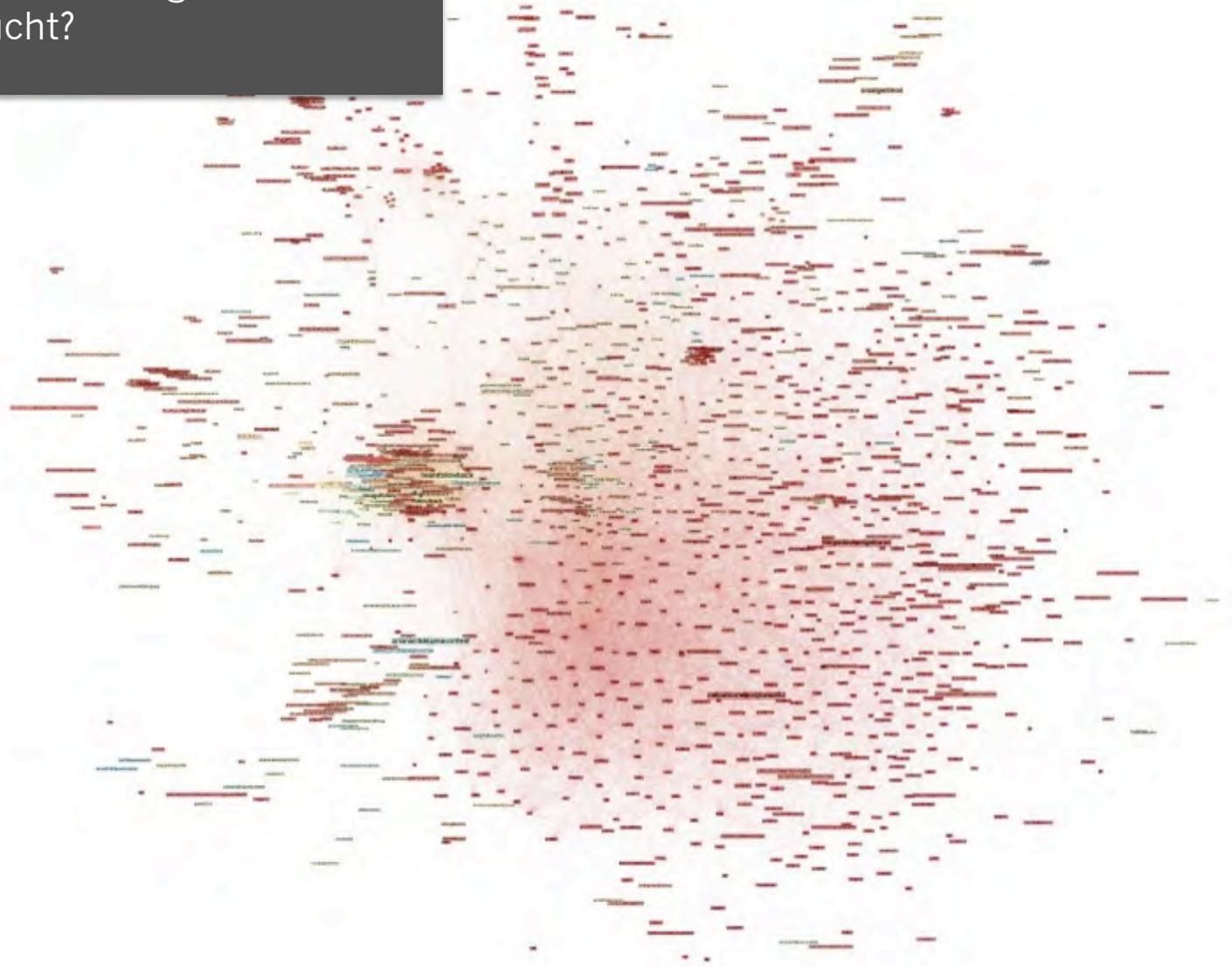
Hashtags



Nutzer-Diversität

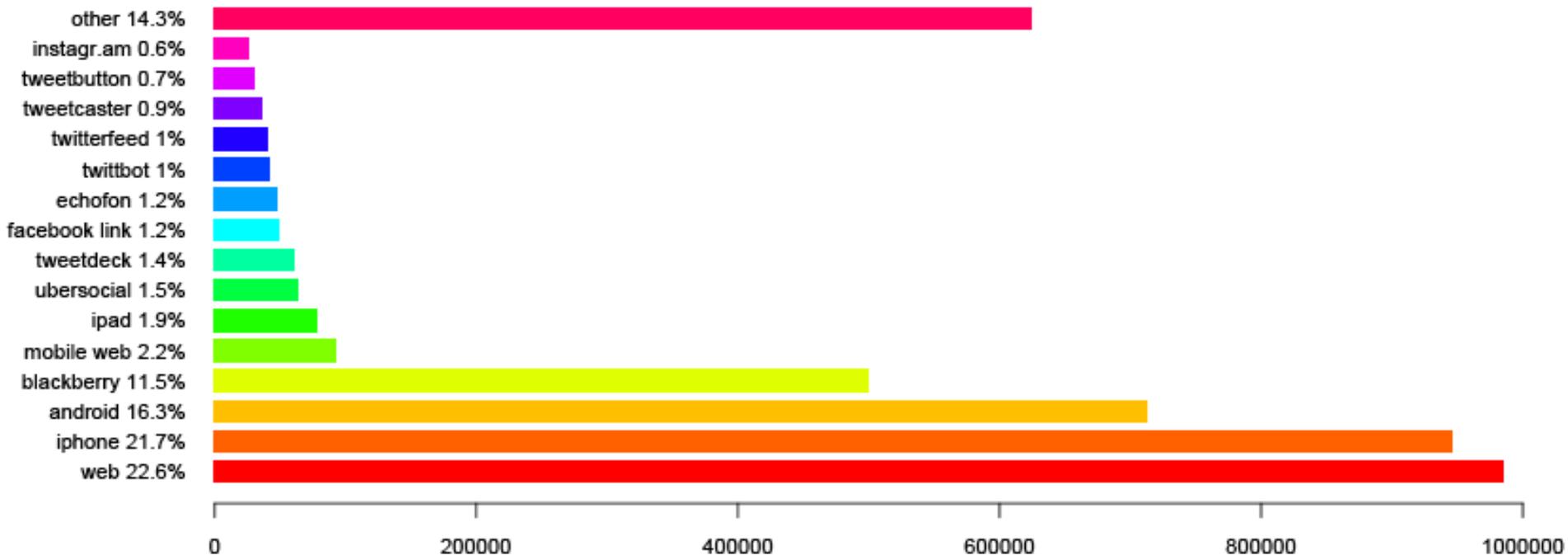
- Wird Hashtags von vielen verschiedenen oder wenigen aktiven Nutzern gebraucht?

Hashtags



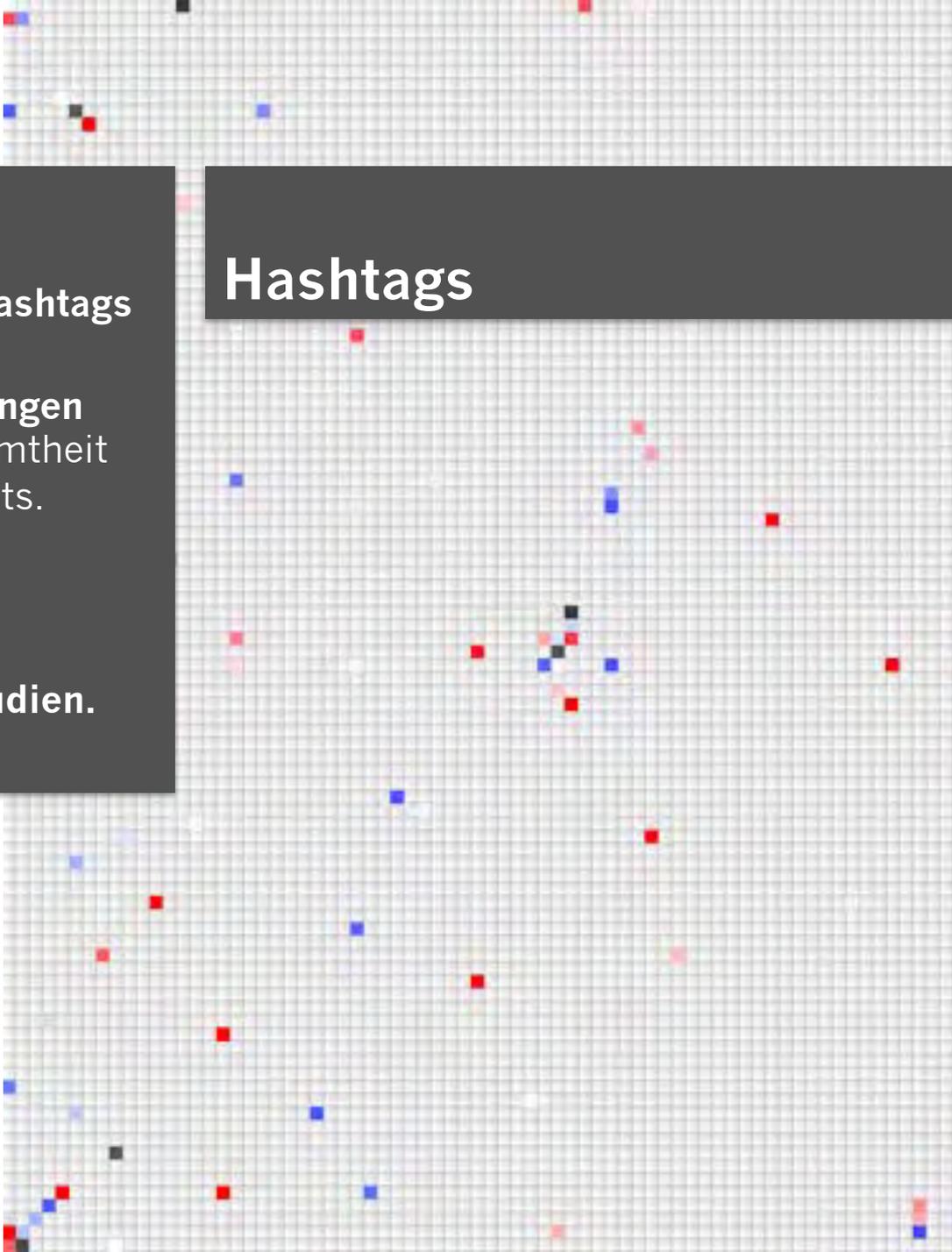
- Nur 22,6% aller Tweets aus dem Web gesendet.
- 18,248 Zugangspunkte.
- Longtail Verteilung.

Zugangspunkte



- Deskriptoren machen auf unterschiedliche **Rollen von Hashtags** aufmerksam.
- Stichproben **schaffen Beziehungen** zwischen Medium, Grundgesamtheit (Big Data) und einzelnen Tweets.
- Basiert auf Wissen über Nutzungspraktiken.
- Zufallsstichprobe ermöglicht **Kontextualisierung von Fallstudien.**

Hashtags



Series: Glenn Greenwald on security and liberty

Previous |

GLENN GREENWALD

ON SECURITY AND LIBERTY



Was the London killing of a British soldier 'terrorism'?

What definition of the term includes this horrific act of violence but excludes the acts of the US, the UK and its allies?

Share
Tweet this
Email



Glenn Greenwald
guardian.co.uk, Thursday 23 May 2013 14.03 BST



Article history

UK news

Woolwich attack

Series

Glenn Greenwald on security and liberty

More from Comment is free on

UK news

Woolwich attack

Series

Glenn Greenwald on security and liberty

More on this story



Woolwich attack: MI5



Klus with football skills

Robin van Persie shows off his incredible football skills



Anti-gravity 3D printing

project allows users to create 3D objects on any given working surface



Bayern Munich brand 'more valuable than Manchester United's'

Brand Finance rate Bayern Munich as the most valuable football brand in the world

About this series

A critical, campaigning column on vital issues of civil rights, freedom of information and justice – and their enemies, from the award-winning journalist, former constitutional litigator and author of three New York Times bestsellers.

Follow @ngreenwald on Twitter or email him at



A man appearing to be holding holding a knife following the Woolwich attack. Photograph: Pixel8000

(updated below)

Two men yesterday engaged in a horrific act of violence on the streets of

Aufgabe

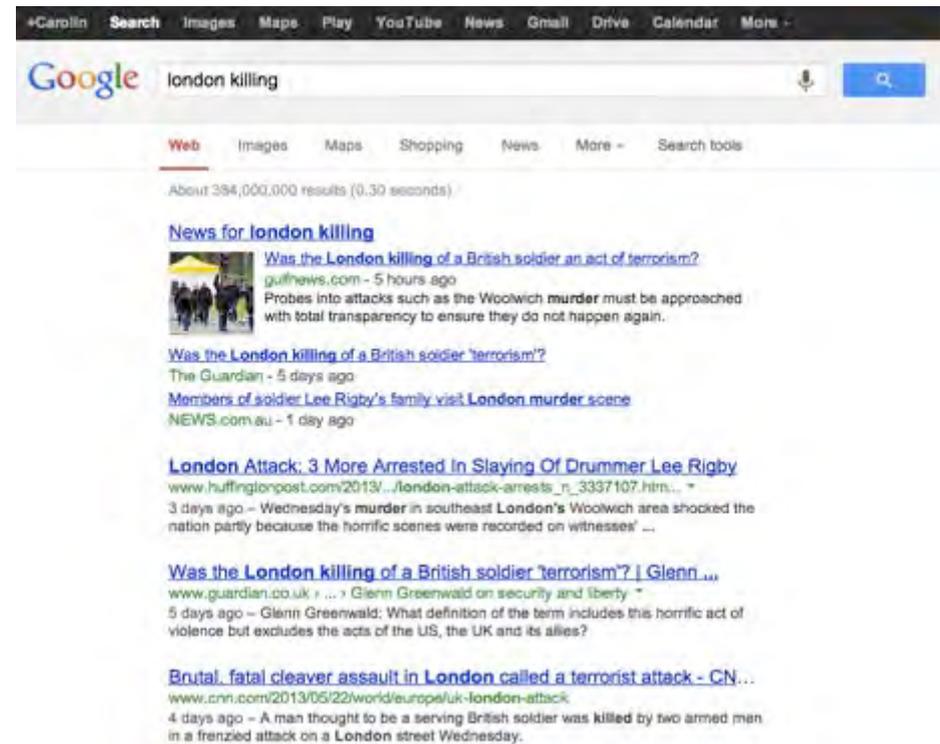
Angenommen, ihr wollt die Diskussion des Londoner Soldatenattentats auf Twitter untersuchen. Welche Stichproben kommen in Frage? Warum?

3. Medien vs. Gesellschaft?



Medium vs. Gesellschaft

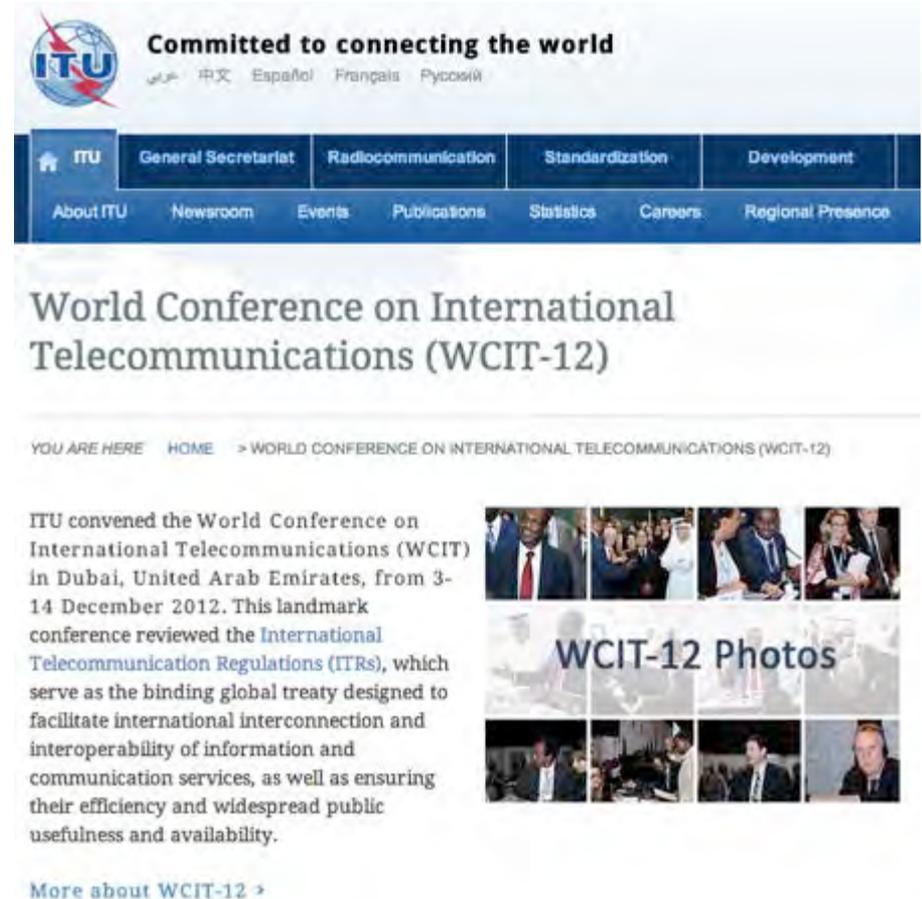
- “When looking at Google, do we see society or Google?” (Rogers 2009)
- Nutze das Medium, um das Medium zu verstehen vs. Nutze das Medium, um gesellschaftliche Fragen zu verstehen.
- Erlauben software-basierte Daten Einsichten in **Medien- oder Themendynamiken?**



The image shows a screenshot of a Google search results page for the query "london killing". The search bar at the top contains the text "london killing" and a search button. Below the search bar, there are navigation tabs for "Web", "Images", "Maps", "Shopping", "News", "More", and "Search tools". The search results show approximately 394,000,000 results found in 0.30 seconds. The first result is a news article titled "Was the London killing of a British soldier an act of terrorism?" from GulfNews.com, dated 5 hours ago. The second result is another news article titled "Was the London killing of a British soldier 'terrorism'?" from The Guardian, dated 5 days ago. The third result is a news article titled "Members of soldier Lee Rigby's family visit London murder scene" from NEWS.com.au, dated 1 day ago. The fourth result is a news article titled "London Attack, 3 More Arrested in Slaying Of Drummer Lee Rigby" from HuffingtonPost.com, dated 3 days ago. The fifth result is a news article titled "Was the London killing of a British soldier 'terrorism'? | Glenn ..." from The Guardian, dated 5 days ago. The sixth result is a news article titled "Brutal, fatal cleaver assault in London called a terrorist attack - CN ..." from CNN.com, dated 4 days ago.

Case: WCIT & Twitter

- WCIT Konferenz in Dubai, 3-14 Dezember 2012 organisiert durch die ITU.
 - Interval: 23.11 – 19.12.
 - 108781 Tweets.
-
- Vor der Konferenz: Erhebung von **Issue-Themen** durch NGO & akademischen Experten.
 - **Forschungsfrage:** Wie verändert sich der Diskurs um WCIT und wie werden Themen verschiedener Akteure aufgegriffen?



Committed to connecting the world
عربي 中文 Español Français Русский

ITU General Secretariat Radiocommunication Standardization Development
About ITU Newsroom Events Publications Statistics Careers Regional Presence

World Conference on International Telecommunications (WCIT-12)

YOU ARE HERE [HOME](#) > WORLD CONFERENCE ON INTERNATIONAL TELECOMMUNICATIONS (WCIT-12)

ITU convened the World Conference on International Telecommunications (WCIT) in Dubai, United Arab Emirates, from 3-14 December 2012. This landmark conference reviewed the [International Telecommunication Regulations \(ITRs\)](#), which serve as the binding global treaty designed to facilitate international interconnection and interoperability of information and communication services, as well as ensuring their efficiency and widespread public usefulness and availability.

[More about WCIT-12 >](#)

WCIT-12 Photos

Co-Wort Analyse

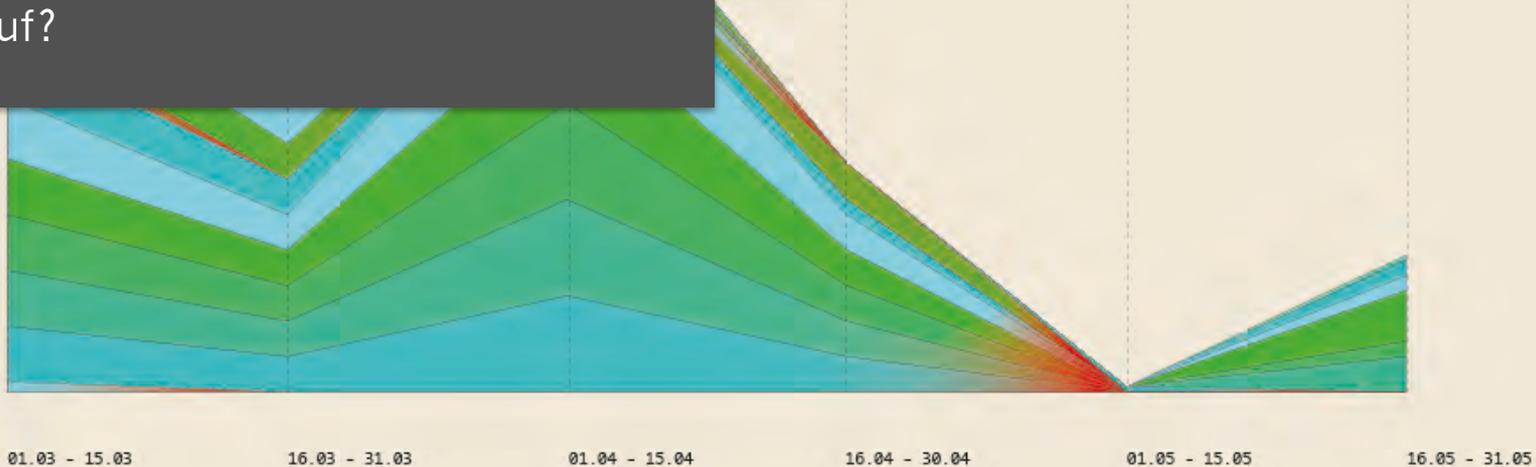
- Co-Wort Analyse untersucht welche Begriffe in der Nähe von einander auftauchen (Callon et al 1983, Danowski 2009).
- Hashtag Co-Occurrence Analysis & “Associational Profiles”: Welche Hashtags tauchen zusammen mit anderen in bestimmten Intervallen auf?

⚙ #emissions 💡 #uk
⚙ #adaptation 💡 #change 💡 #yorkshire

💡 #rivers
💡 #agriculture
💡 #education
💡 #extremeweather
💡 #environment
💡 #africa
💡 #mexico
💡 #maize
⚙ #flood

💡 #cancun
💡 #co2
💡 #cop16
💡 #copenhagen
💡 #economics
💡 #mudslide
💡 #flood
⚙ #africa

💡 #africa
⚙ #adaptation
⚙ #kiswahili
⚙ #tanzania
⚙ #sindebele
⚙ #zimbabwe
⚙ #poverty

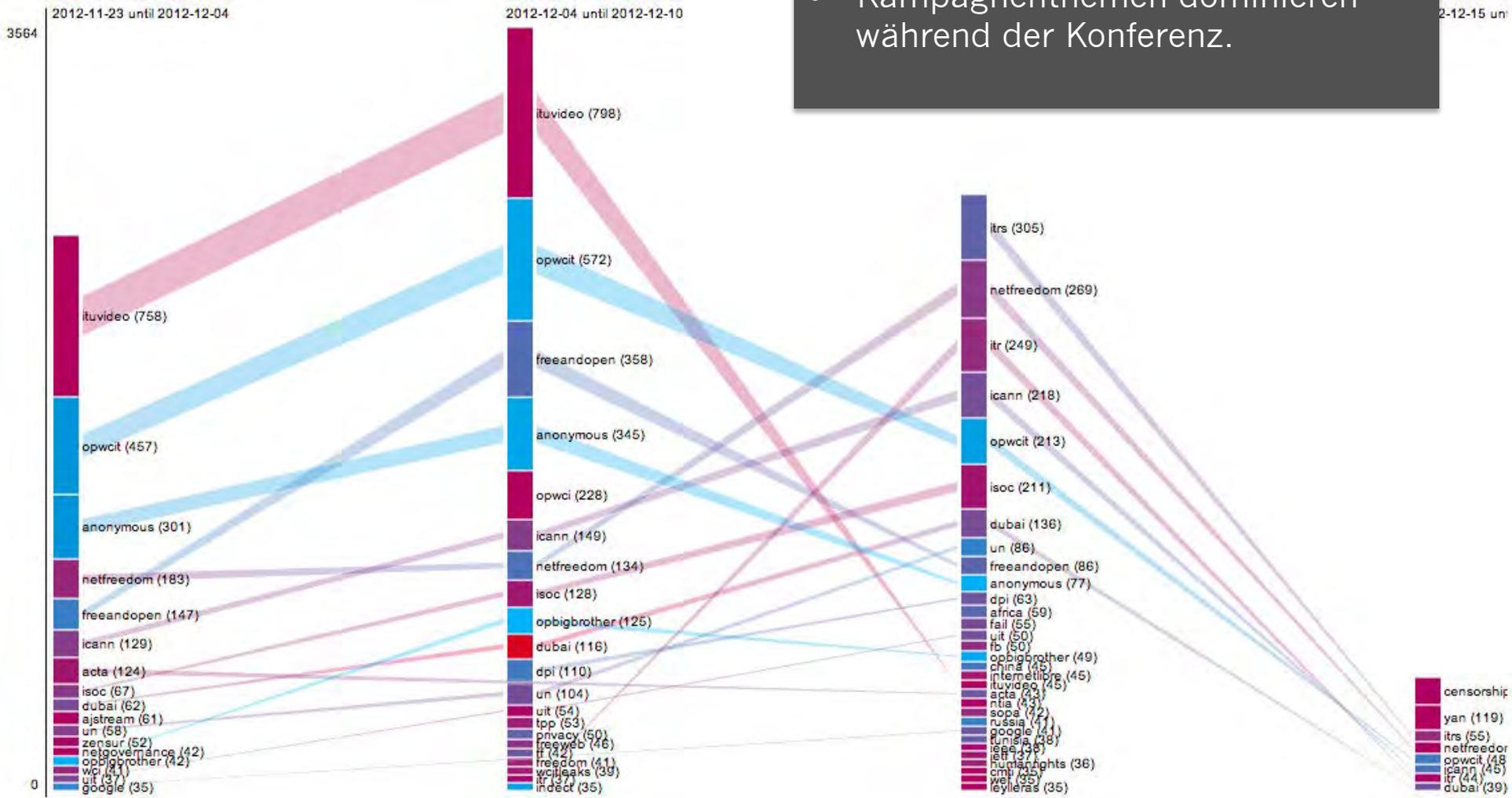


#WCIT Profil

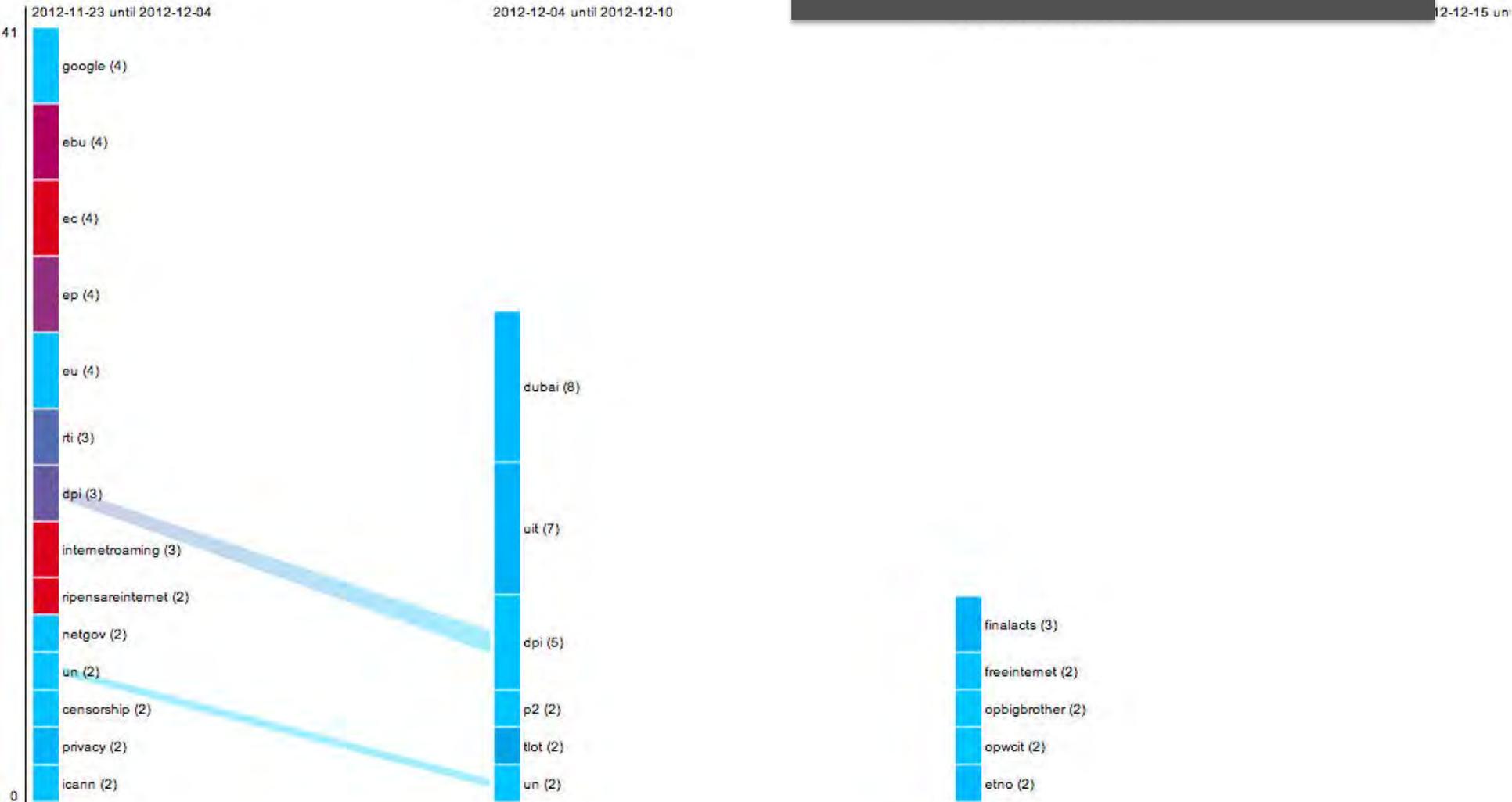
- **Kampagnen Hashtags:** #netfreedom, #ituvideo, #anonymous, #opbigbrother.

- **Institutions Hashtags:** #isoc, #ican.

- Kampagnenthemen dominieren während der Konferenz.



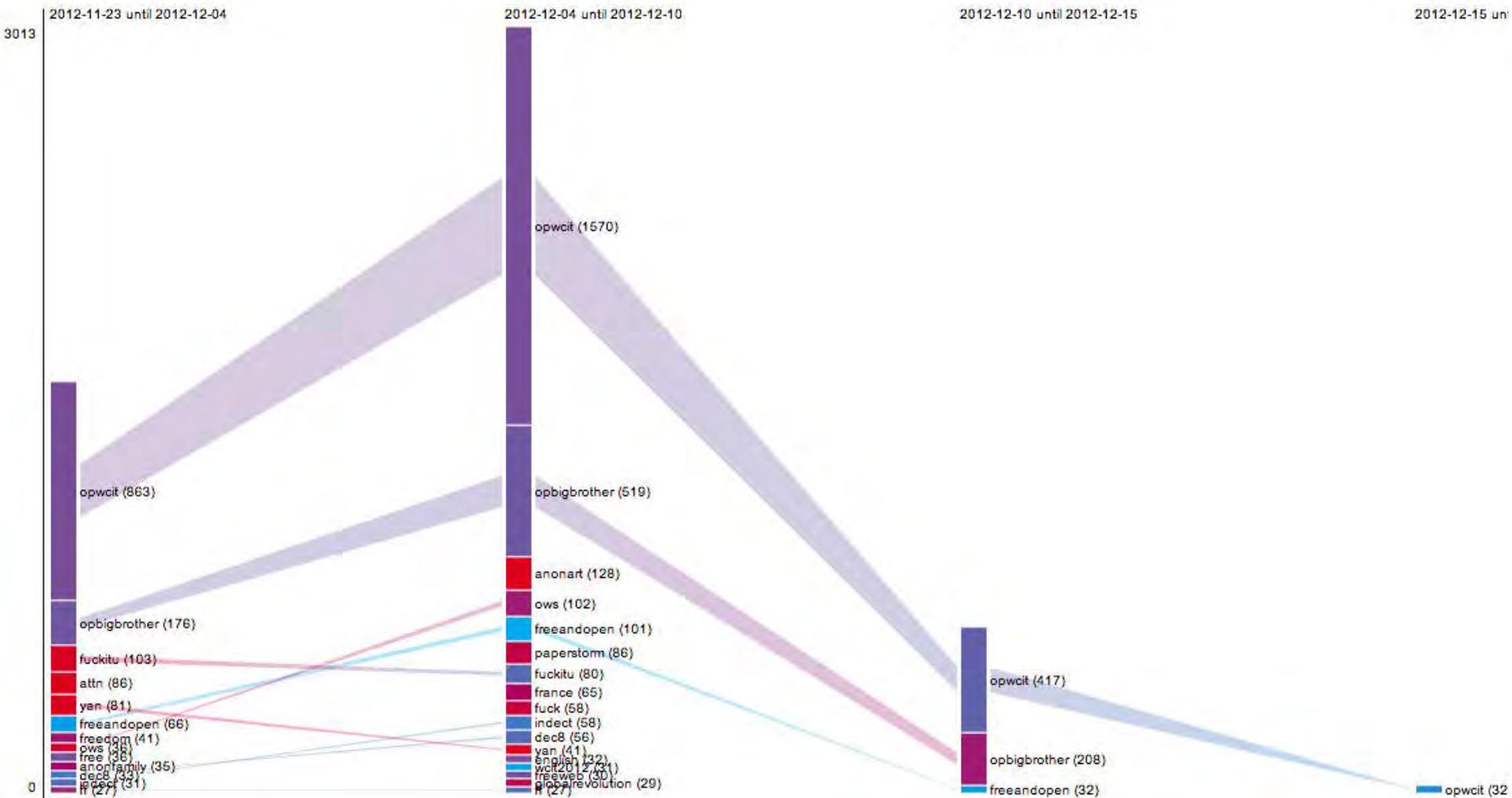
Vor der Konferenz



- **#netneutrality**
- Start vor, relativ schwach während der Konferenz.

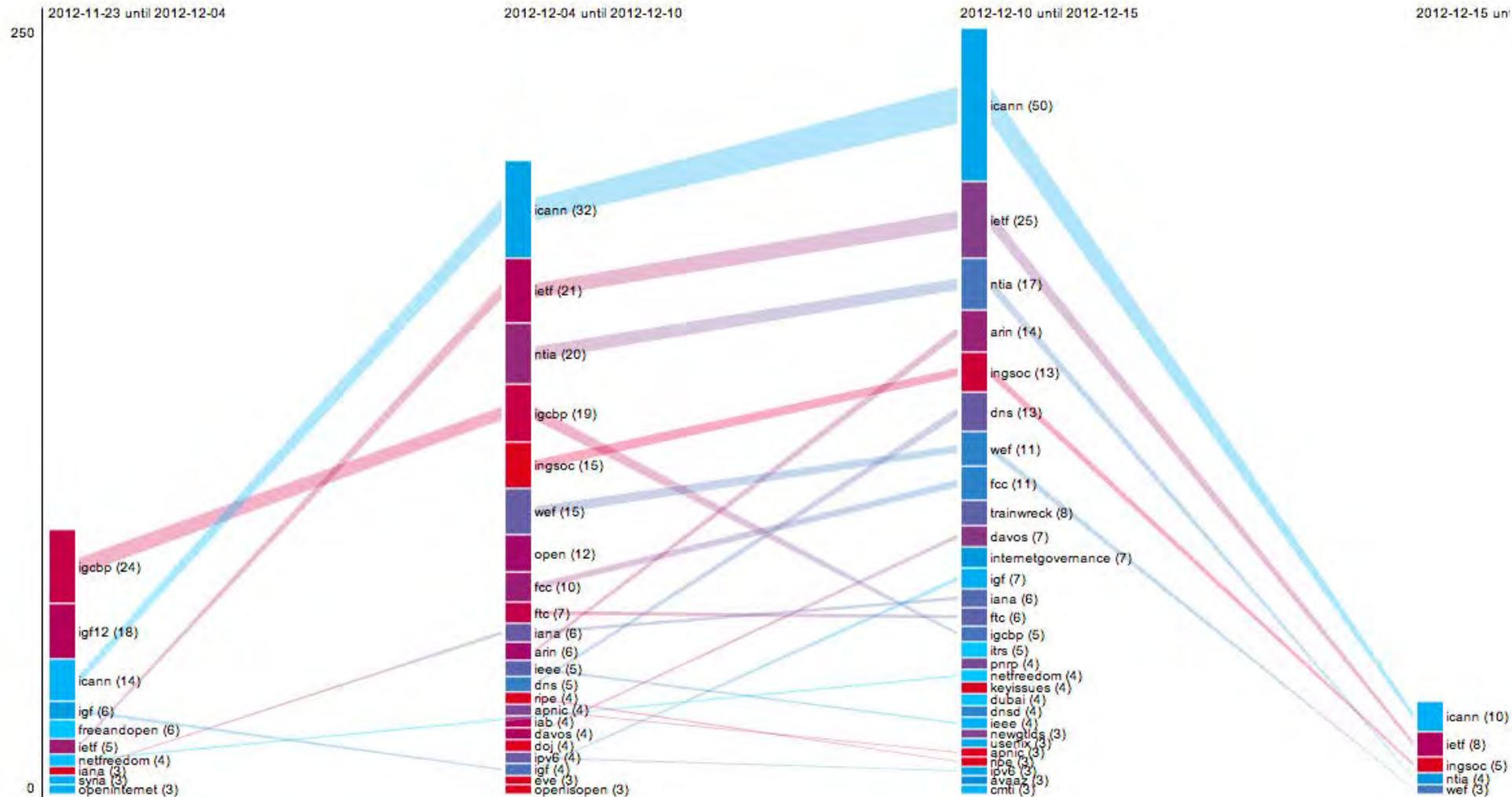
Beginn der Konferenz

- #anonymous
- Kampagnenhashtag.



Ende der Konferenz

- #isoc
- Institutions Hashtag.



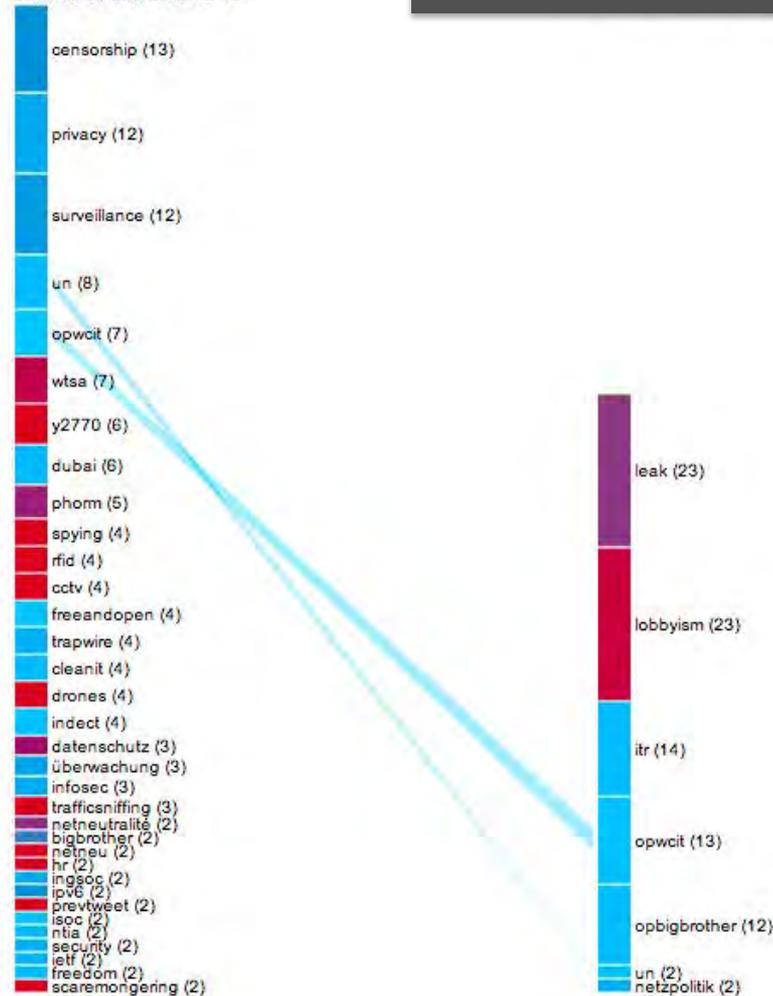
Varianz der Assoziationen

- #netneutrality
- Experten Hashtag.
- Fallende Bedeutung und wandelnde Assoziationen.

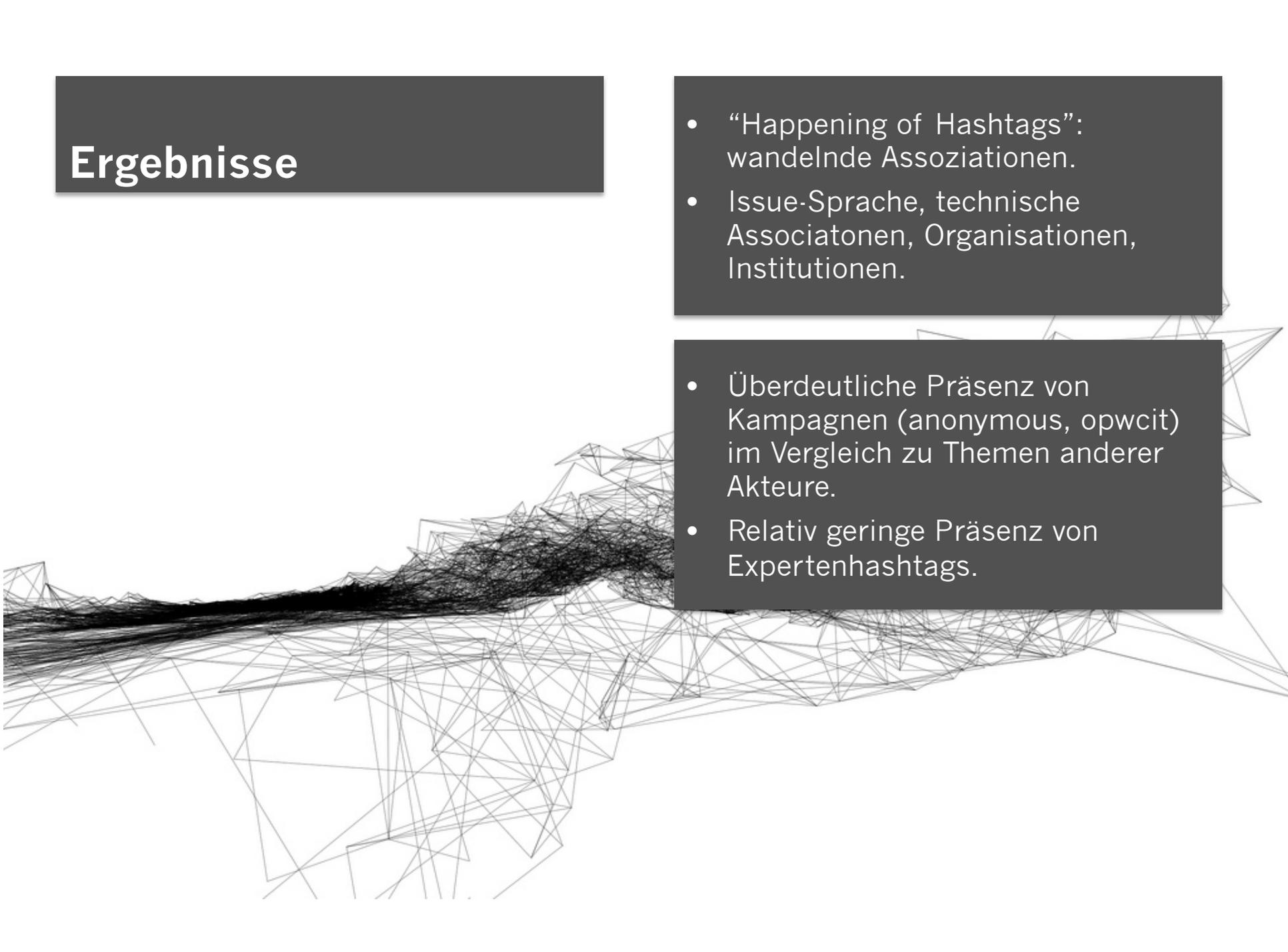
2012-11-23 until 2012-12-04

2012-12-04 until 2012-12-10

2-15 un



Ergebnisse



- “Happening of Hashtags”: wandelnde Assoziationen.
- Issue-Sprache, technische Assoziationen, Organisationen, Institutionen.

- Überdeutliche Präsenz von Kampagnen (anonymous, opwcit) im Vergleich zu Themen anderer Akteure.
- Relativ geringe Präsenz von Expertenhashtags.

Team

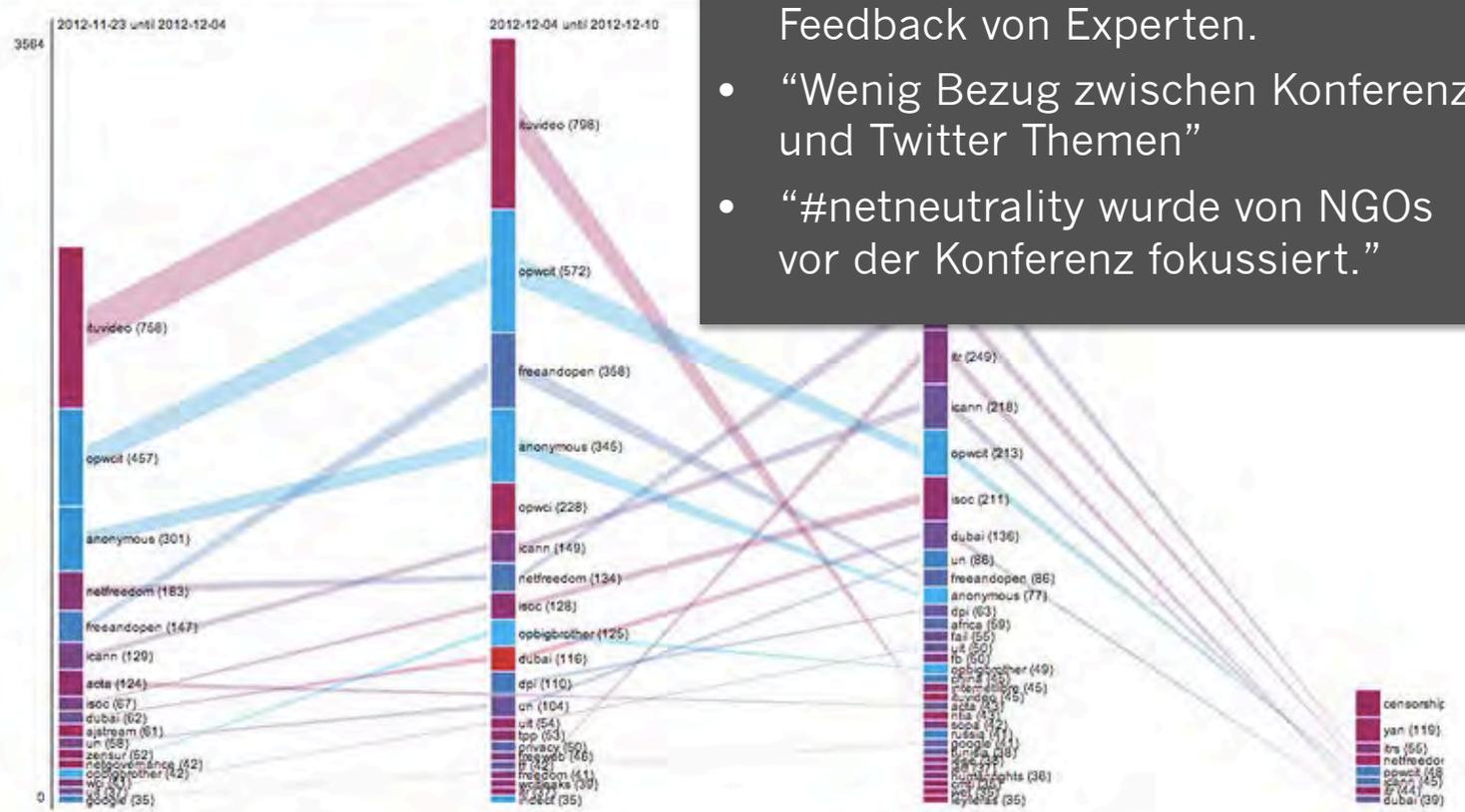
Medien vs. Themen

Below we present the Twitter profiles of a selection of these issue terms. We produced this data by using an experimental online tool for textual analysis, called the Twitter Profiler. To analyse this data, we used an experimental online tool for textual analysis, called the Twitter Profiler. The corresponding hashtag for each of the issues suggested by our respondents for profiling according to Twitter itself. That is, we also profiled hashtags that are most often used to refer to the issue.

In order to analyse issue dynamics over time, we carved up the Twitter data set in four intervals, each of the four intervals, and shows how frequently these words and hashtags co-occur.

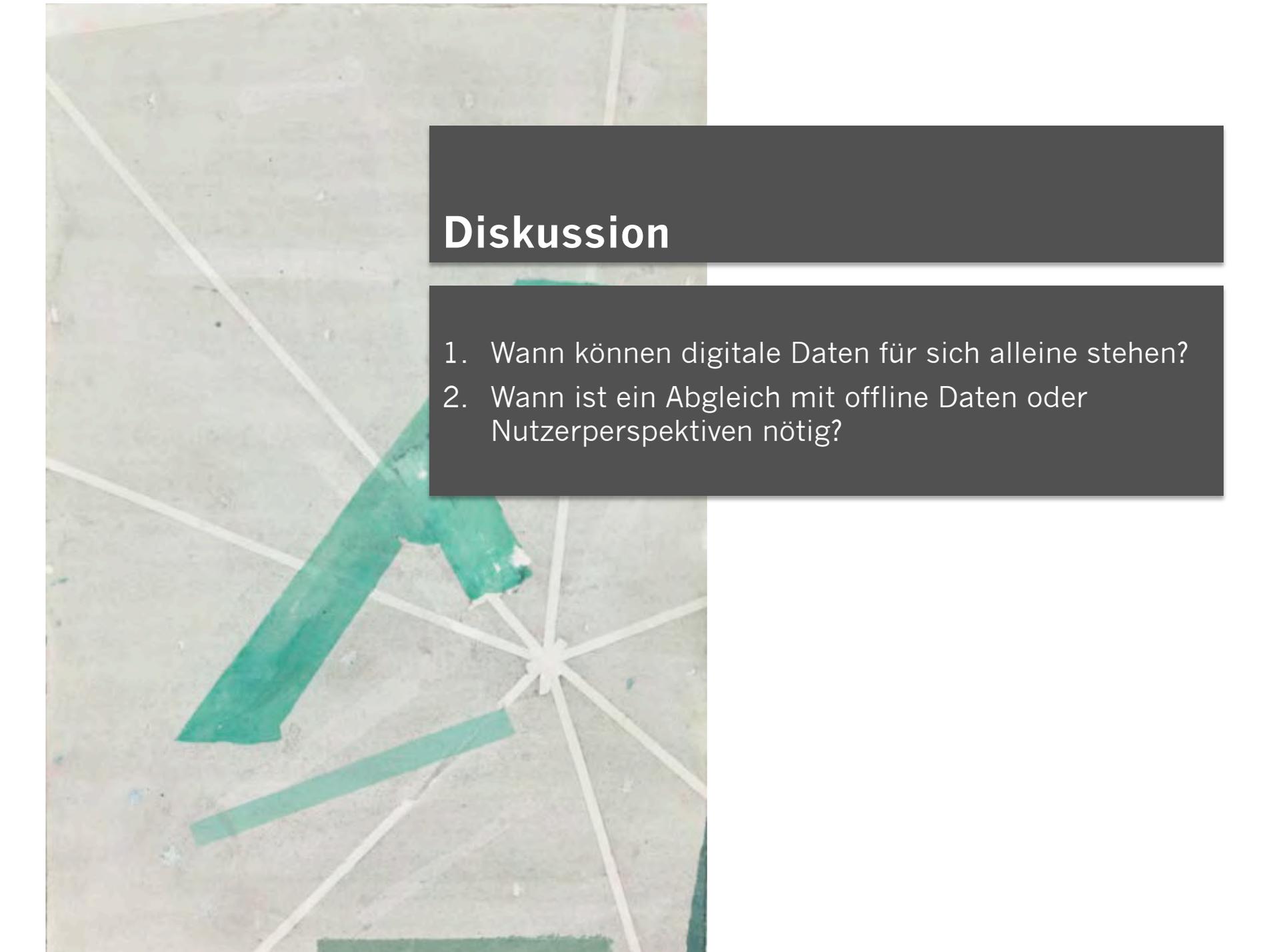
As we worked with Twitter data in this initial study, a key question for us is whether and primarily inform us about the role of Twitter in communicating the summit?

Profiling WCIT Hashtags



- **Medien-Effekte:** Starke Präsenz und Verknüpfung von Kampagnen-Worten.
- **Themen-Effekte:** Anstieg und Abstieg von #netneutrality & #dpi.

- **Online & Offline Abgleich:** Feedback von Experten.
- “Wenig Bezug zwischen Konferenz und Twitter Themen”
- “#netneutrality wurde von NGOs vor der Konferenz fokussiert.”



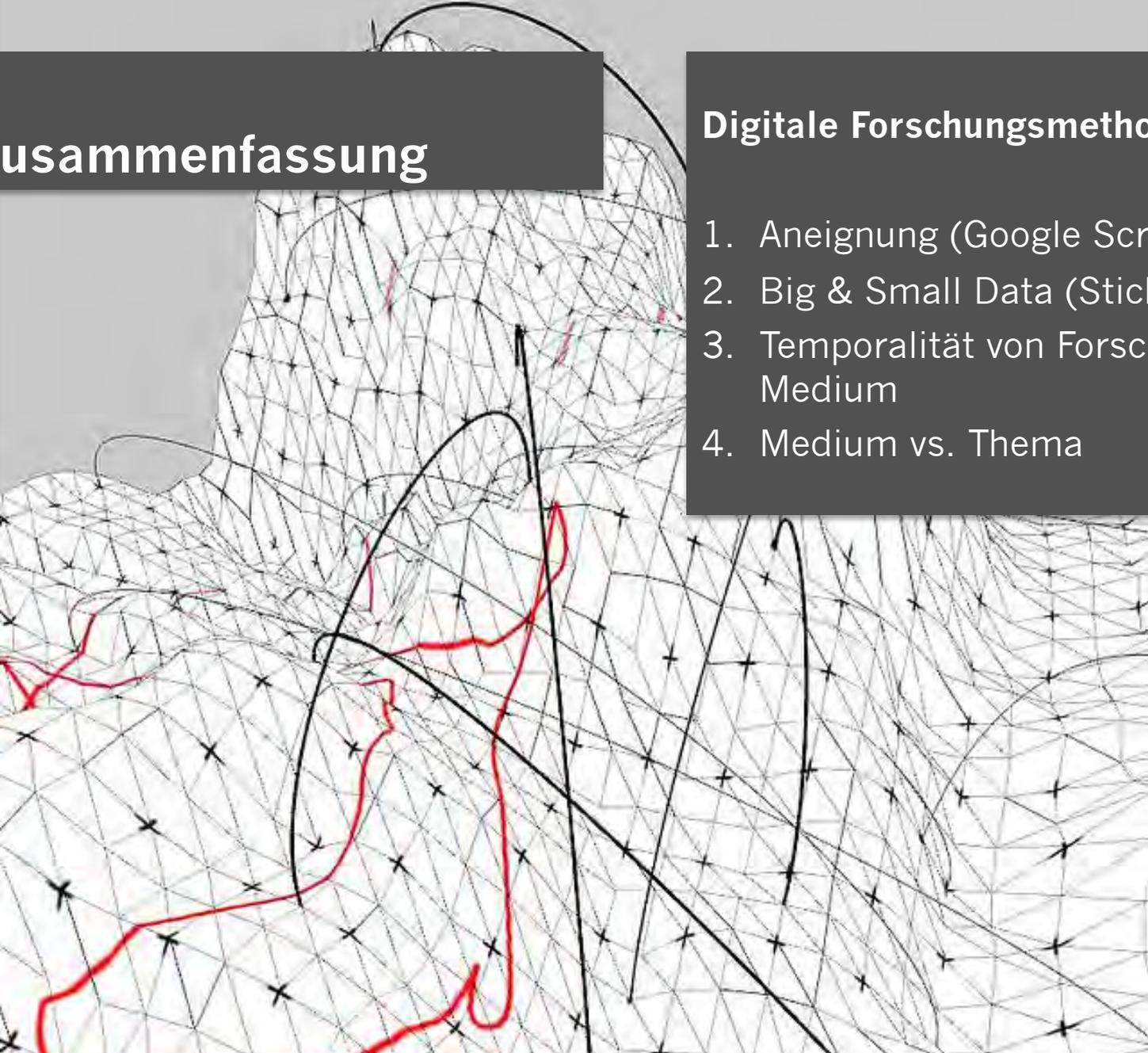
Diskussion

1. Wann können digitale Daten für sich alleine stehen?
2. Wann ist ein Abgleich mit offline Daten oder Nutzerperspektiven nötig?

(Mehr als) medienspezifische Methoden

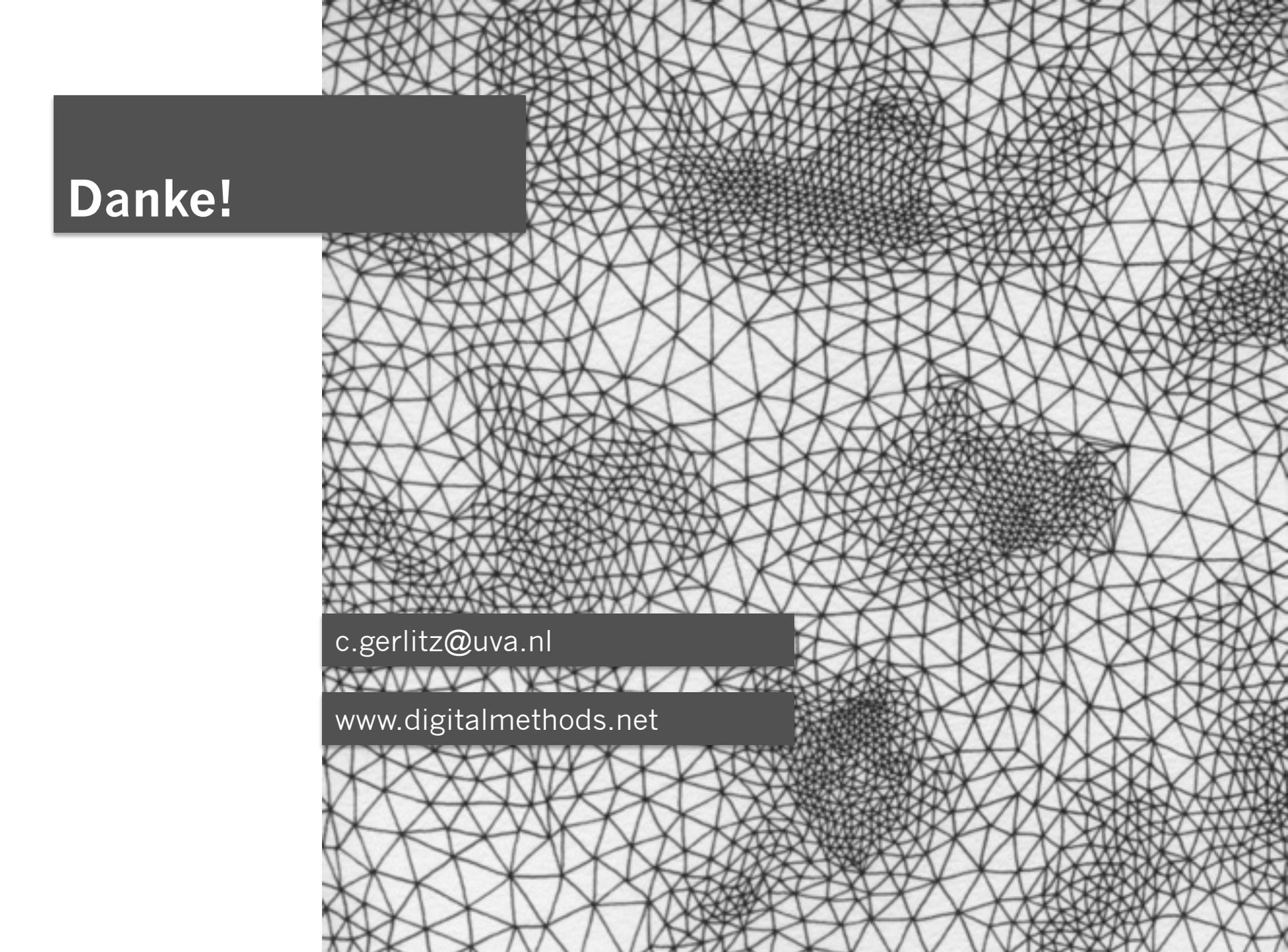
- Digital Methods beruhen nicht nur aus medienspezifischen Methoden.
 - Beinhalten Expertenmeinungen.
 - Abgleich mit offline Daten.
 - Manuelle Kategorisierung und manuelle Testprojekte.
-
- Digital Methoden als **Interface Methoden**, als Schnittstelle zwischen Disziplinen und Kompetenzen.

Zusammenfassung



Digitale Forschungsmethoden

1. Aneignung (Google Scraper)
2. Big & Small Data (Stichproben)
3. Temporalität von Forschung und Medium
4. Medium vs. Thema



Danke!

c.gerlitz@uva.nl

www.digitalmethods.net