

SPIEGEL MINING

Was Big Data ist.

DAVID KRIESEL

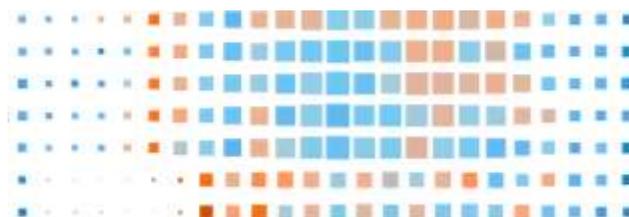
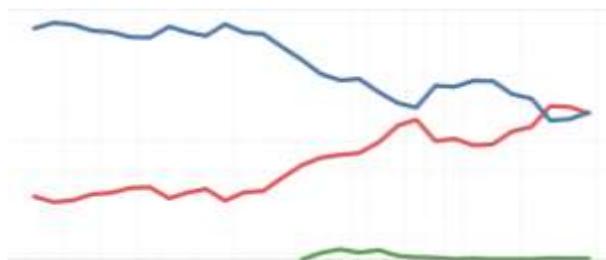
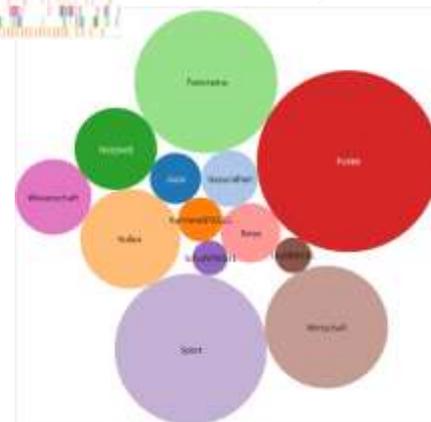
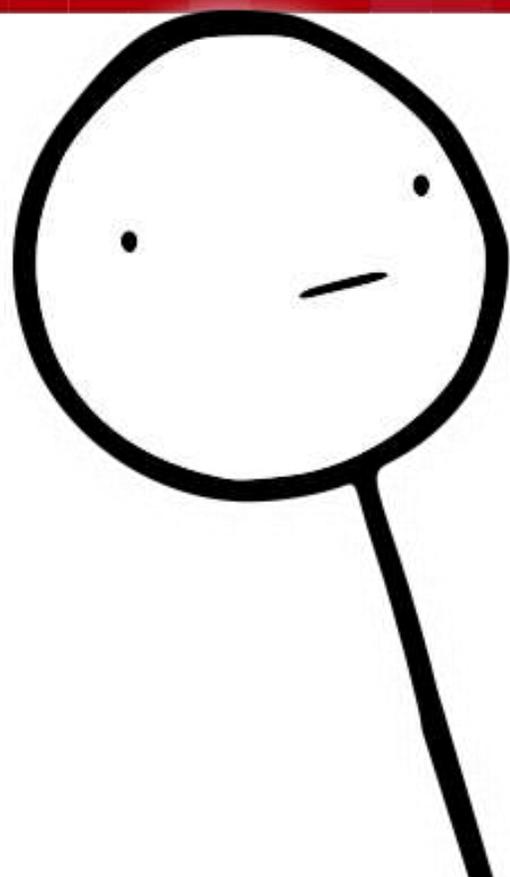
DKRIESEL.COM

FACHTAG, NEULAND BFI, BREMEN, 23. NOVEMBER 2017

SÄMTLICHE UNGEKENNZEICHNETEN ABBILDUNGEN STAMMEN VON MIR SELBST, SIND FREI VERFÜGBAR ODER SIND BROWERSCREENSHOTS.

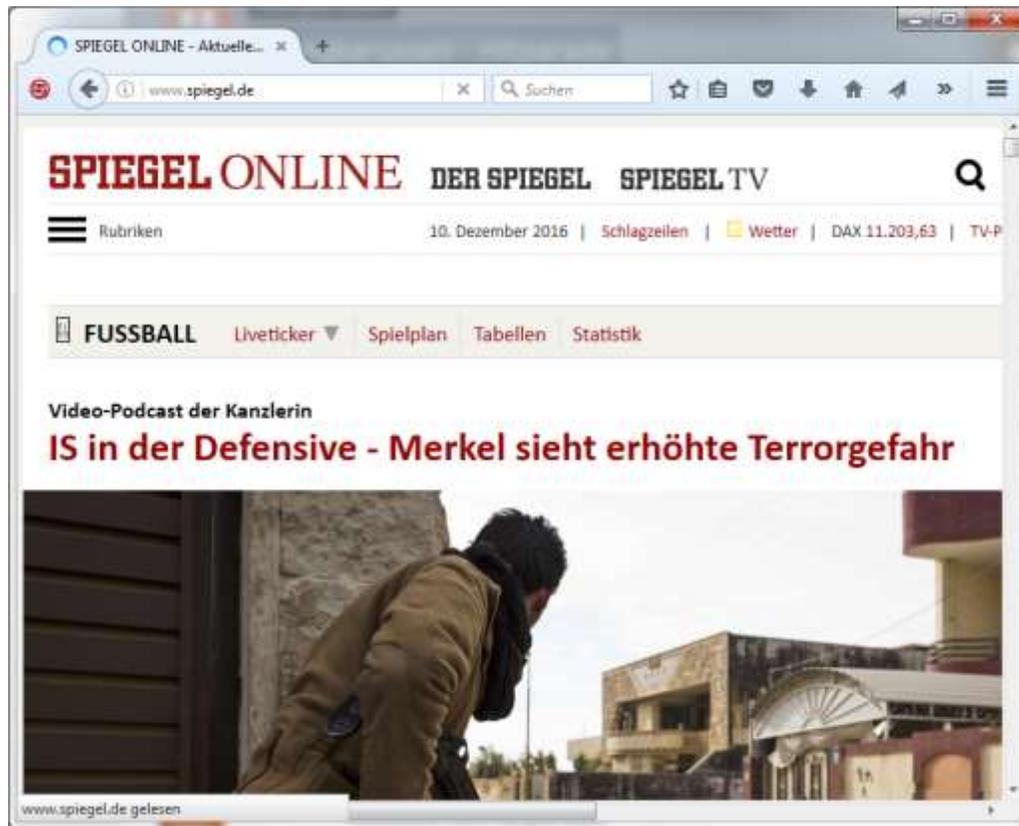
SPIEGEL

MINING



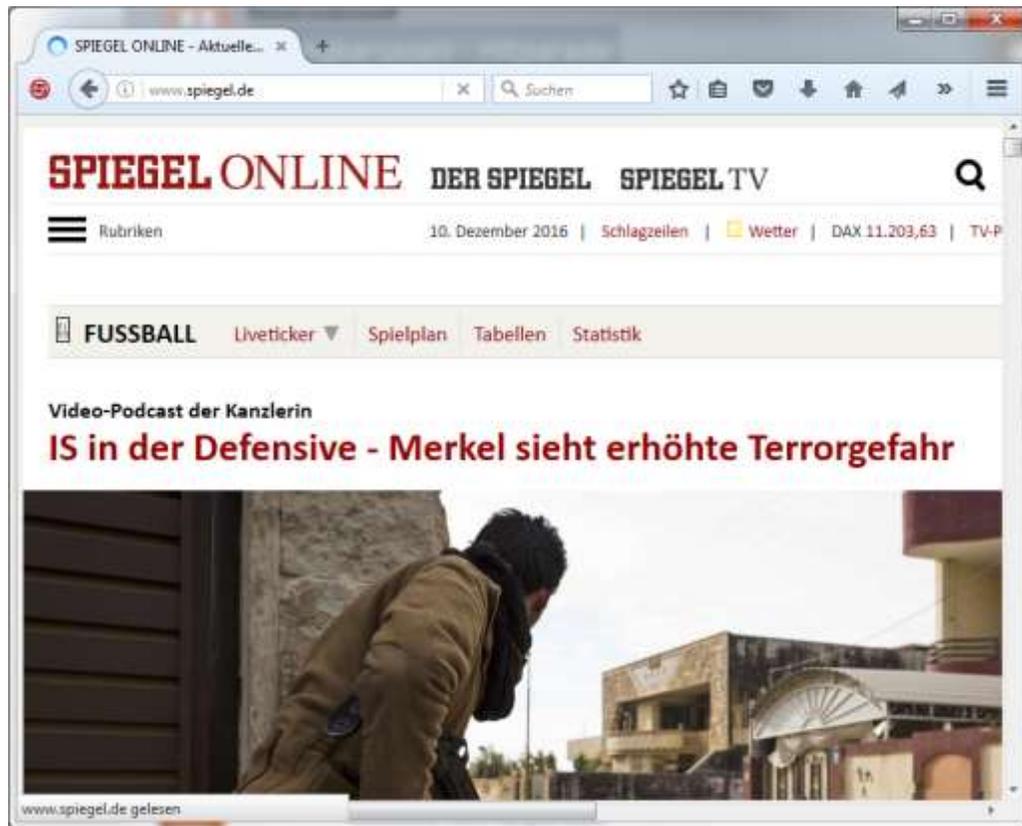
Was ist Spiegelmining?

SpiegelOnline

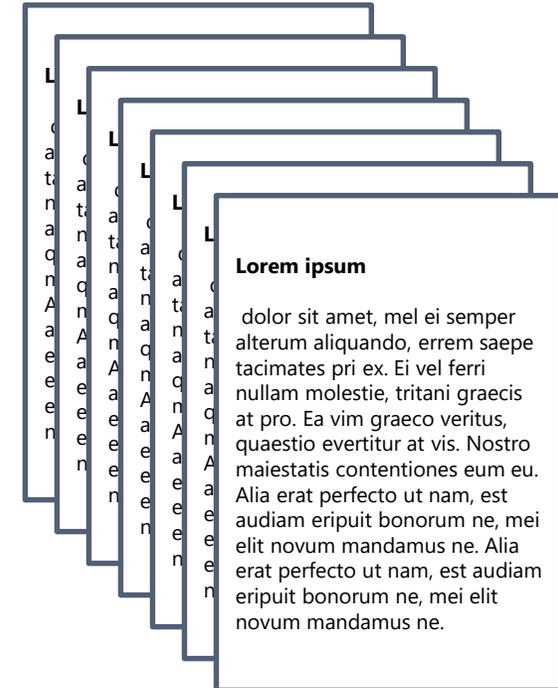
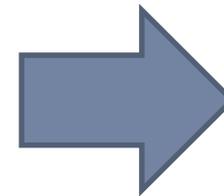


Was ist Spiegelmining?

SpiegelOnline



Artikelsammlung

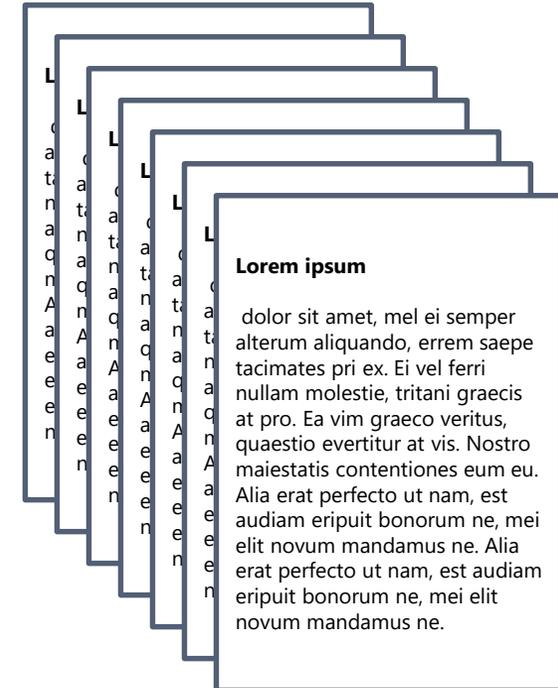
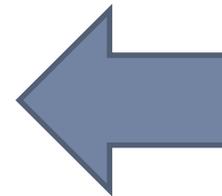


Was ist Spiegelmining?

Features

Rubrik	Veröffentlichung
Panorama	2014-07-16 13:46:23
Netzwelt	2015-01-19 09:17:33
Politik	2015-01-14 16:30:55
Sport	2014-09-30 13:41:21
Sport	2014-07-07 11:06:12
Panorama	2014-08-02 09:23:59
...	...

Artikelsammlung

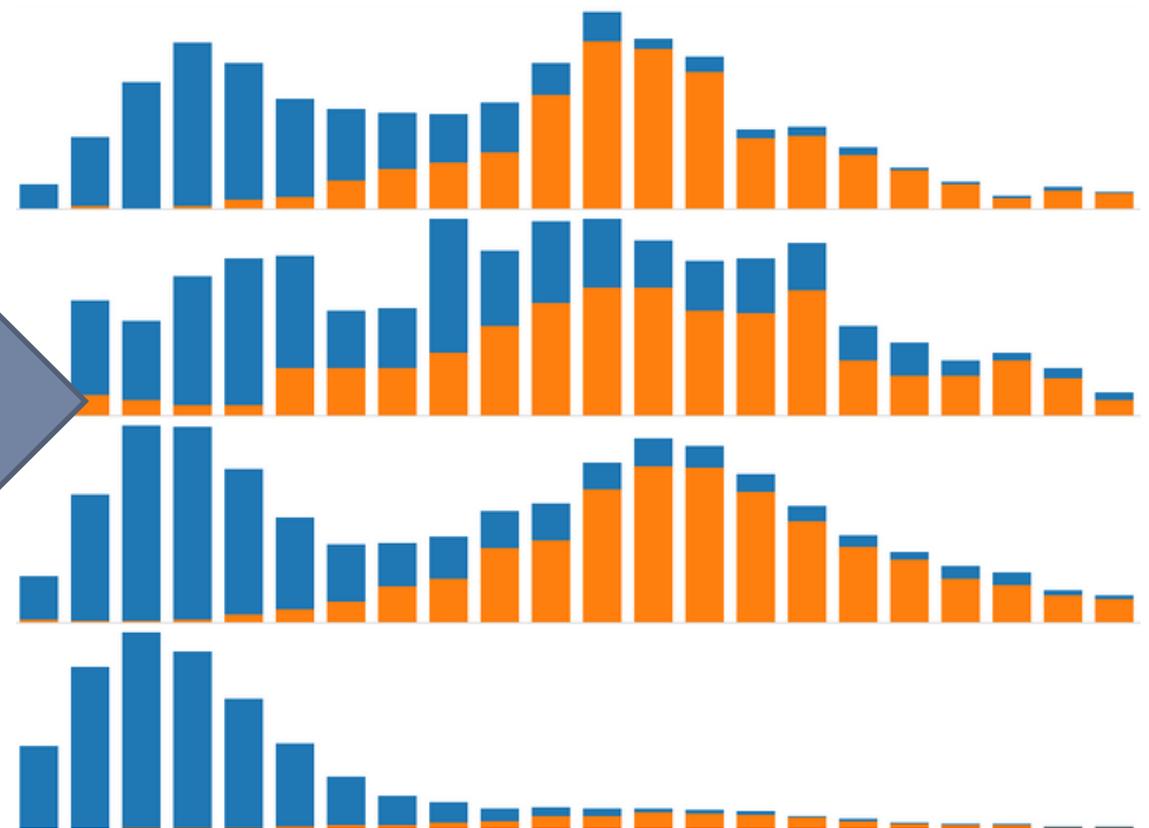


Was ist Spiegelmining?

Features

Rubrik	Veröffentlichung
Panorama	2014-07-16 13:46:23
Netzwelt	2015-01-19 09:17:33
Politik	2015-01-14 16:30:55
Sport	2014-09-30 13:41:21
Sport	2014-07-07 11:06:12
Panorama	2014-08-02 09:23:59
...	...

Auswertungen



Was ist Spiegelmining?

Blogartikel-Serie



News [D. Kriesel]

www.dkriesel.com/start?btng[]

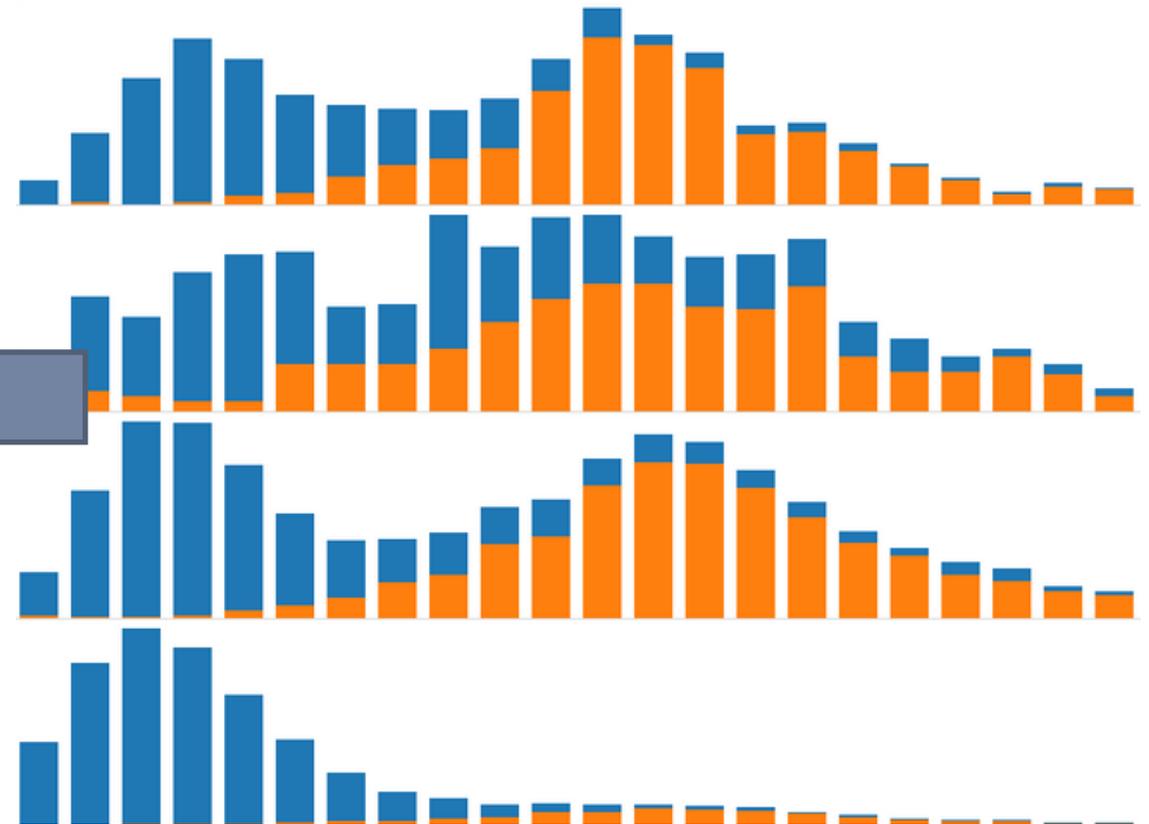
Suchen

SpiegelMining. Auch Spiegelredakteure feiern Weihnachten. Eine Analyse von 70.000 SpiegelOnline-Artikeln

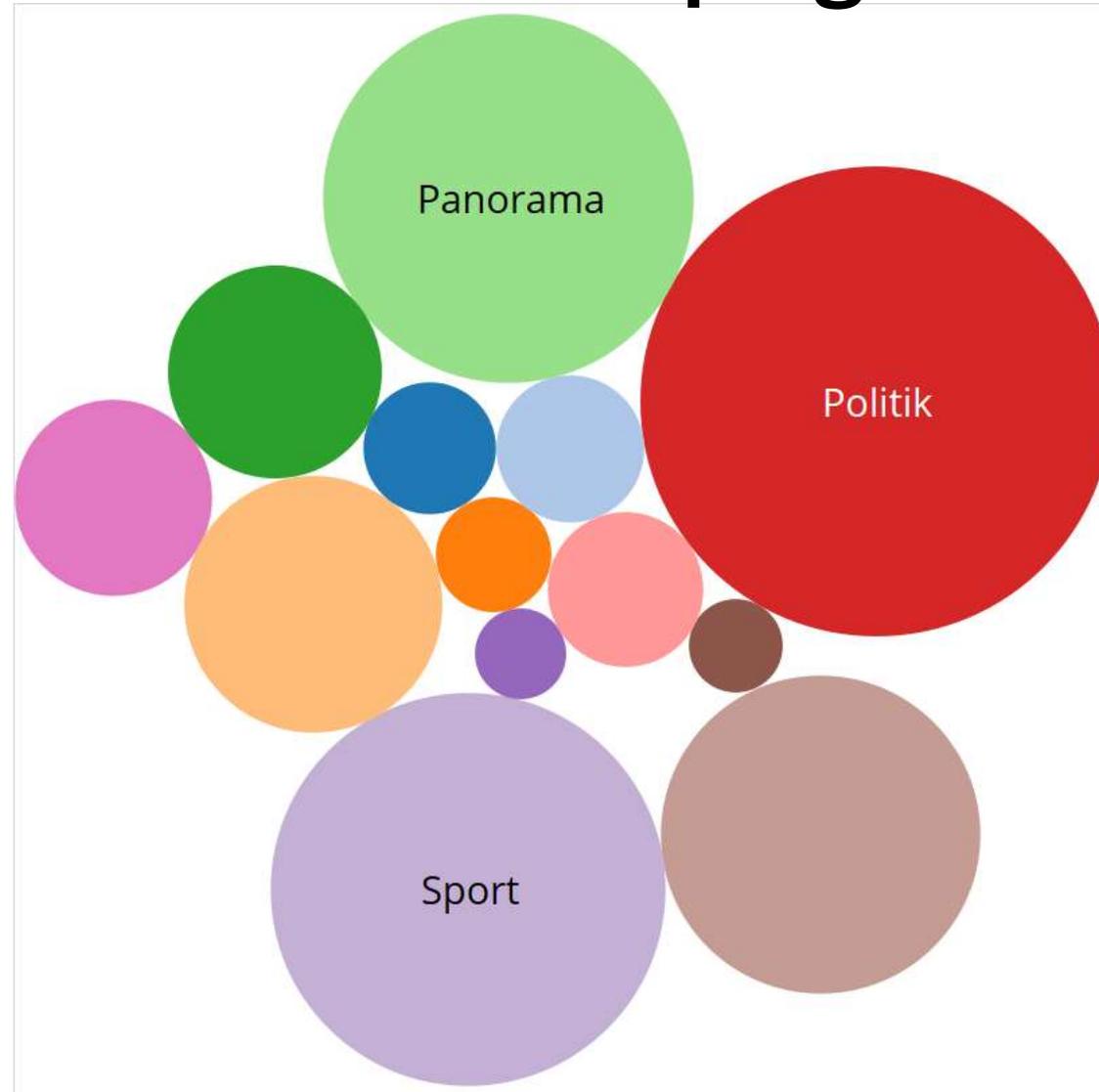
SPIEGEL MINING

Seit Mitte 2014 habe ich mehr als 70.000 Artikel von SpiegelOnline systematisch gespeichert. Jeden Tag kommen im Schnitt 100 dazu. Diese Artikelmasse werden wir in der nächsten Zeit auswerten und erforschen. Was herauskommt, ist eine tiefgreifende Analyse des Publikationsverhaltens des vielleicht größten Meinungsmachers Deutschlands.

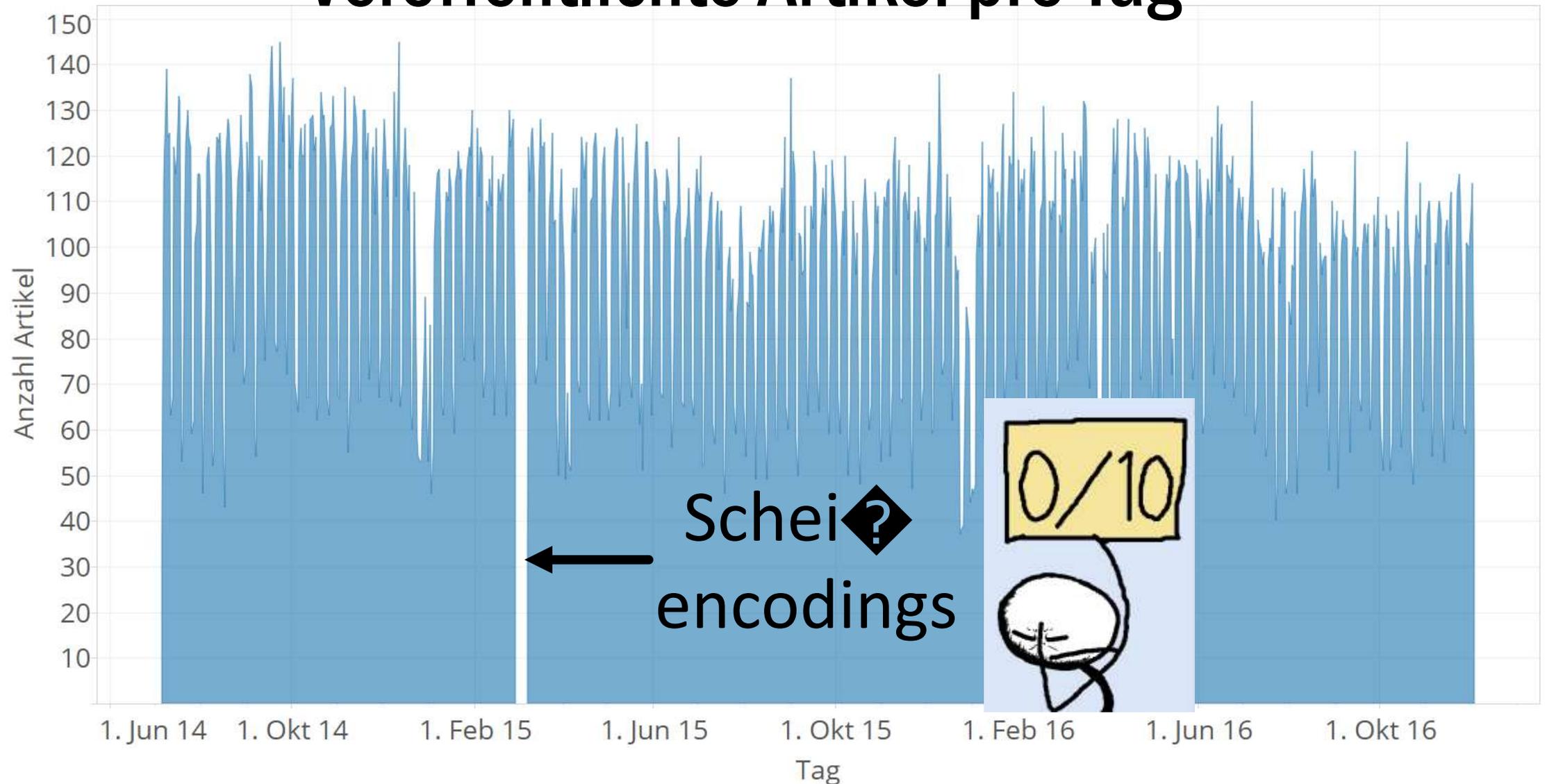
Auswertungen



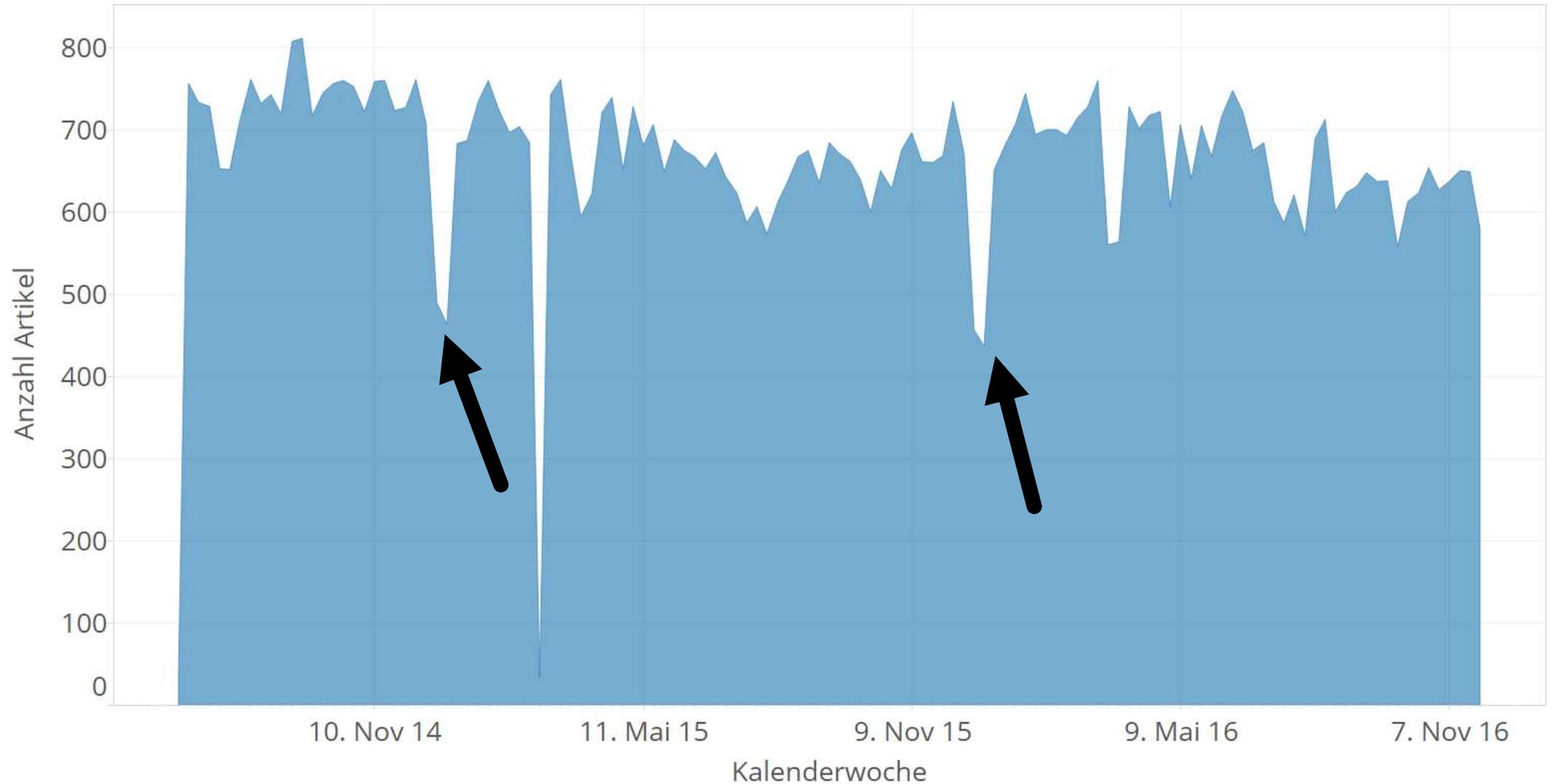
Die Rubriken von SpiegelOnline



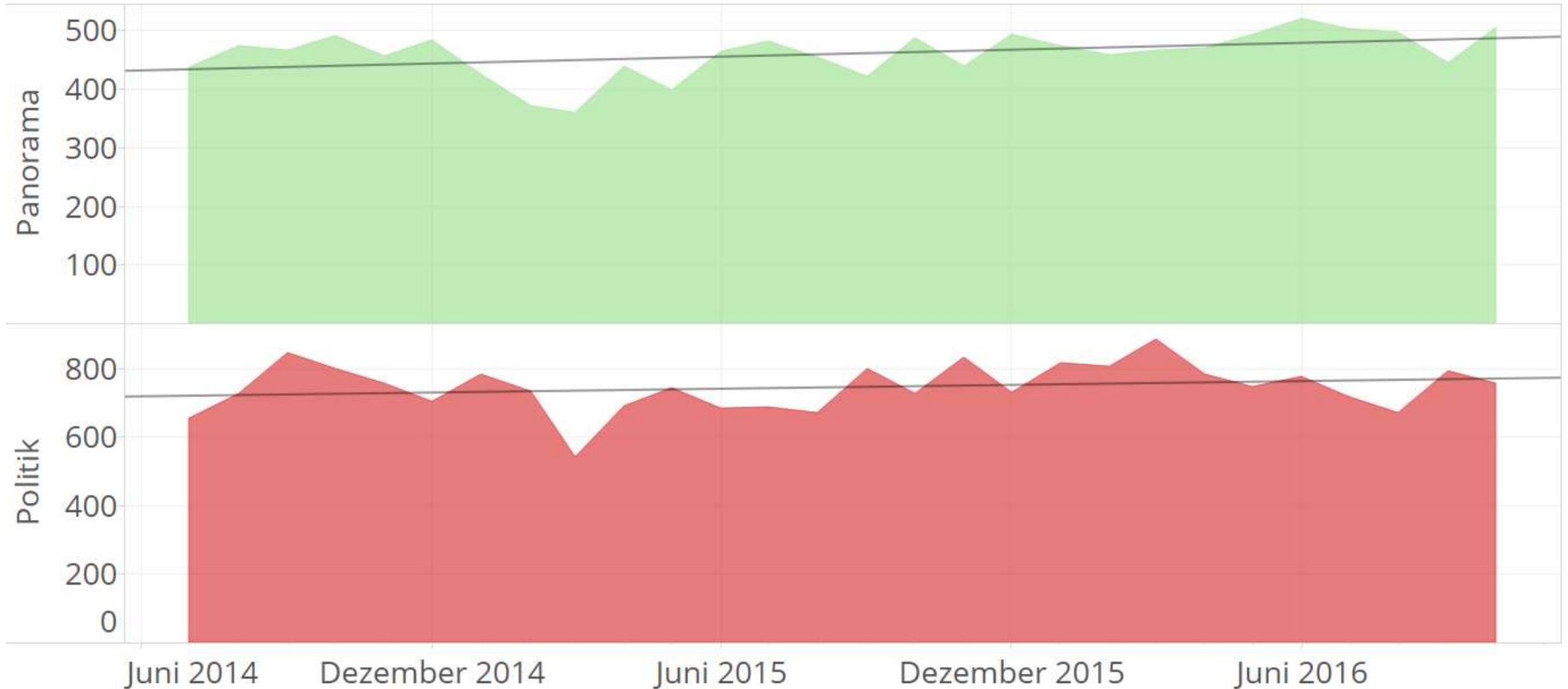
Veröffentlichte Artikel pro Tag



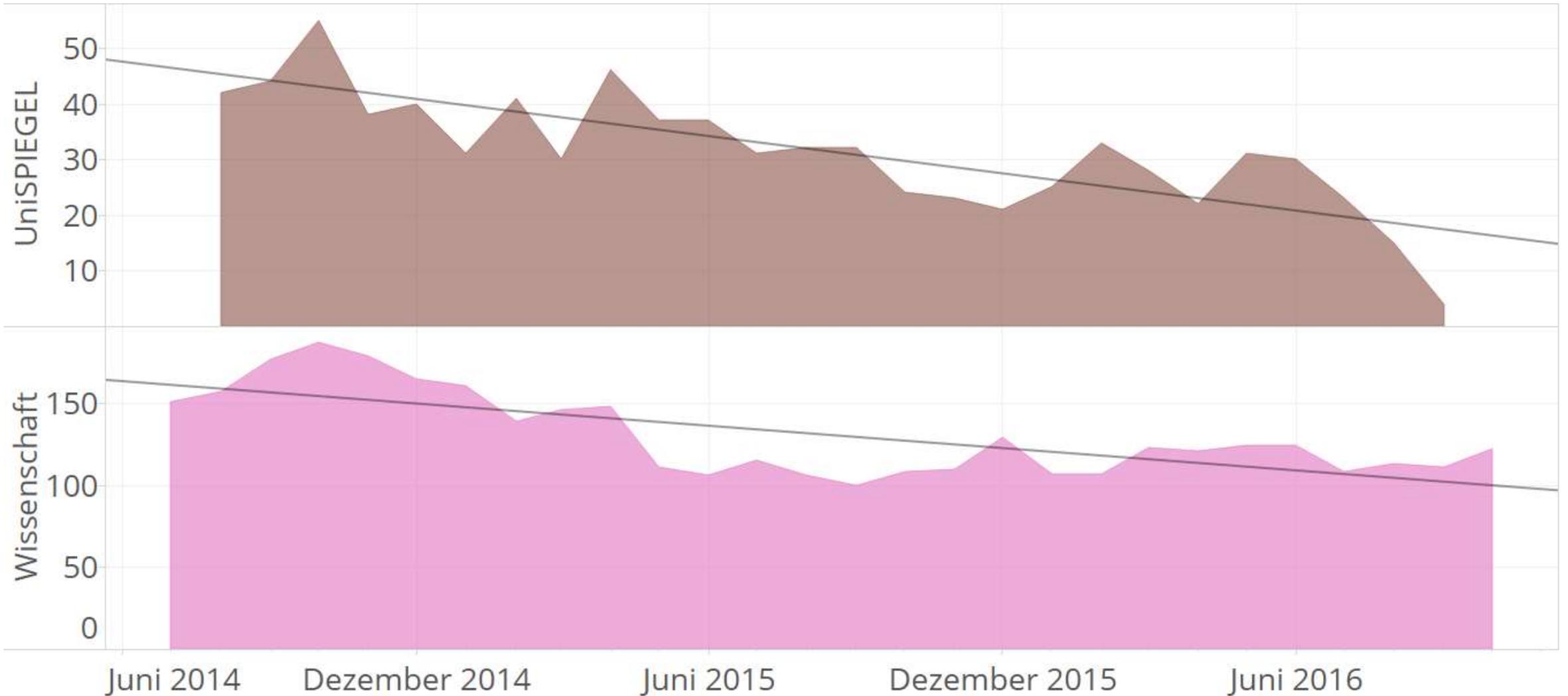
Veröffentlichte Artikel pro Woche



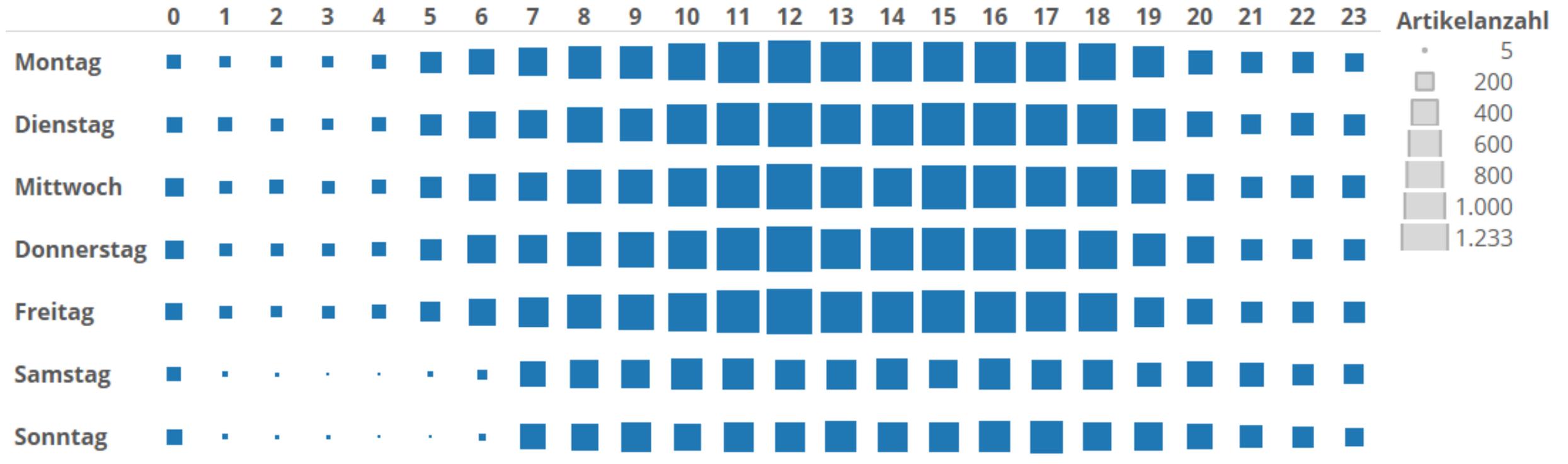
Panorama und Politik veröffentlichen stabil ...



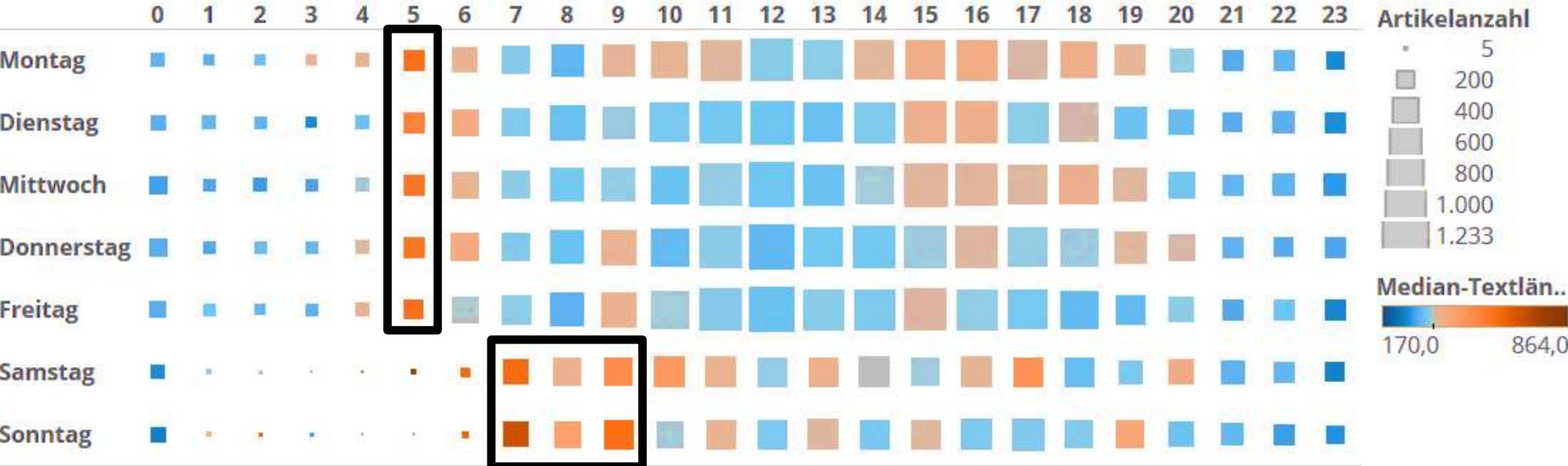
... andere Ressorts werden gerupft.



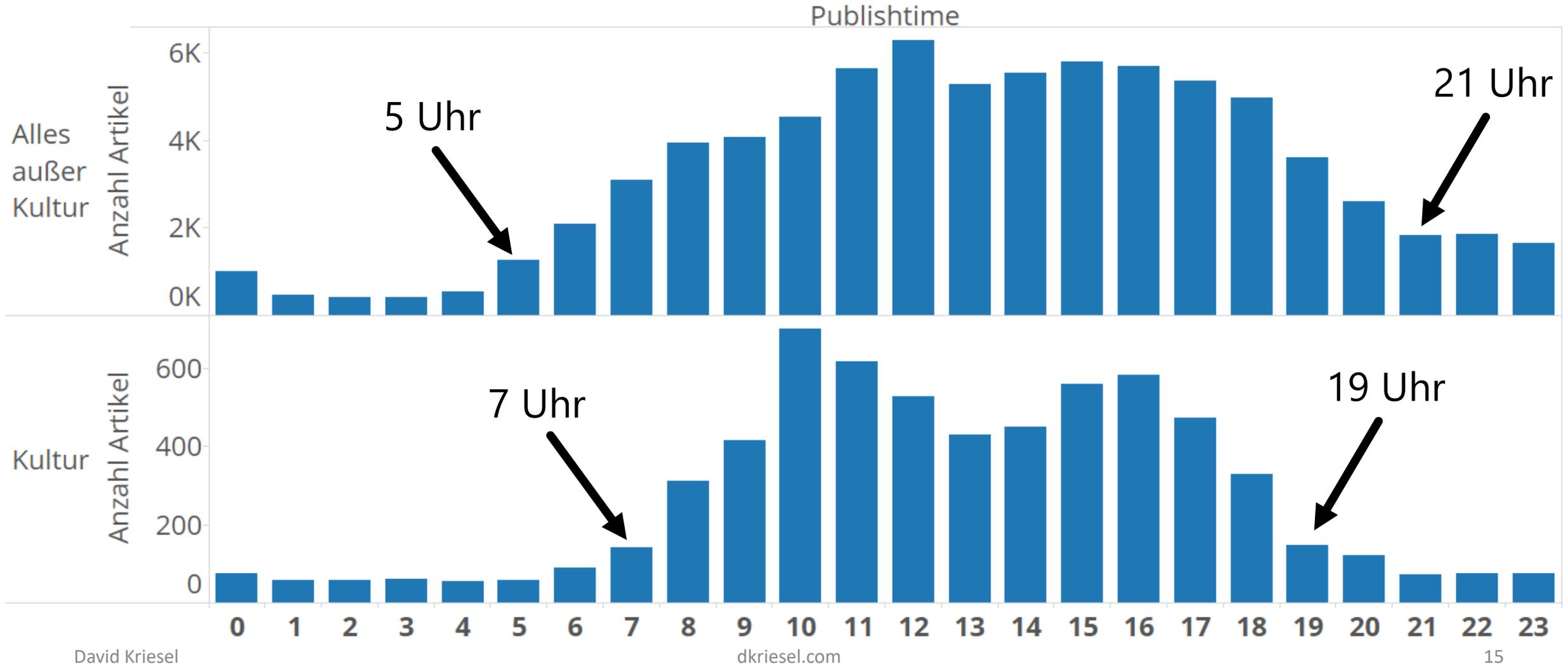
Veröffentlichungen nach Wochentag und Zeit

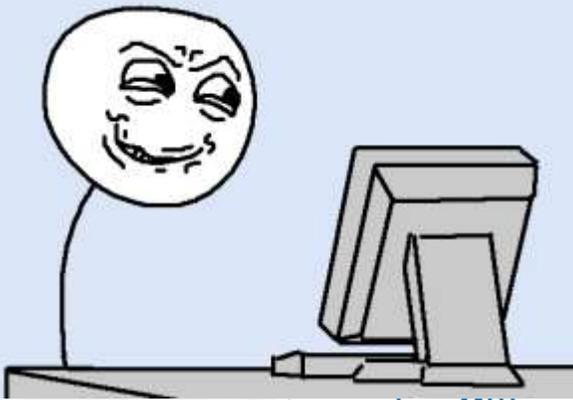


Veröffentlichungen/Textlänge nach Wochentag und Zeit



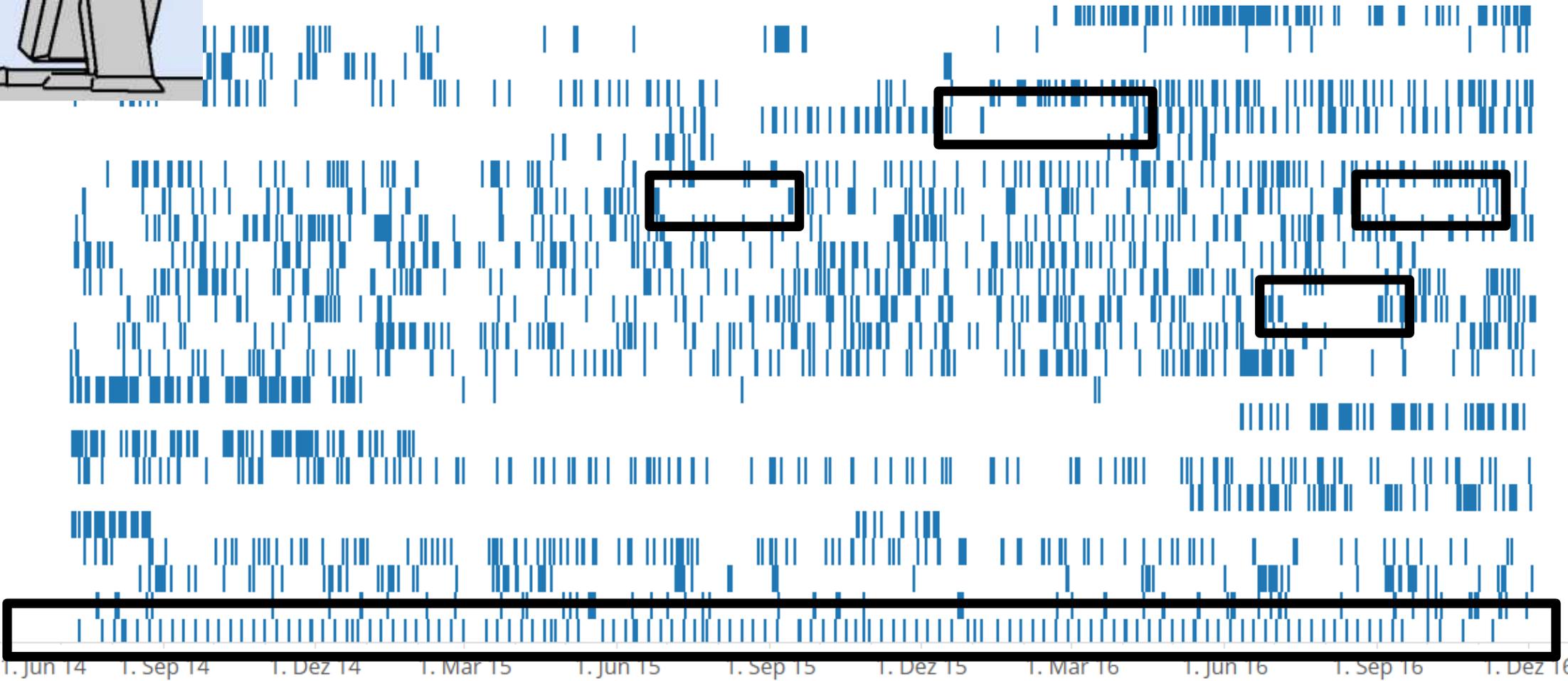
Artikel nach Zeit (Kultur)



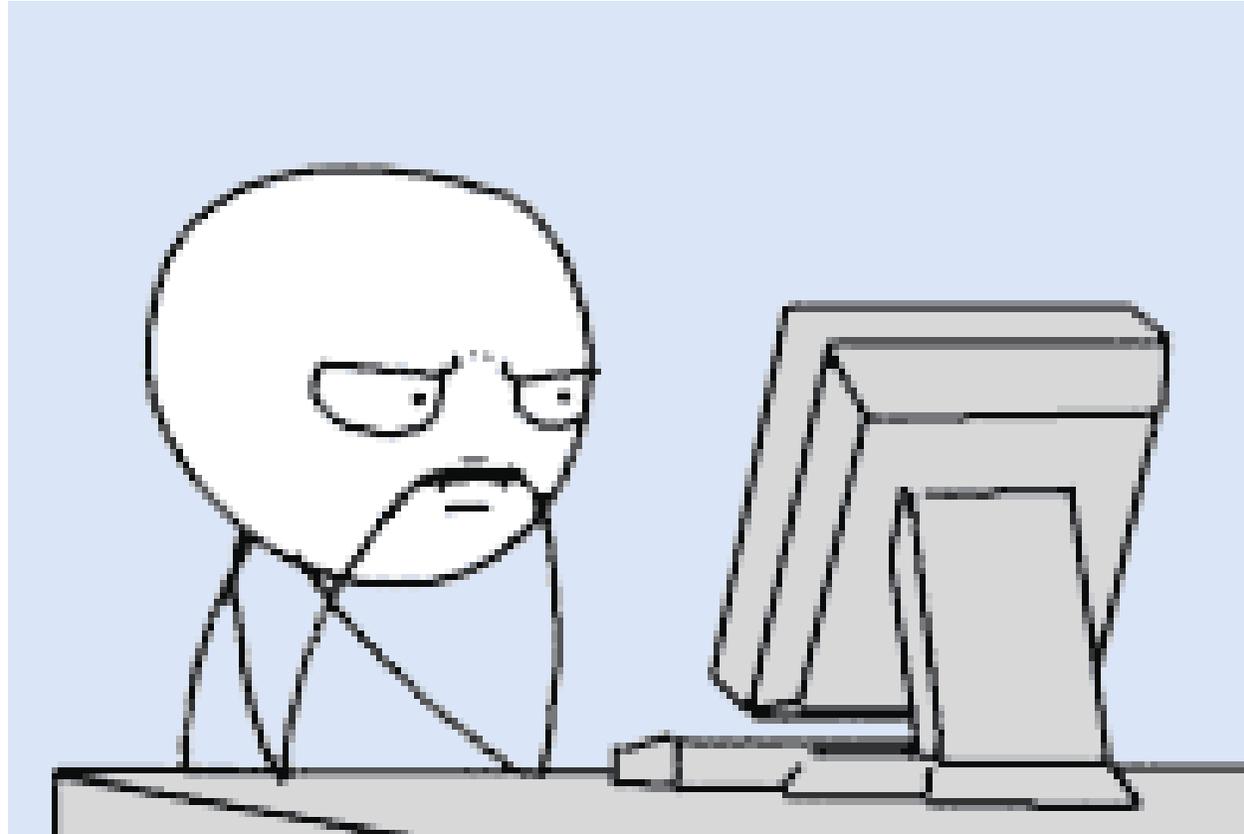


Wer, wann, was, mit wem?

Blurred text on the left side of the slide, likely representing a list of names or identifiers.



Wir alle, wann, was, mit wem?



Artikel haben Keywords

Politik, Ausland, Saudi-Arabien, König
Salman von Saudi-Arabien

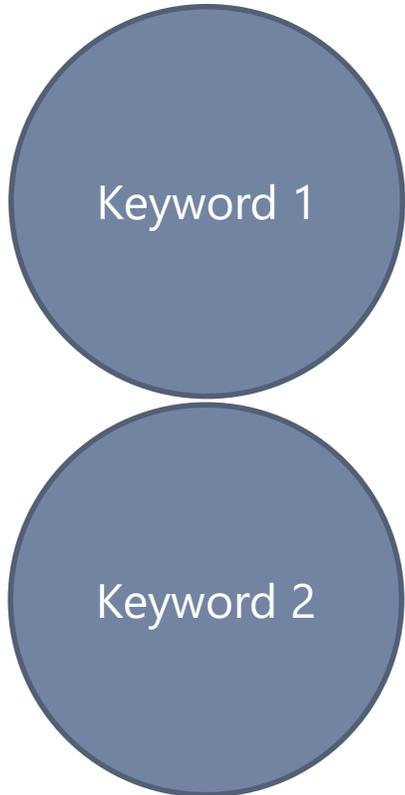


Panorama, Drogen in Kolumbien,
Kolumbien, Kokain, Drogen, Farc



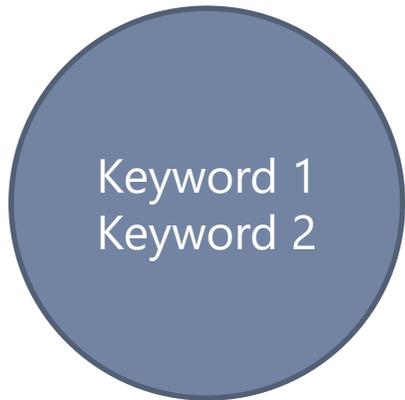
Keyword-Verwandtschaften

Verwandt



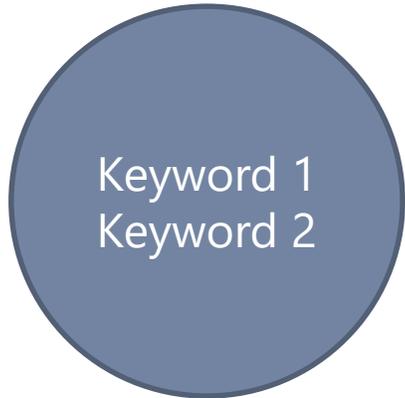
Keyword-Verwandtschaften

Verwandt

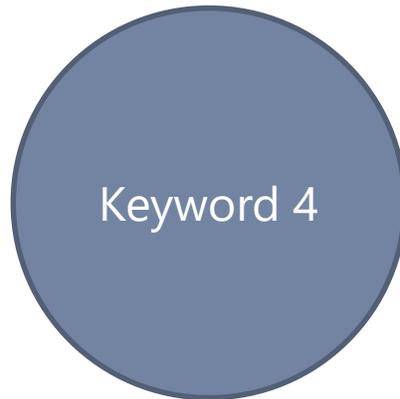
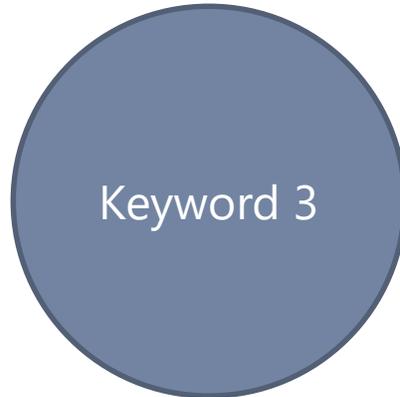


Keyword-Verwandtschaften

Verwandt

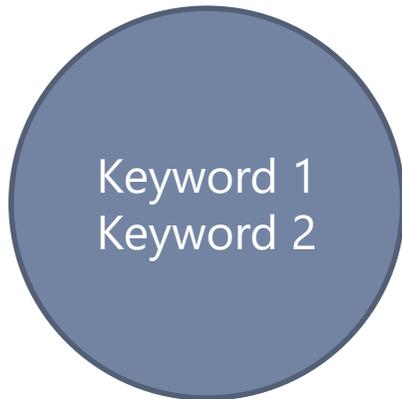


Unverwandt

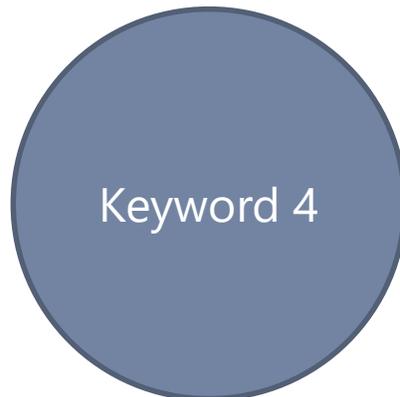
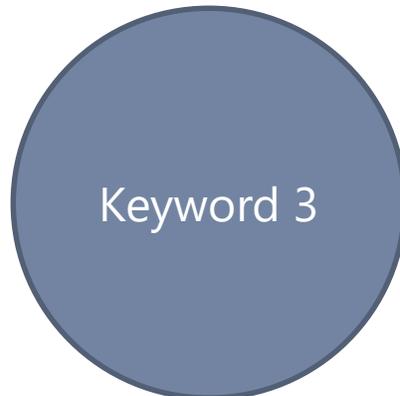


Keyword-Verwandtschaften

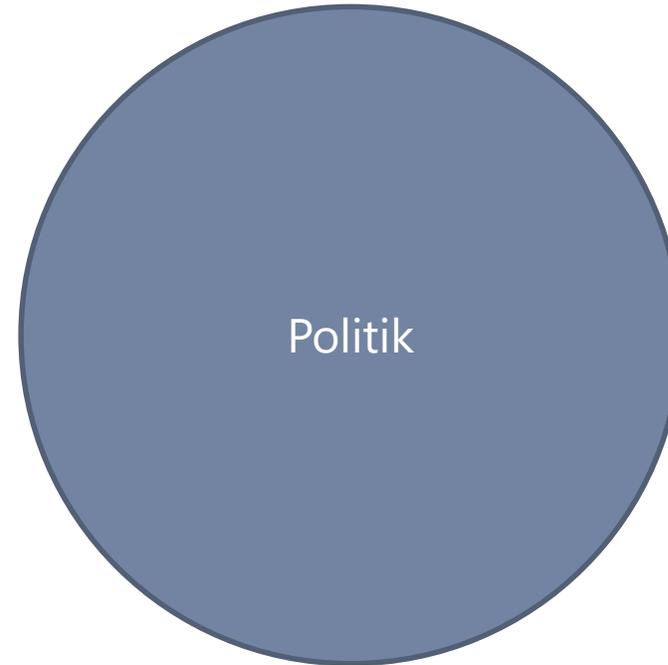
Verwandt



Unverwandt

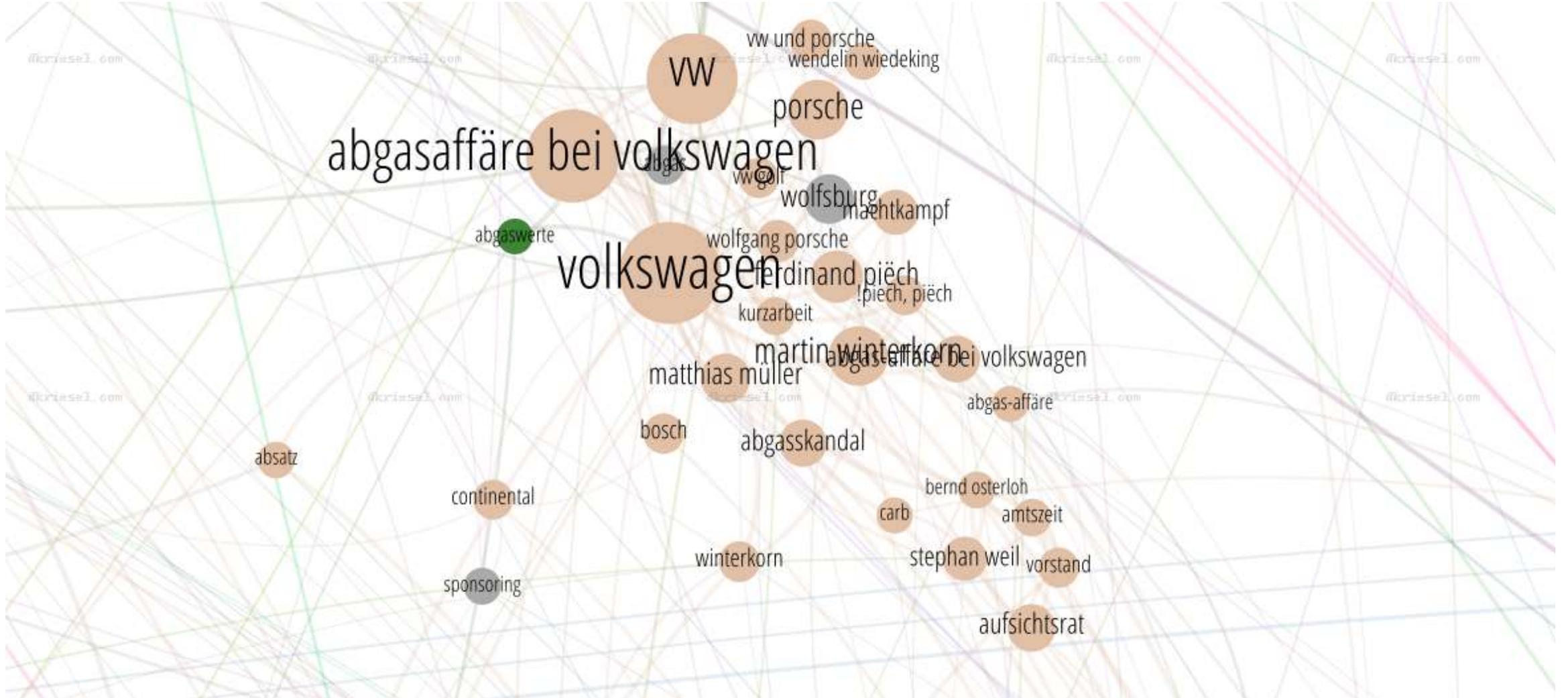


Mittelverwandt

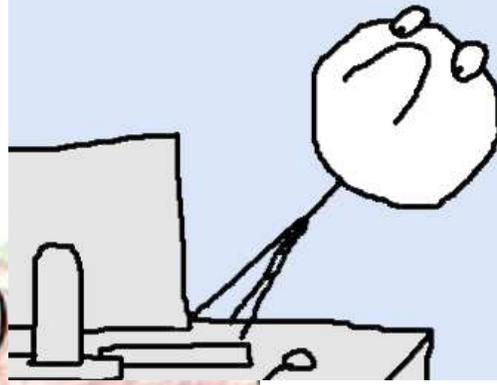
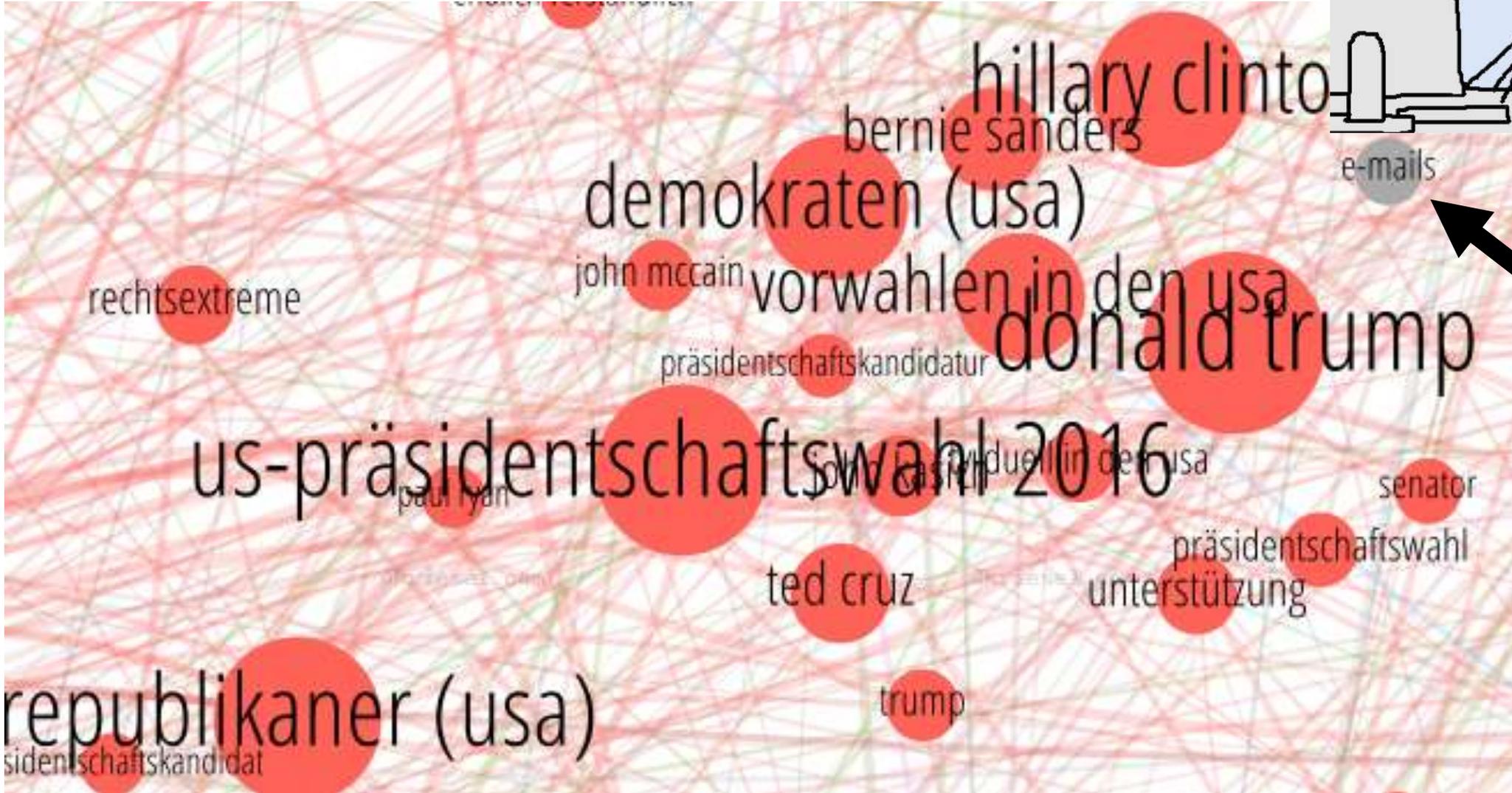


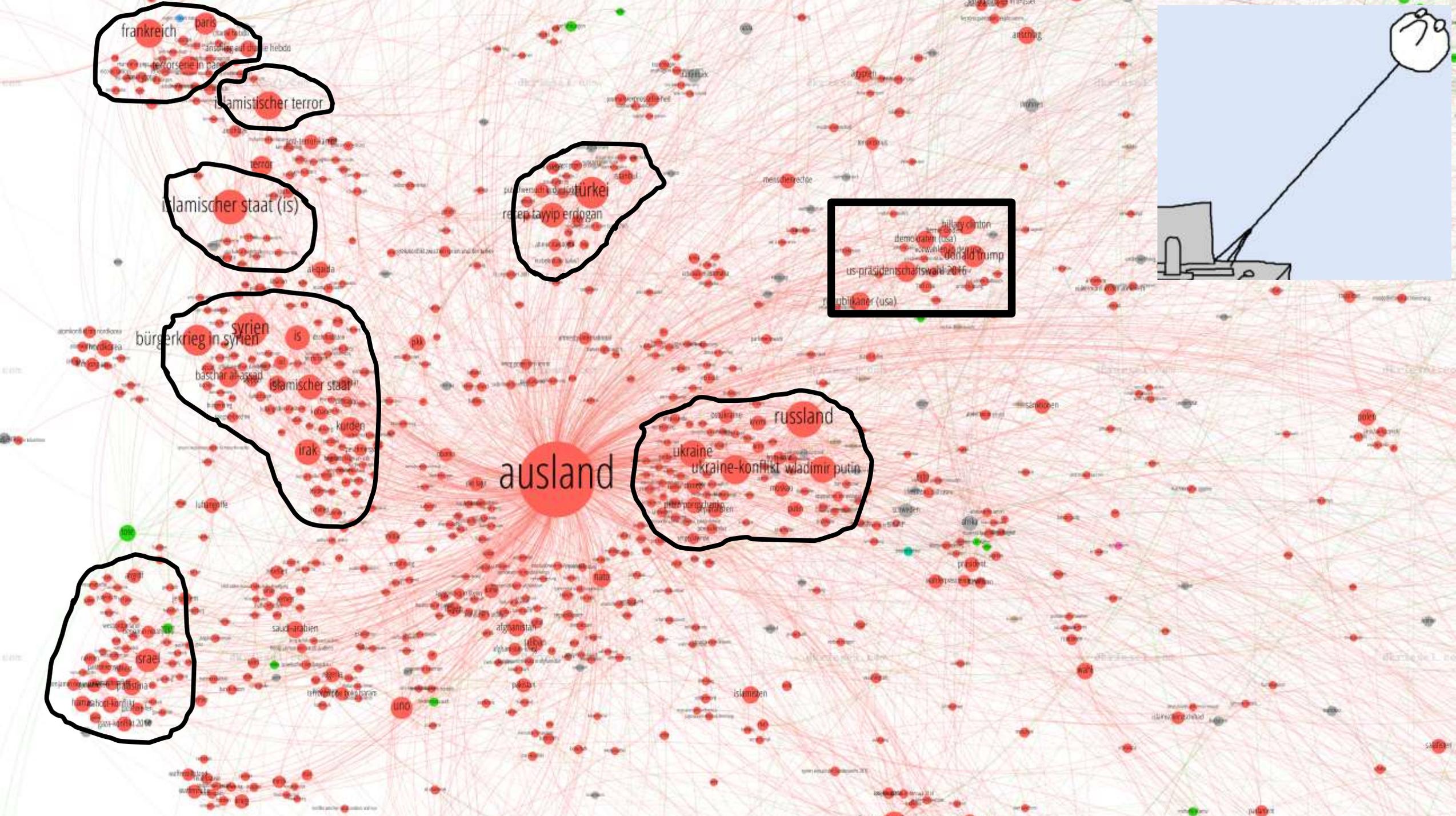


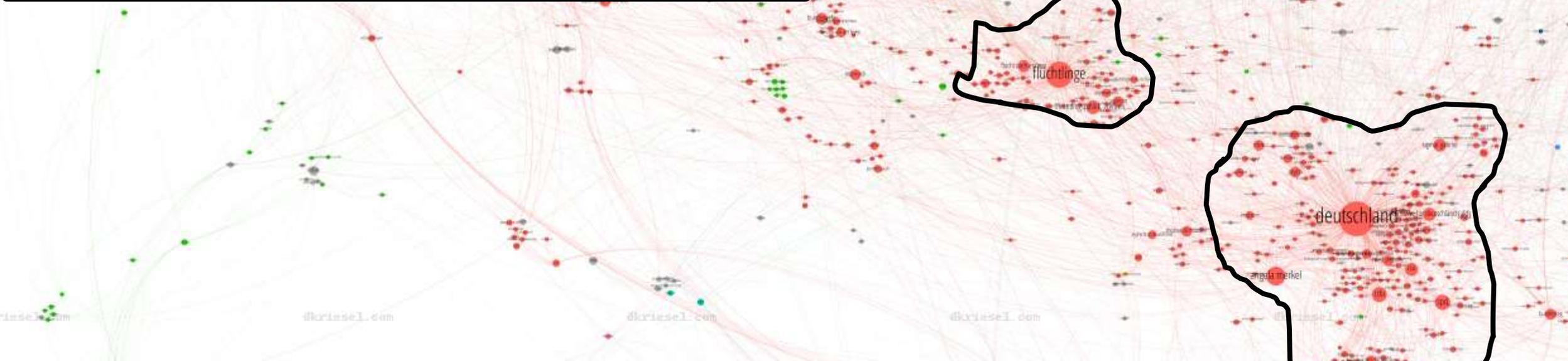
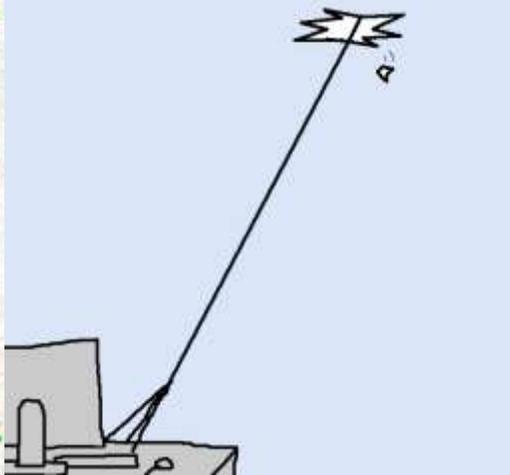
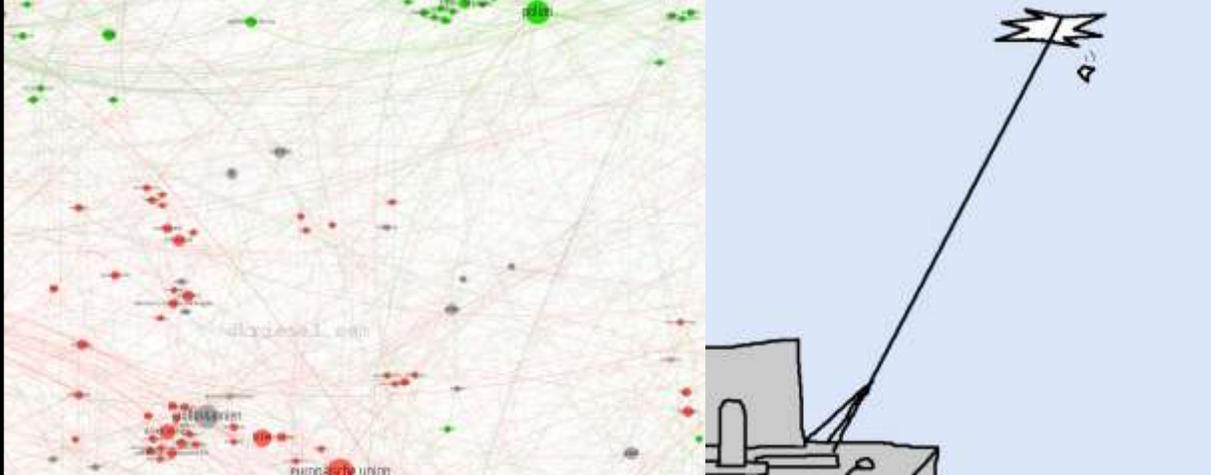
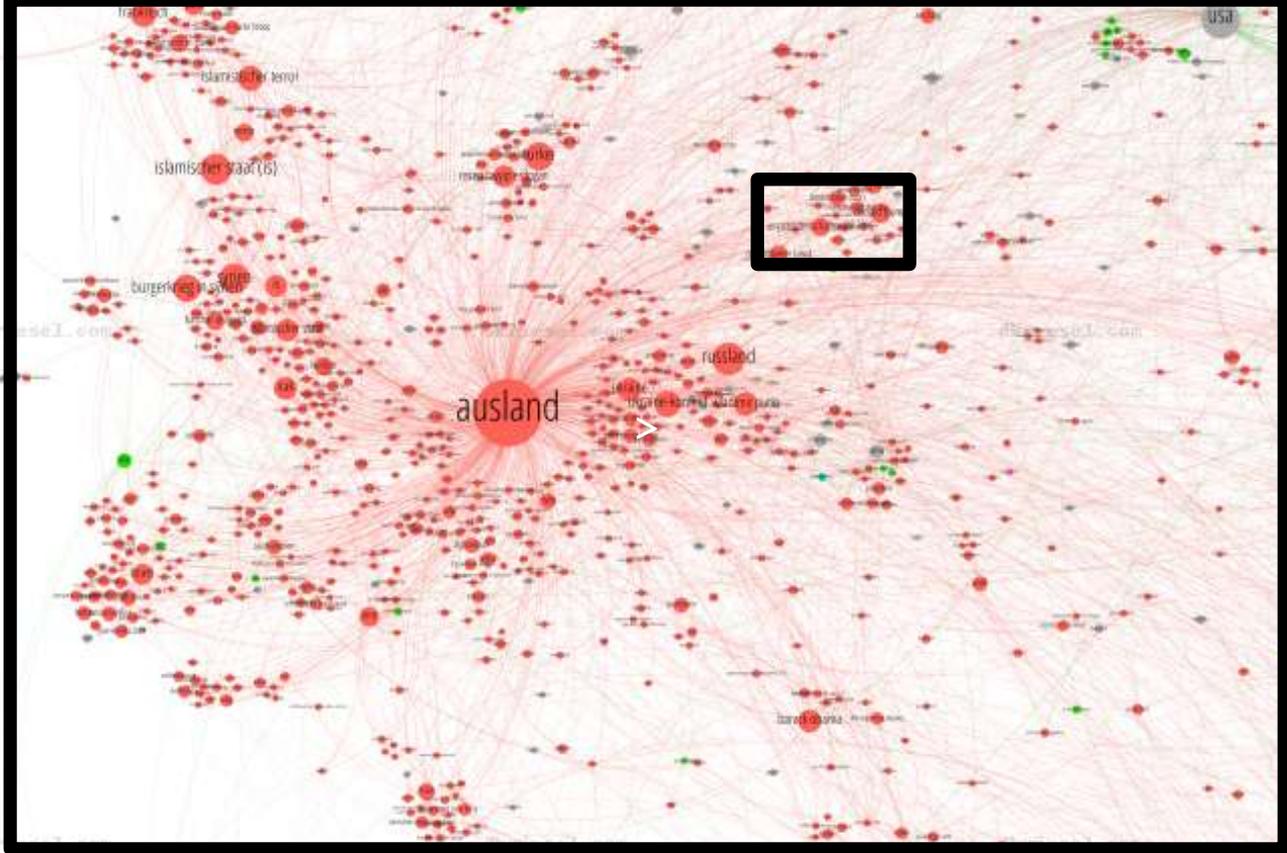
Die VW-Abgas-Affäre

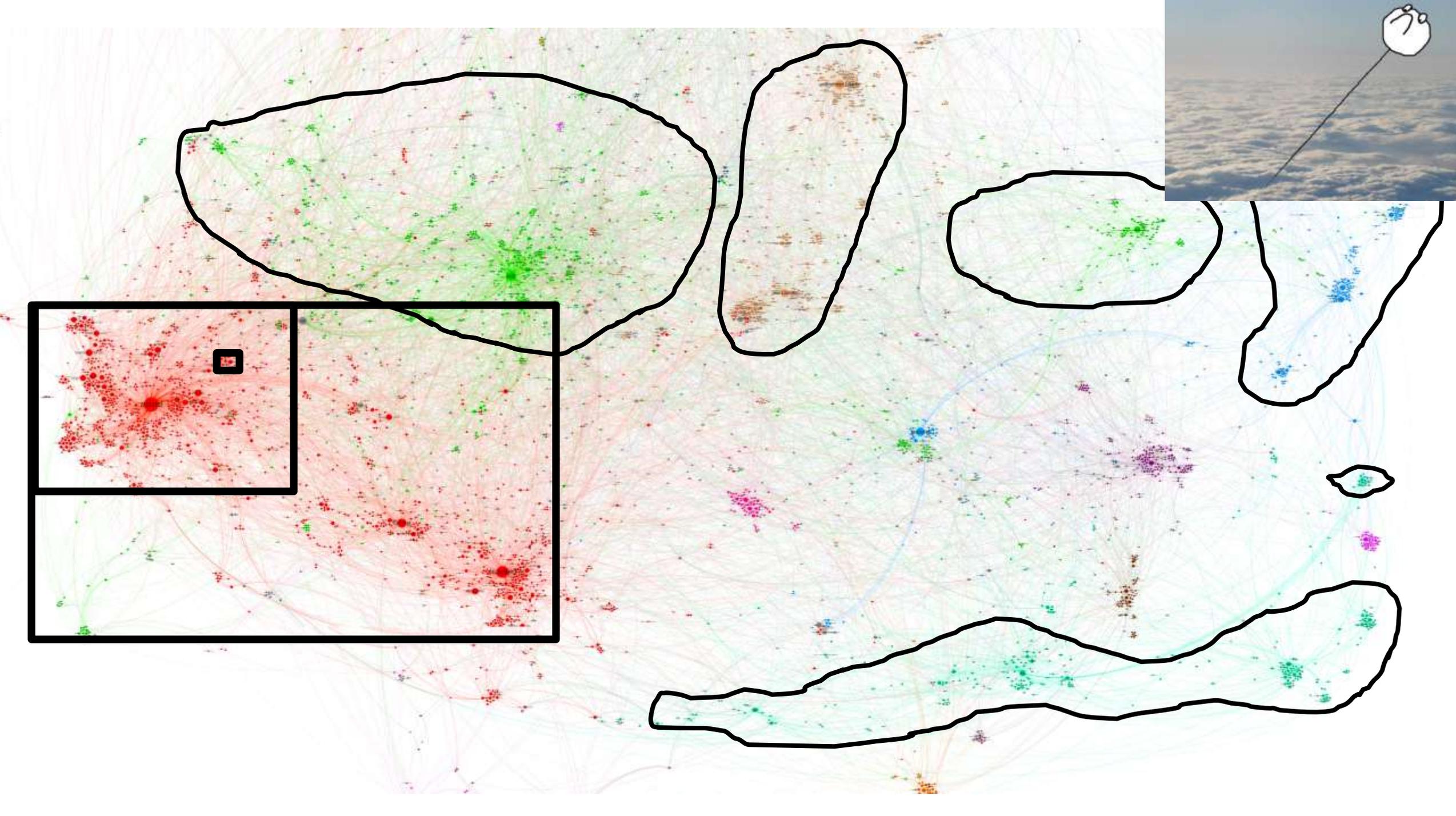


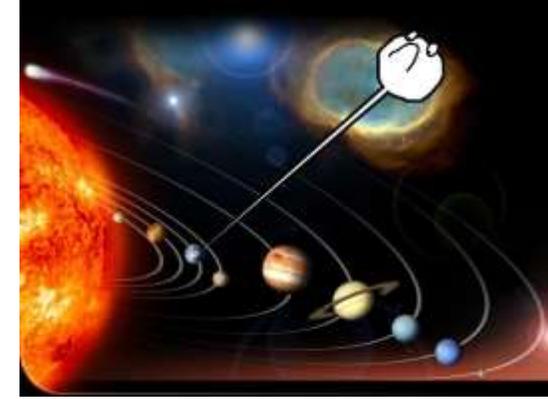
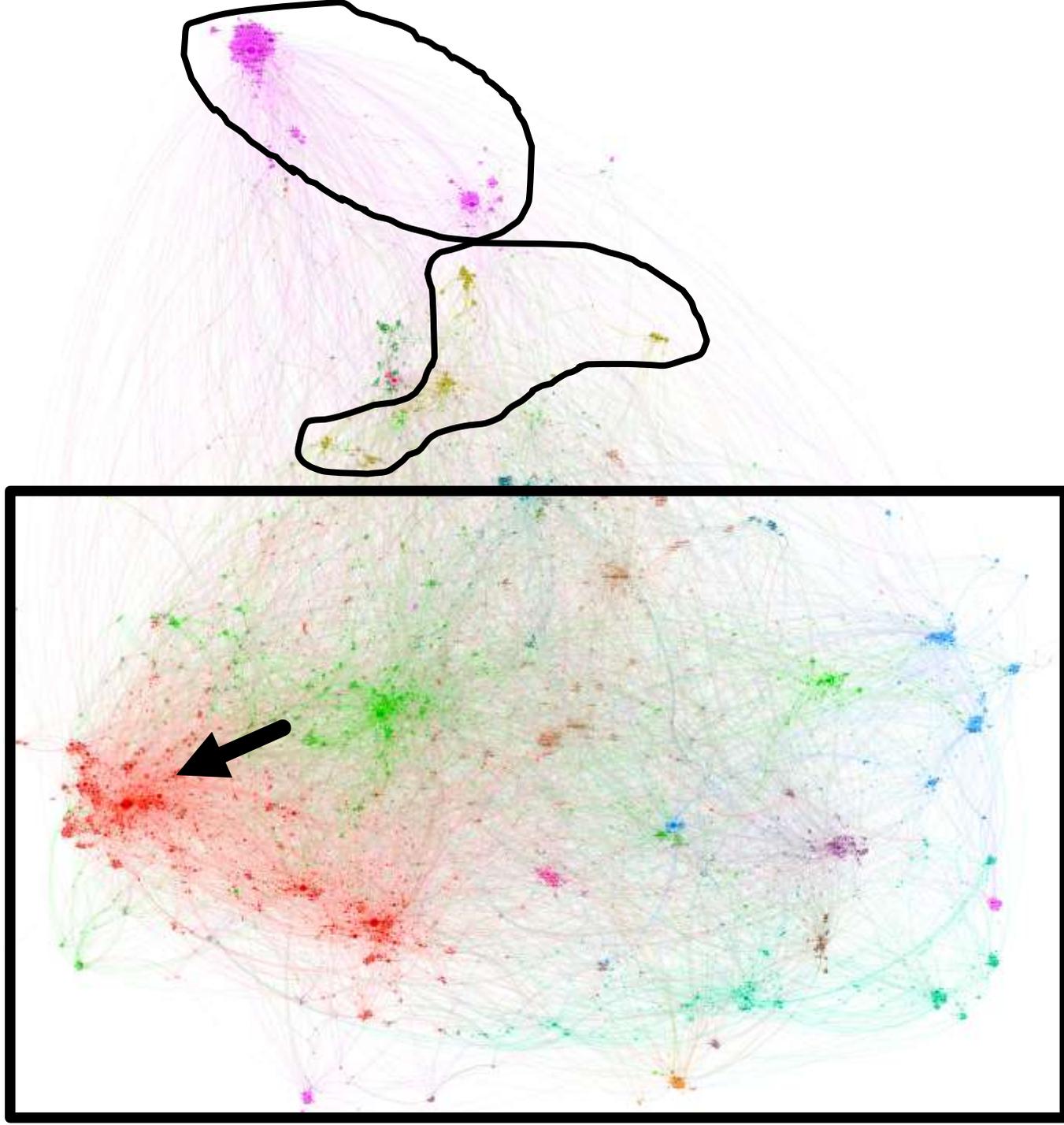
US-Präsidentschaftswahl 2016









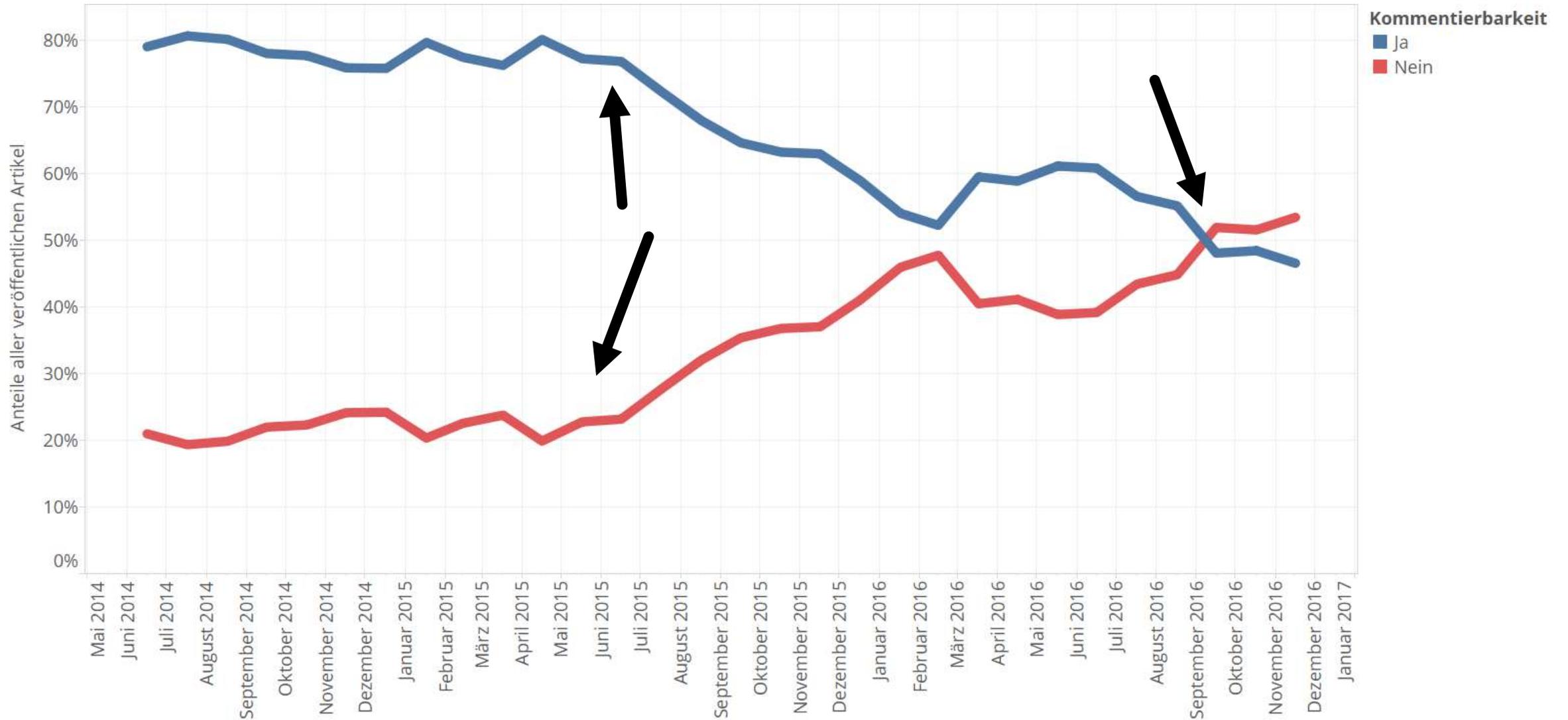


„Diskutieren Sie über diesen Artikel“



The screenshot shows a web browser window with the URL www.spiegel.de/sport/fussball/football-leaks-ron. The page header includes the Spiegel Online logo and navigation links. The main content area features a forum section titled "Forum >" with a sub-header "Diskutieren Sie über diesen Artikel" and a main title "Forum - Diskutieren Sie über diesen Artikel". A large black arrow points from the right towards this title. Below the title, it indicates "insgesamt 4 Beiträge" and a red button labeled "+ Alle Kommentare öffnen". The page is identified as "Seite 1 von 1". The first comment is by user "ArnoNyhm1984" from "heute, 19:53 Uhr", with the text "1. legal legitim ethisch" and a paragraph starting with "Ich habe weder in der Berichterstattung des jetzigen SPIEGELS noch der letztwöchigen Ausgabe zum selben Thema einen Hinweis gefunden, dass das, was Ronaldo, Özil & Co. da gemacht haben, ILLEGAL gewesen wäre. Es war [...]". A second comment by "bambata" from "heute, 21:17 Uhr" is partially visible at the bottom.

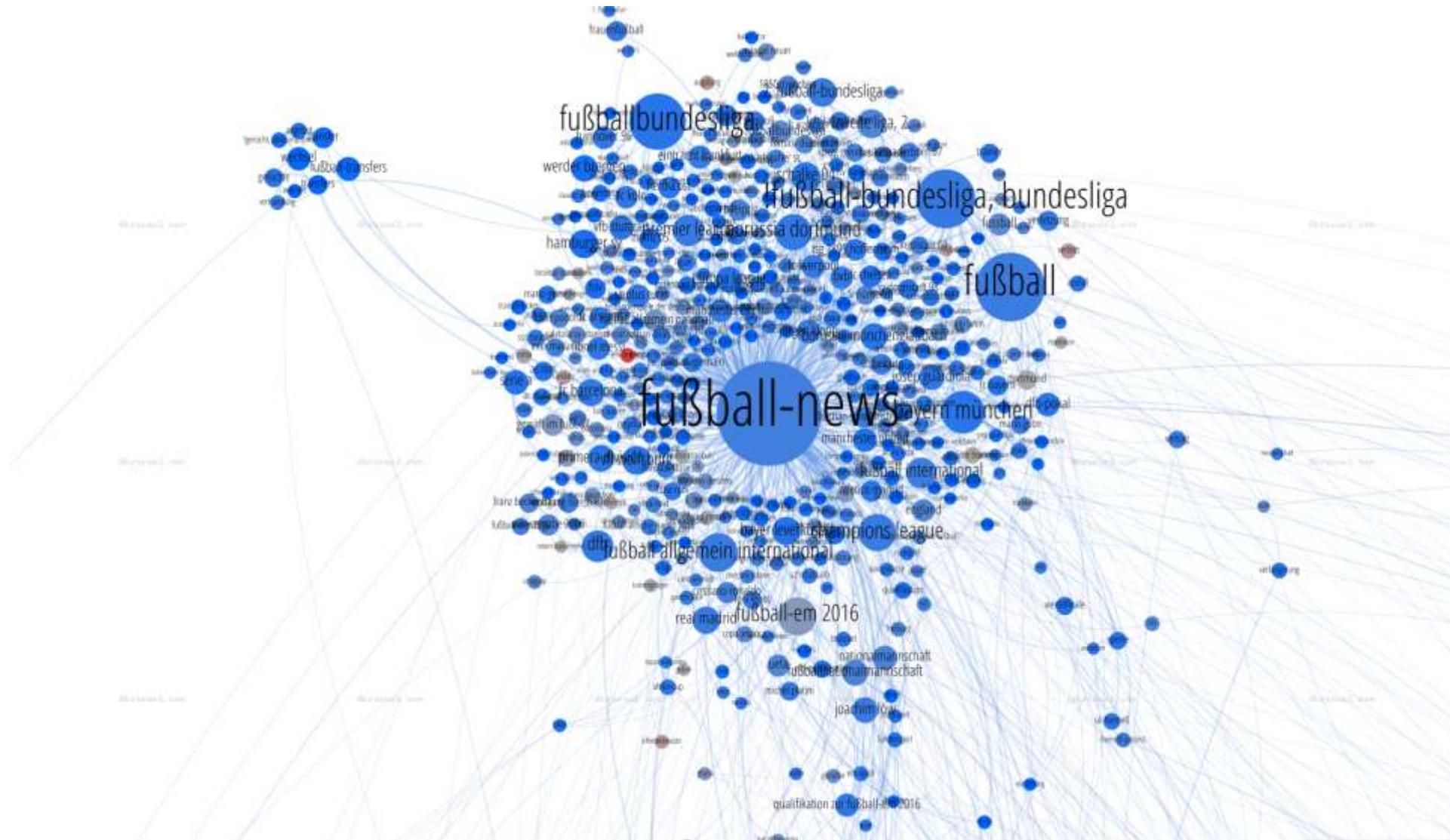
Zeitliche Entwicklung der Kommentierbarkeit



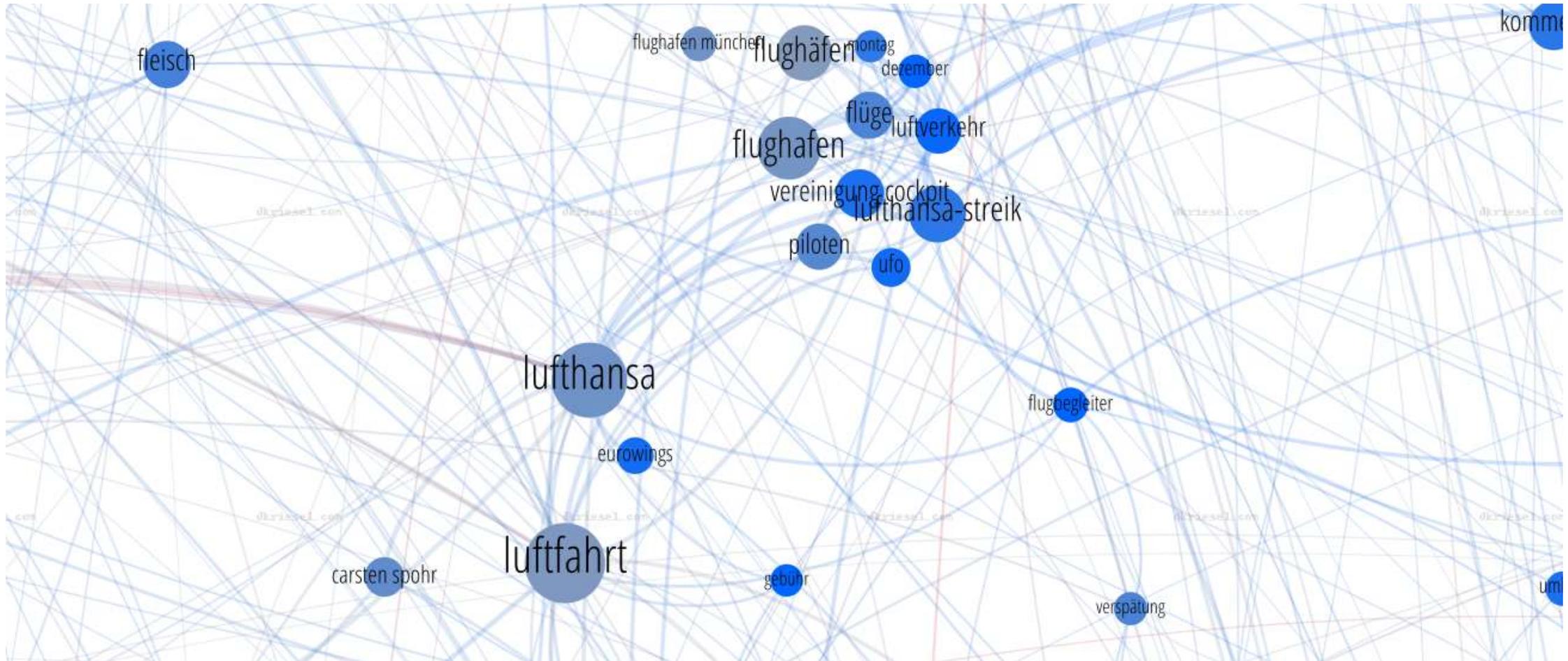
Kommentierbarkeit nach Themen



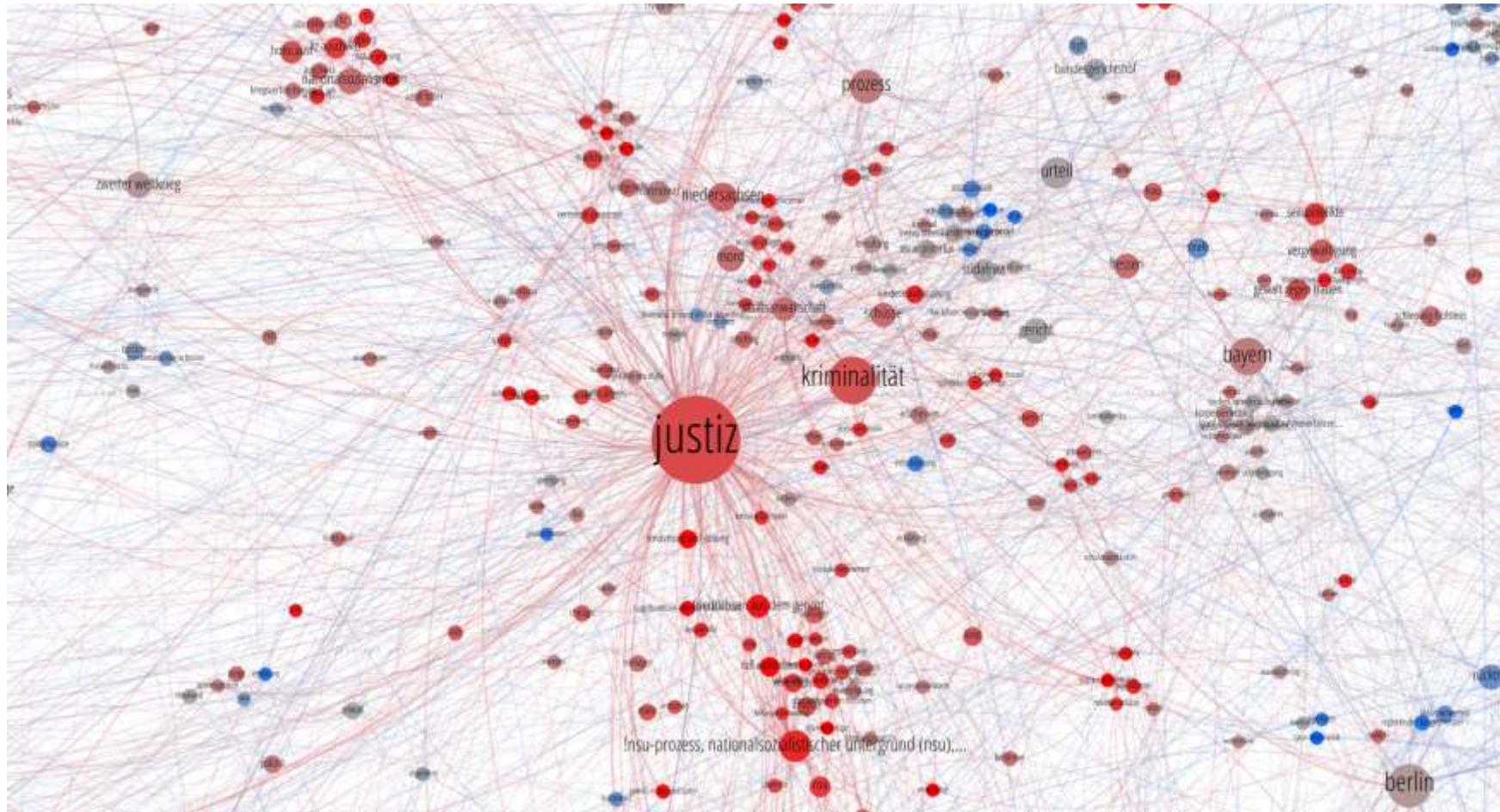
Kommentierbarkeit nach Themen



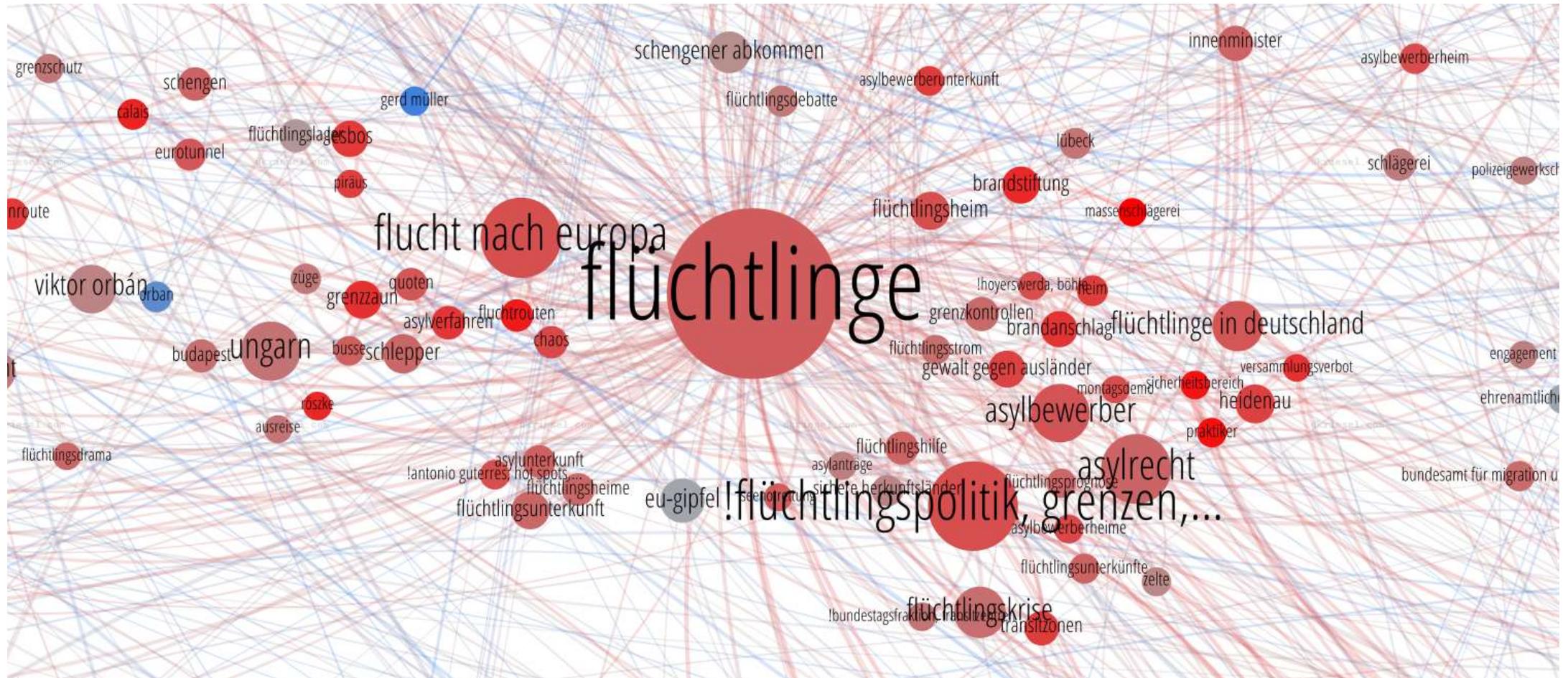
Kommentierbarkeit nach Themen



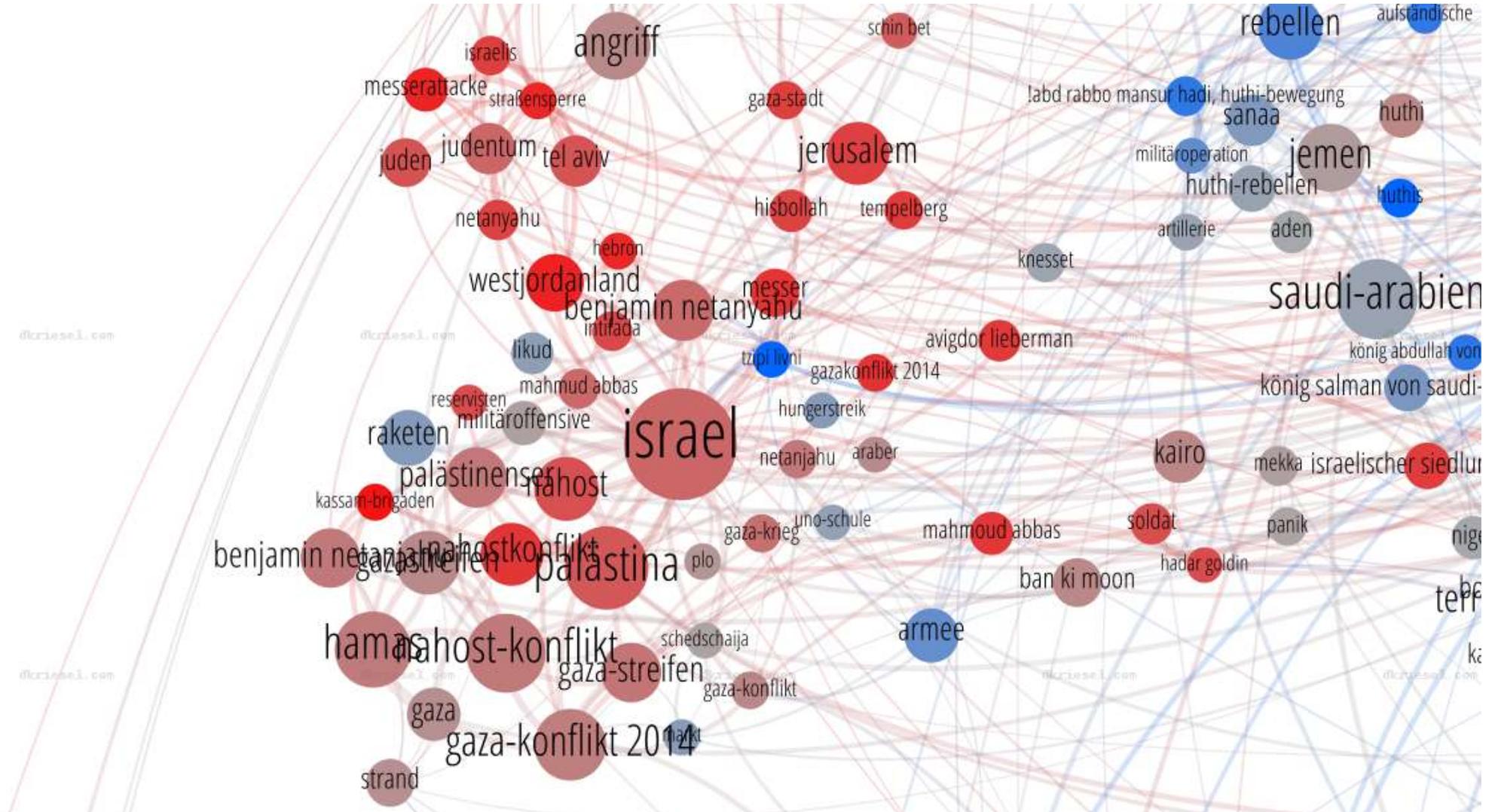
Kommentierbarkeit nach Themen



Kommentierbarkeit nach Themen



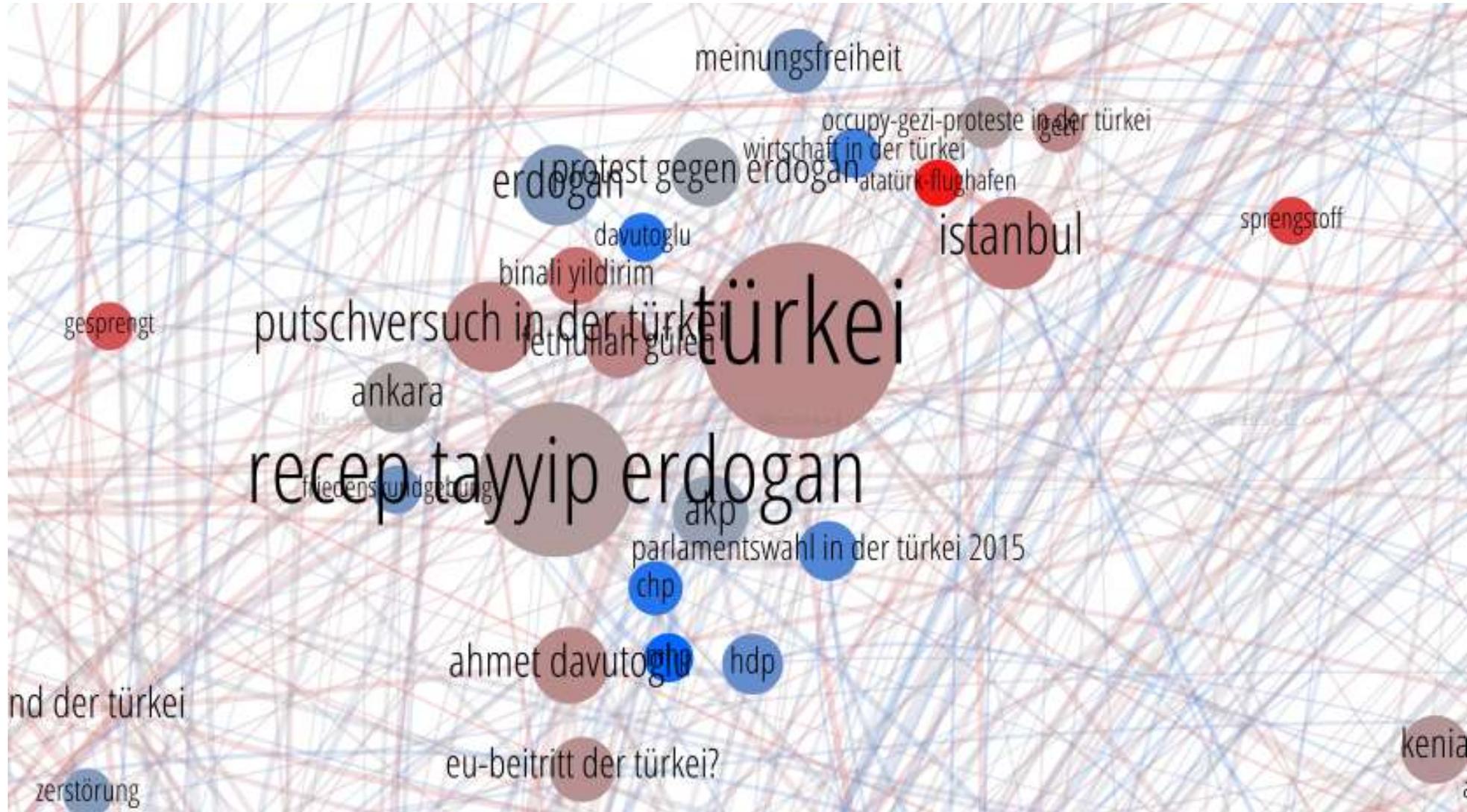
Kommentierbarkeit nach Themen



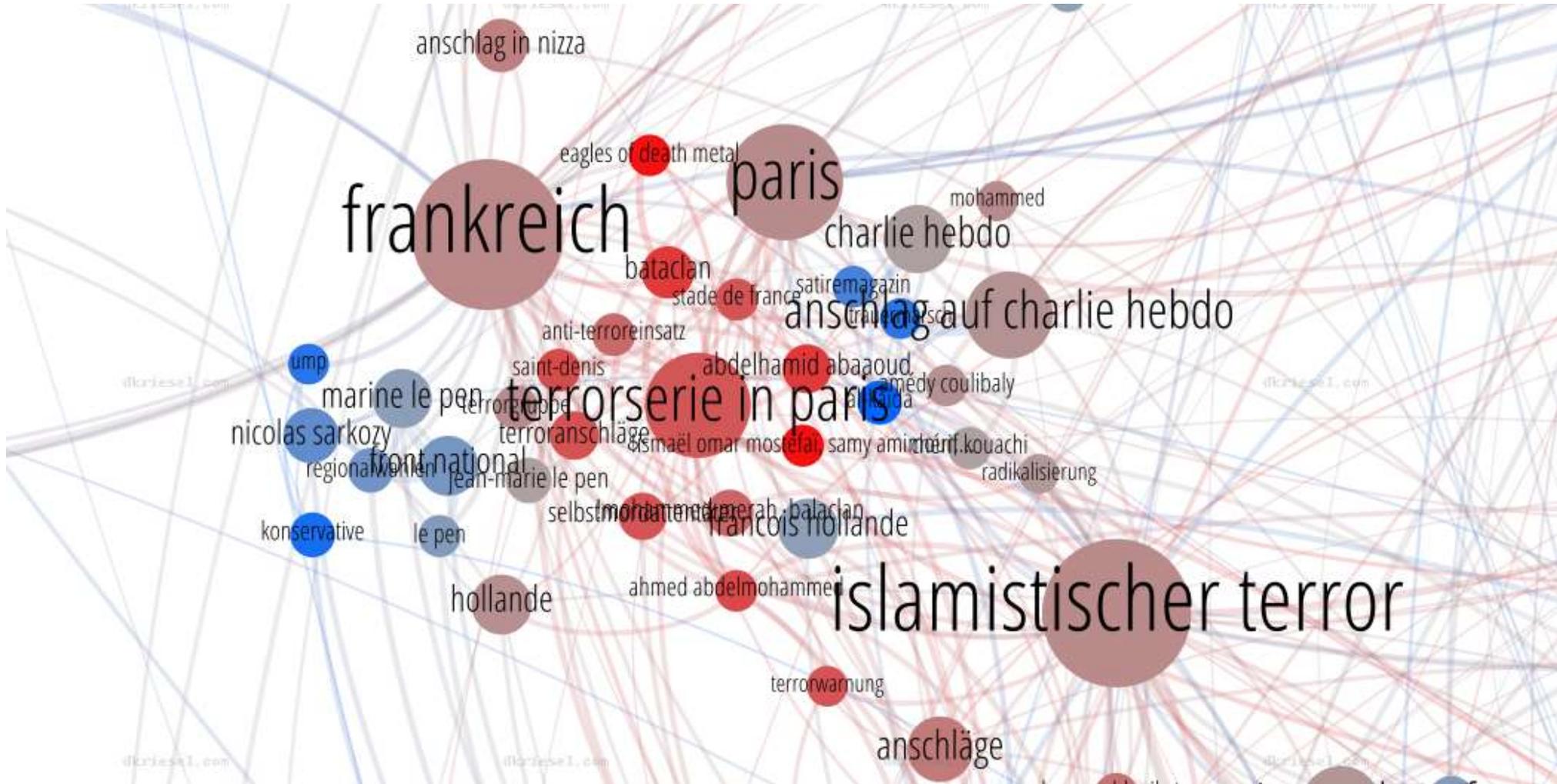
Kommentierbarkeit nach Themen



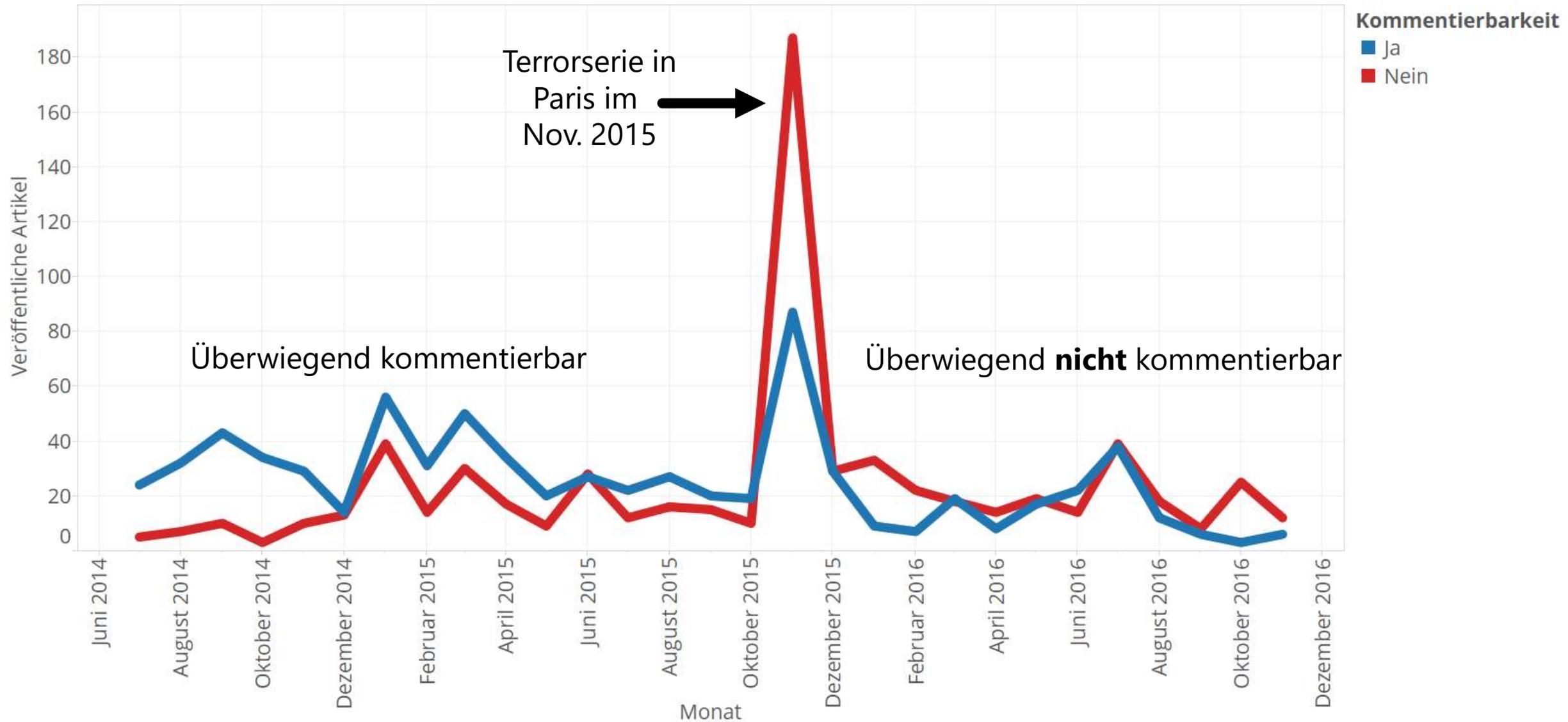
Kommentierbarkeit nach Themen



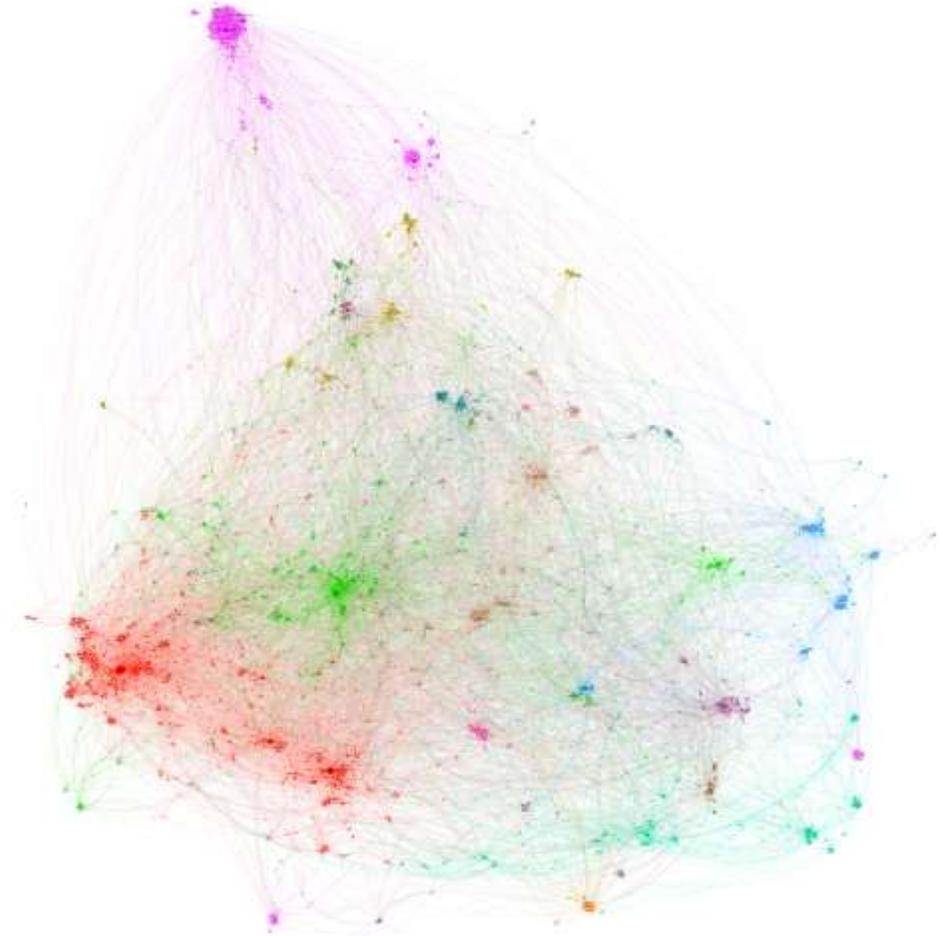
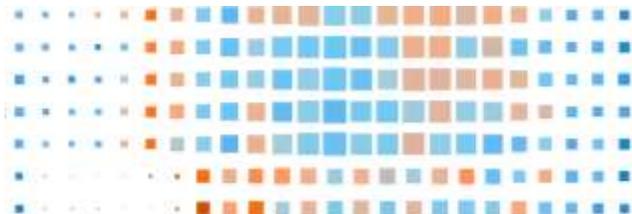
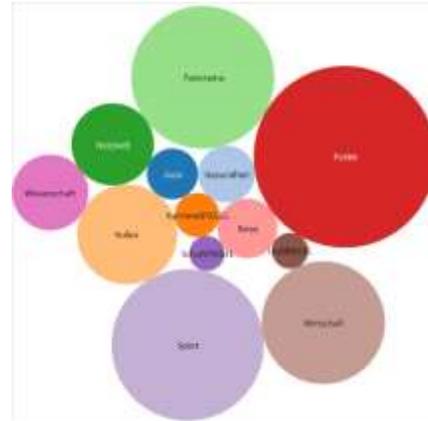
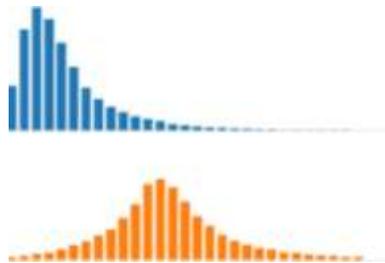
Kommentierbarkeit nach Themen



Kommentierbarkeit von Frankreich-Artikeln nach Zeit



Einfache und komplexe Datenunterteilungen



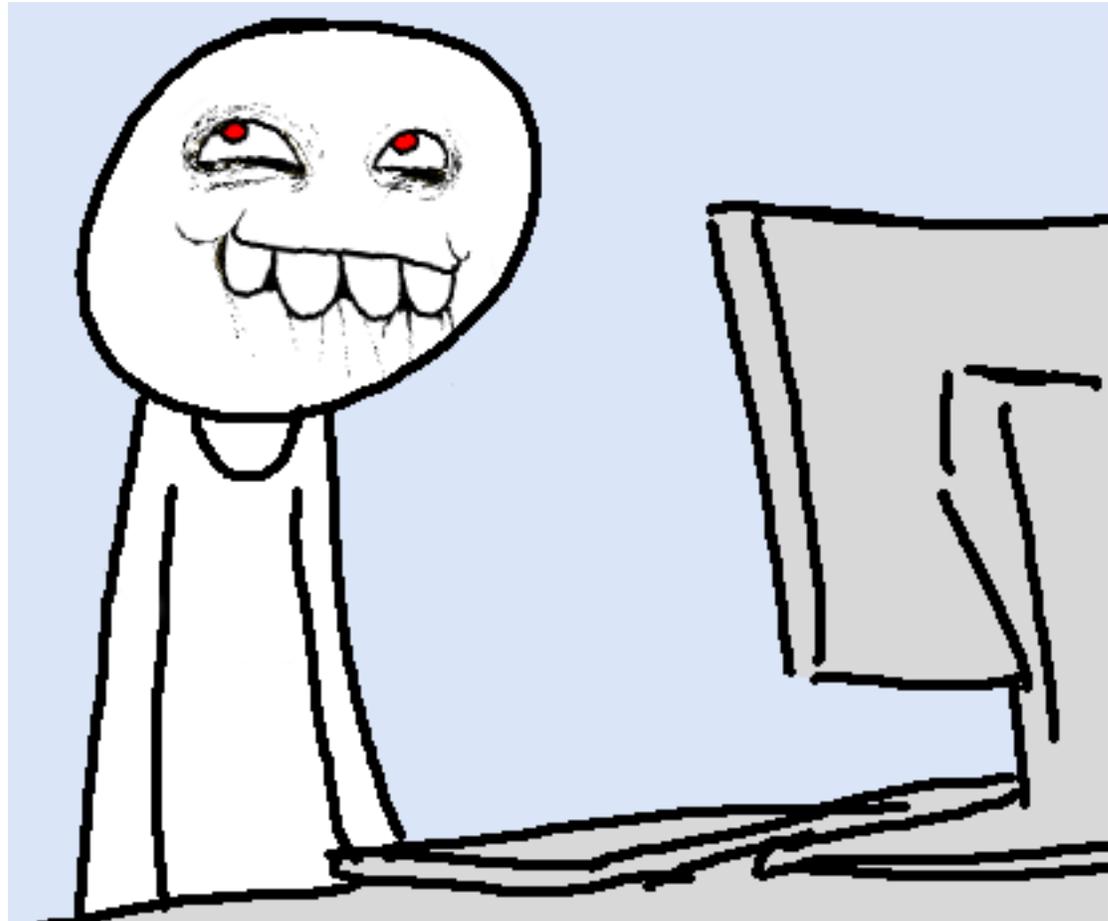
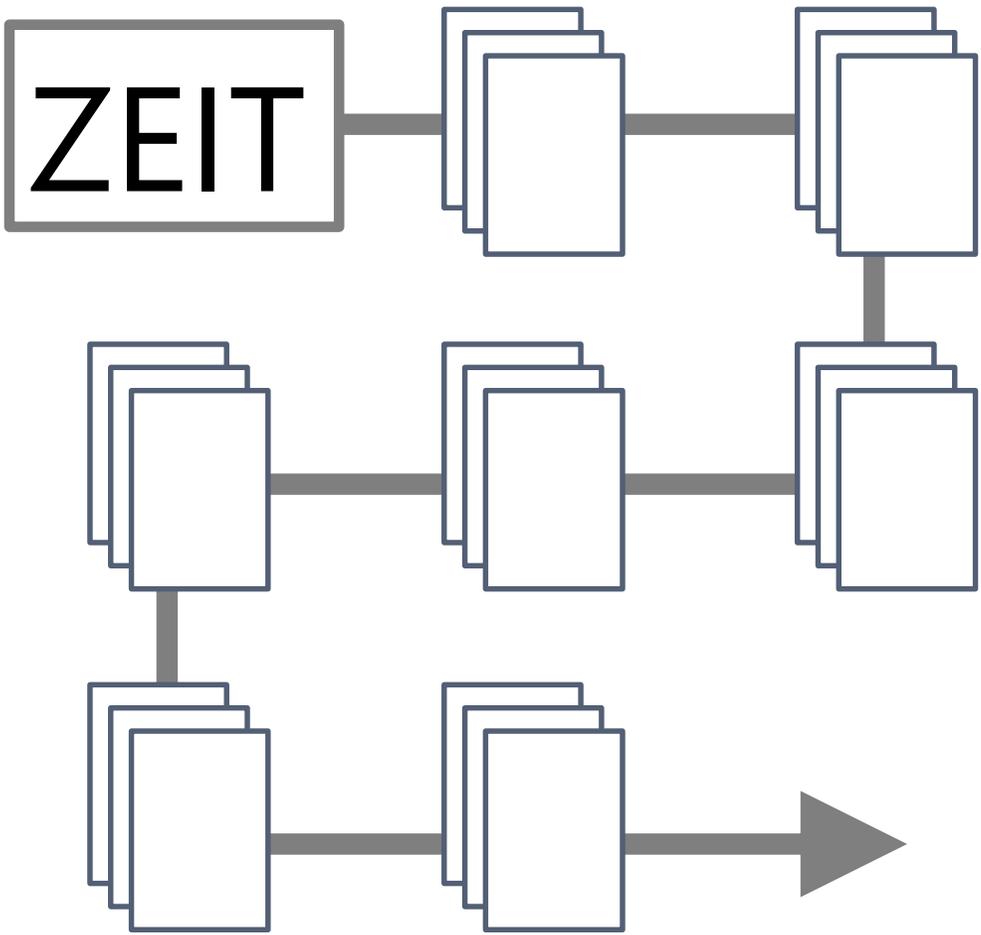
Data Science im Wahlkampf



Sagte ich: 100.000 Artikel?



Ich meinte 800.000.



Zeitliche Veränderungen in Artikeln sind meßbar



20. Januar 2015

SAP-Chef wächst 2014 langsamer als geplant

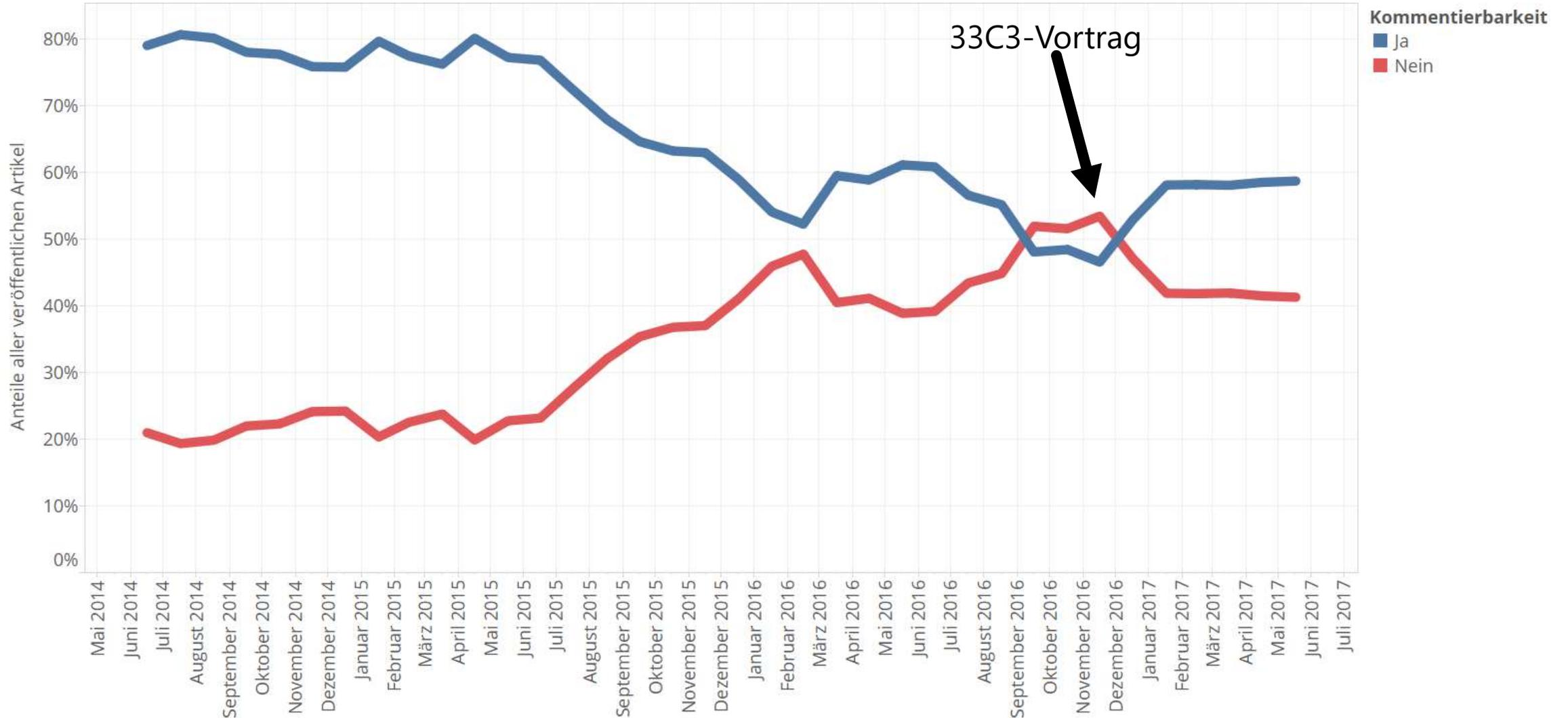
21. Januar 2015:

SAP wächst 2014 langsamer als geplant

Jetzt:

SAP kann Wachstums- und Gewinnziele nicht erfüllen

Zeitliche Entwicklung der Kommentierbarkeit, continued



HERZLICHEN DANK!

ALLE SPIEGELMINING-ARTIKEL AUF EINEN BLICK (BIS JETZT 5):

[HTTP://WWW.DKRIESEL.COM/SPIEGELMINING](http://www.dkriesele.com/spiegelmining)

KONTAKTAUFNAHME / FRAGEN / BUCHUNG:

MAIL@DKRIESEL.COM / [XING](#) / [LINKEDIN](#)