

- Wirtschaftsbranche Services und Dienstleistungen
- Wirtschaftsbranche Presse, Druck und Verlagswesen
- Wirtschaftsbranche Medizin, Pharmazie, Gesundheit
- Wirtschaftsbranche Medien, Internet, Information und Kommunikation
- Wirtschaftsbranche Kunst, Kultur, Musik
- Wirtschaftsbranche Informationstechnologie
- Wirtschaftsbranche Freizeit, Lifestyle, Unterhaltung
- Wirtschaftsbranche Fernsehen
- Wirtschaftsbranche Energieerzeugung und Versorgung
- Wirtschaftsbranche Elektronik und Elektrotechnik
- Weinbranche
- Pharmabranche
- Medienbranche
- Luftfahrtbranche
- Getränkebranche
- Fahrzeugbranche
- Buchbranche
- Kochbranche

TopicZoom Themen-Galerie zu deutschen Städten

Published by **levin** on 05/07/2020

Die folgenden Grafiken zeigen verschiedene Analysen von Artikeln der Wikipedia zu 14 deutschen Städten (Download der Artikel am 1.5.2020).

Nach einer automatischen Verschlagwortung der Städte-Artikel mit der TopicZoom Engine wurden für drei Beispiellisten von Themen (allgemeine Themen, Wirtschaftsbranchen und Sportarten) und zwei Gewichtungen (Weight und Significance) die jeweiligen Scores ermittelt und in einer Heat-Map Grafik dargestellt.

Die Schaubilder dieser Galerie illustrieren eine Weise, wie man mit Hilfe der TopicZoom Technologie automatisch einen Überblick über das Vorkommen vorgegebener Themen in einer Textkollektion erhalten kann. Anstelle der hier vorgestellten drei Beispiellisten von Themen könnte man beliebige andere Themenlisten als Grundlage nehmen.

Ergänzend zeigen wir unten, wie man mit Hilfe der TopicZoom Techniken für Textdokumente ein themenspezifisches "Screeningbild" ermitteln kann. Dabei werden für ein vorgegebenes Thema alle Passagen aller Eingabetexte mit einem Bezug zu diesem Thema optisch sichtbar gemacht.

Aktuell

- [TopicZoom Themengalerie: Analyse von Wikipediaseiten zu Pharmafirmen](#)
- [TopicZoom-Themengalerie: Suchmaschinentechnologie zur thematischen Suche in Studiengangsbeschreibungen](#)
- [TopicZoom Themen-Galerie zu deutschen Städten](#)

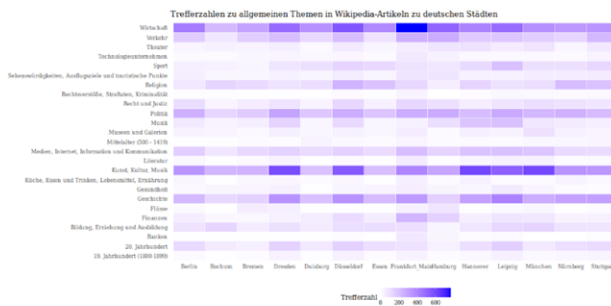
Search

Search 

1. Allgemeine Themen

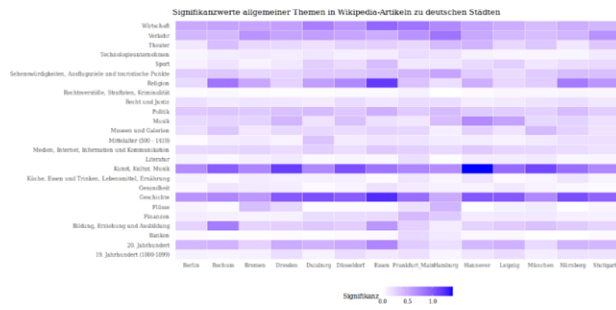
1a. Gewichte allgemeiner Themen in Wikipediatexten zu Städten

In der ersten Anwendung schauen wir auf allgemeine Themen wie "Wirtschaft", "Verkehr", Musik" oder "Gesundheit". Ergänzend werden einige weitere Themenbereiche wie Zeitperioden ("Geschichte", "20. Jahrhundert") oder geographische Themen ("Flüsse") mit aufgeführt. Das Gewichtungsprinzip der ersten Grafik bezieht sich auf Zahl der Treffer von Textbausteinen, die nach der TopicZoom-Ontologie den genannten Themenbereichen zuzurechnen sind. Hierbei erfahren sehr allgemeine Themen eine hohe Gewichtung, da diese viele Teilthemen mit spezifischen Begriffen abdecken und entsprechend häufig auftreten. So wurden beispielsweise im Artikel zu "Frankfurt am Main" über 600 Textpassagen mit einem Bezug zum Thema "Wirtschaft" gefunden.



1b. Signifikanz allgemeiner Themen in Wikipediatexten zu Städten

Die zweite Grafik stellt dieselbe Themenmenge unter einem anderen Gewichtungsprinzip dar. Die Angaben beziehen sich nun auf die "Signifikanz" der jeweiligen Themenbereiche. Bei gleicher Trefferzahl erhalten häufig auftretende Themen eine kleinere Signifikanz, seltener auftretende Themen eine höhere. Damit werden häufig auftretende, "banale" Themen bestraft, seltene Themen "belohnt". Im vorliegenden Fall erhalten hierdurch beispielsweise geschichtliche Themen in erkennbarer Weise ein erhöhtes Augenmerk.



1c. Themenspezifische "Screeningbilder" der Textdokumente: Nachweise für die Treffer zu den allgemeinen Themen

Der Blick auf die Grafiken löst oft ein Interesse aus, ein näheres Bild zu den dort vermittelten Informationen zu erhalten. Schaut man auf die vorausgegangene Grafik, so mag man sich fragen, warum "Religion" und "Geschichte" gerade im Artikel zu Essen, oder warum "Kunst, Kultur, Musik" im Artikel zu Hannover eine hohe Signifikanz besitzen.

Alle TopicZoom Ergebnisse sind vollständig rekonstruierbar, da sie aus einzelnen rekonstruierbaren Textbausteinen resultieren.

Nutzer können stets im Detail nachsehen, welche Textpassagen aller Texte Inhalte zu den genannten Themenbereichen enthalten. Diese Inspektionstechnik wird hier anhand einiger Beispielthemen illustriert.

Klickt man nachfolgenden Link, so werden die Treffer zu Religion dargestellt. Mit einem Mouseover über die gelben Bereiche im Screeningbild von Essen sieht man u.a., dass Essen eine Rolle in der Reformation spielte. Andere Textteile gehen auf die Novemberprogramme 1938 ein.

Inspektion der Treffer zu Religion

Ein näherer Blick auf das Thema "Kunst, Kultur, Musik" führt uns in Hannover u.a. auf die Beschreibungen des Historischen Museums, des Schauspielhauses Hannover sowie auf viele weitere Einrichtungen.

Inspektion der Treffer zu Kunst, Kultur, Musik

Einen kleinen "Reiseführer deutscher Städte" erhalten wir durch den Blick auf die Treffer zu Sehenswürdigkeiten.

Inspektion der Treffer zu Sehenswürdigkeiten

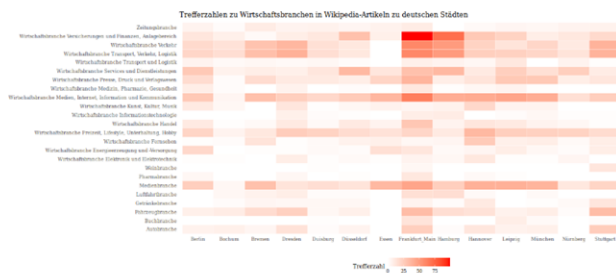
2. Wirtschaftsbranchen in deutschen Städten

Unsere zweite Themenauswahl ist deutlich spezieller. Die Ergebnisse zeigen die Abdeckungsbreite der TopicZoom-Ontologie, die den Hintergrund für alle ermittelten Ergebnisse darstellt. Wir schauen hier auf diverse Wirtschaftsbranchen. Damit erhalten wir ökonomische Porträts der Städte.

2a. Gewichte von Wirtschaftsbranchen in Wikipediatexten zu Städten

Die erste Grafik zu Wirtschaftsbranchen bezieht sich wieder auf reine Trefferzahlen. Man beachte die veränderte Grundskala:

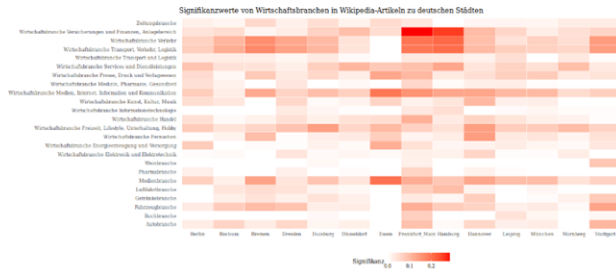
Für die Wirtschaftsbranchen ergeben sich deutlich kleinere Trefferzahlen als für die oben betrachteten allgemeinen Themen.



2b. Signifikanz von Wirtschaftsbranchen in Wikipediatexten

zu Städten

Auch hier ergibt sich eine leicht veränderte Gewichtung, wenn wir die Signifikanz der Themen anstelle der Trefferzahlen zur Grundlage der Bewertung machen. So wird z.B. die Medienbranche aufgewertet.



2c. Themenspezifische "Screeningbilder" der Textdokumente: Nachweise für die Treffer zu Wirtschaftsbranchen

Ein näherer Blick auf die Treffer zum Thema "Medienbranche" führt uns u.a. in Hannover auf das Funkhaus des NDR, in Essen auf die Funke Mediengruppe (ehemals WAZ-Mediengruppe) oder in Bremen auf die Korrespondentenbüros von RTL und Sat1:

[Inspektion der Treffer zur Medienbranche](#)

Schauen wir auf die Treffer zu "Pharmabranche", so erfahren wir z.B. dass keiner der großen Pharmakonzerne mehr seinen Sitz in Frankfurt hat, aber Merz Pharma in Frankfurt ist. Der Blick auf Hannover legt offen, dass dort Boehringer Ingelheim 2012 das Europäische Forschungszentrum für Tierimpfstoffe, Boehringer Ingelheim Veterinary Research Center, eröffnet hat.

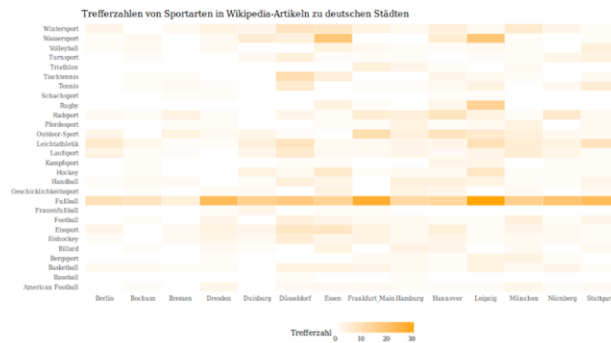
[Inspektion der Treffer zur Pharmabranche](#)

3. Sportarten in deutschen Städten

Unsere letzte Themenliste ist dem Sport gewidmet.

3a. Gewichte einzelner Sportarten in Wikipediatexten zu Städten

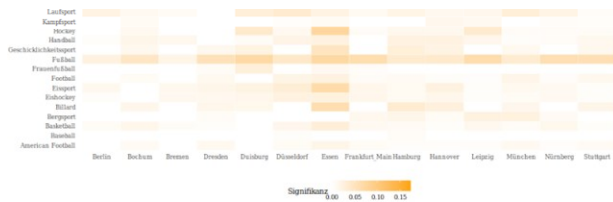
Zunächst zeigen wieder die reinen Trefferzahlen zu einzelnen Sportarten in den Stadtbeschreibungen. Wie nicht anders zu erwarten, spielt der Bereich "Fußball" hier eine dominierende Rolle.



3b. Signifikanz einzelner Sportarten in Wikipediatexten zu Städten

Da über Fußball häufig geschrieben wird, erscheint dieser Themenbereich bei Blick auf die Signifikanz weniger dominant.





3c. Themenspezifische "Screeningbilder" der Textdokumente: Nachweise für die Treffer zu Sportarten

Der unvermeidbar Blick auf "Fußball" führt uns u.a. nicht ganz unerwartet in München auf den TSV 1860 München, den FC Bayern München und die Allianz-Arena, in Leipzig auf den RB Leipzig.

Inspektion der Treffer zum Thema "Fußball"

Beim Blick auf die vorausgegangene Grafik fällt die Rolle des Wassersports in Essen auf. Die Ursache hierfür kann der Leser nun leicht selbst ermitteln:

Inspektion der Treffer zu Wassersport

Unsere letzte Darstellung zeigt auf einen Blick die Treffer zu drei Themen Volleyball, Turnsport und Triathlon.

Inspektion der Treffer zu Volleyball, Turnsport und Triathlon

Quellennachweis:

Die in diesem Post als Datengrundlage verwendeten Texte stammen aus der Wikipedia und unterliegen den Lizenzbestimmungen Creative Commons Attribution-ShareAlike 3.0 Unported

Categories: **NEUIGKEITEN**

Tags: Heatmap Thematischer Vergleich

Thematisches Bild Themenerkennung

Themenvergleich



Related Posts



NEUIGKEITEN

TopicZoom Themengalerie: Analyse von Wikipediaseiten zu Pharmafirmen

Automatische Sortierung von medizinischen Texten nach Fachgebieten, Krankheiten, Medikamenten und anderen Gesundheitsthemen



NEUIGKEITEN

TopicZoom-Themengalerie: Suchmaschinentechnologie zur thematischen Suche in Studiengangsbeschreibungen

Die thematische Suche, bei welcher übereinstimmende Themen zwischen Anfragen und Antworten die Grundlage für die Ermittlung von Antwortlisten darstellt, ist eine interessante Alternative zu üblichen keywordbasierten Suchverfahren. Hervorzuheben ist, dass bei dieser Art der Suche [Read more...](#)



NEUIGKEITEN

TopicZooms Text KI-Technologie

Künstliche Intelligenz, KI, Artificial Intelligence oder AI ist eines der Schlagwörter, die im Zusammenhang mit der Digitalisierung unterschiedlicher Bereiche immer wieder genannt wird. Leider bleibt oft unklar, worin diese "Intelligenz" denn eigentlich besteht. In dieser [Read more...](#)