

Tracking Discourses on Antibiotic Resistance with Corpus Linguistic Methods

Dr. Julia Krasselt, Digital Discourse Lab ZHAW

In collaboration with: Prof. Dr. Peter Stücheli-Herlach (Project Lead),
Birgitta Borghoff, Elise Beauverd, Loïse Bilat, Natalie Schwarz

Applied NLP Technologies for Physical & Mental Health in Practice, Lausanne, 28.03.2022

Outline

Project Background

Corpus Centered Discourse Analysis

Data: Swiss-AL

Methods & Results

Discussion

Project Background

Project Team at ZHAW:

Prof. Dr. Peter Stücheli-Herlach

Dr. Julia Krasselt

Birgitta Borghoff

Elise Beauverd

Loïse Bilat

Natalie Schwarz

Matthias Fluor

Klaus Rothenhäusler

Project Background

- Swiss Strategy on Antibiotic Resistance (StAR), implemented by four federal offices (FOPH, FSVO, FOAG, FOEN)
- launched in 2015
- Main Objective: «Maintaining the effectiveness of antibiotics for humans and animals, and combating resistance»



Project Background

- Two consecutive research projects, funded by FOPH
- 2017–2019: Situated corpus based discourse analysis on the Strategy on Antibiotic Resistance (StAR)
- 2019–2021: Corpus based discourse tracking on StAR

What are the linguistic characteristics (patterns of language use) of public communication in the Swiss discourse on antibiotic resistance by actors relevant to the policy field?

Project Background

- Applied research project: derivation of empirically based recommendations for communication
- Language regions: Romandie & German speaking Switzerland
- Framework: applied discourse analysis
- Method: corpus linguistics

Corpus centered discourse analysis

Corpus centered discourse analysis

«How is it that one particular statement and no other
appeared in its place?»

(Foucault [1968], 2001, p. 899)

Corpus centered discourse analysis

- Discourses are characterized by **patterns of language use**
- Language is embedded in **social, cultural and historical contexts**
- Language not only represents **knowledge**, but also **creates** it
- Discourse: «thematically constrained ('virtual') text corpus from which the researcher has to extract an (necessary smaller) 'actual' corpus of analysis.» (Spitzmüller & Warnke, 2001, p. 76)
- **Discourse as a linguistic object**

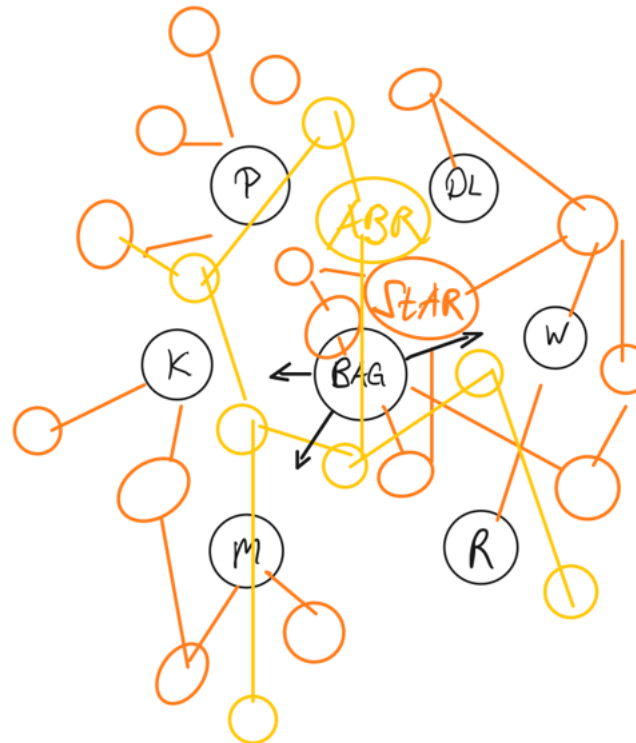
Corpus centered discourse analysis

Risiko für Antibiotikaresistenz senken: St.Galler Ärzte erzielen Forschungserfolg

Ärzte des Kantonsspitals St.Gallen haben die ideale Dauer der Antibiotikatherapie bei einer häufigen Form der Blutvergiftung untersucht und einen Erfolg erzielt.

Abstimmungen
Bundesrat lehnt Pestizid- und

DISKURSNETZWERK



triebe
den und
sch nur,

Resistenzen können sich auch ohne Antibiotika-Einsatz verbreiten

05.09.2019 | News
Von: Fabio Bergamin

Antibiotikaresistenzen verbreiten sich nicht nur dort, wo viel Antibiotika eingesetzt werden, schliessen ETH-Forschende aus Laborexperimenten. Das bedeutet: Um Resistenzen einzudämmen reicht es nicht, den Antibiotikaeinsatz zu reduzieren. Man sollte auch die Verbreitung resistenter Keime blockieren.

ETHZ

BAZ

AGRICULTURE

Actualisé 12 novembre 2018, 16:31

Antibiotiques: le projet de Fribourg devient fédéral

Le projet de diminution de l'usage des antibiotiques, déjà mis en oeuvre dans le canton de Fribourg, est devenu un projet-phare au niveau fédéral.

20 Minutes

Lehren aus der Corona-Krise

Lebenswichtige Medikamente sollten immer verfügbar sein

Immer billiger, irgendwo produziert, keine Lager: Nicht erst Corona zeigt die Risiken von Arbeitsteilung und Preiskampf.

In den vergangenen Monaten waren lebenswichtige Medikamente wie Antibiotika, Betäubungsmittel oder Schmerzmittel immer wieder mal sehr knapp. Ärztinnen und Apotheker behelfen sich teils mit Medikamenten in anderen Grössen oder Konzentrationen.

EN

IKA NICHT

leudet, dass
gar nicht mehr
ien. Der falsche
von Antibiotika
nzen führen.
nnen auf andere
erden.

en wenden Sie
1 Arzt, Ihre

Zahnärztin / Ihren Zahnarzt oder Ihre
Apothekerin / Ihren Apotheker.

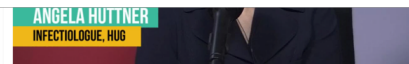
Künstliche Intelligenz sagt Antibiotikaresistenzen schneller vorher

Forschende des Universitätsspitals Basel (USB) und der ETH Zürich haben gezeigt, dass sich Resistenzen von Bakterien mittels neuartigen Computeralgorithmen deutlich schneller ermitteln lassen als bisher. Dies könnte helfen, schwere Infekte in Zukunft effizienter zu behandeln – und wäre ein grosser Fortschritt im Kampf gegen antibiotikaresistente Bakterien.

Universitätsspital Basel

SRF

Zurich Universities of Applied Sciences and Arts



RTS

FOPH

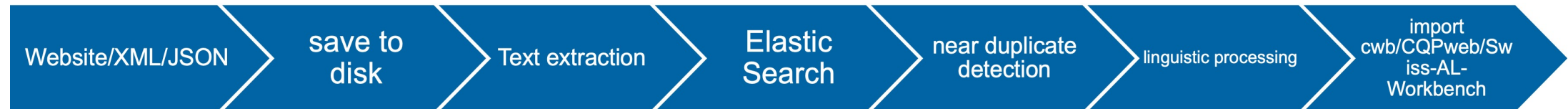
Data: Swiss-AL

Data: Swiss-AL

- Swiss-AL: A multilingual corpus for Applied Linguistics
- The largest text corpus in Switzerland
- Covers all 4 Swiss language regions
- Designed for discourse analysis, but applicable also in other research disciplines
- Sources
 - texts published on websites of publicly relevant actors (politics, authorities, industry, science, civil society)
 - Journalistic media, access via Swiss Media Database

Data: Swiss-AL

Corpus Linguistic Pipeline



- for websites: source-specific scraping of title, lead and content with Xpath)

- POS-Tagging
- Lemmatization
- Named Entity Recognition
- Dependency Parsing
- language specific annotations

Krasselt, J., Dreesen, P., Fluor, M., Mahlow, C., Rothenhäusler, K., & Runte, M. (2020). Swiss-AL: A Multilingual Swiss Web Corpus for Applied Linguistics. *Proceedings of the 12th Language Resources and Evaluation Conference*, 4138--4144. <https://www.aclweb.org/anthology/2020.lrec-1.510/>

Data: Swiss-AL

subcorpus	content	size (in words)
Swiss-AL-Base	media, websites from politics, industry, science, civil society	DE: 2.15 B. FR: 1 B. IT: 14 M.
Swiss-AL-Media	media (national, regional, local)	DE: 1.2 B. FR: 381 M. IT: 150 M. RM: 5.9 M.
Swiss-AL-Specialized	Discourse specific corpora (e.g., covid-19, right-wing populism, renewable energy, antibiotic resistance)	-

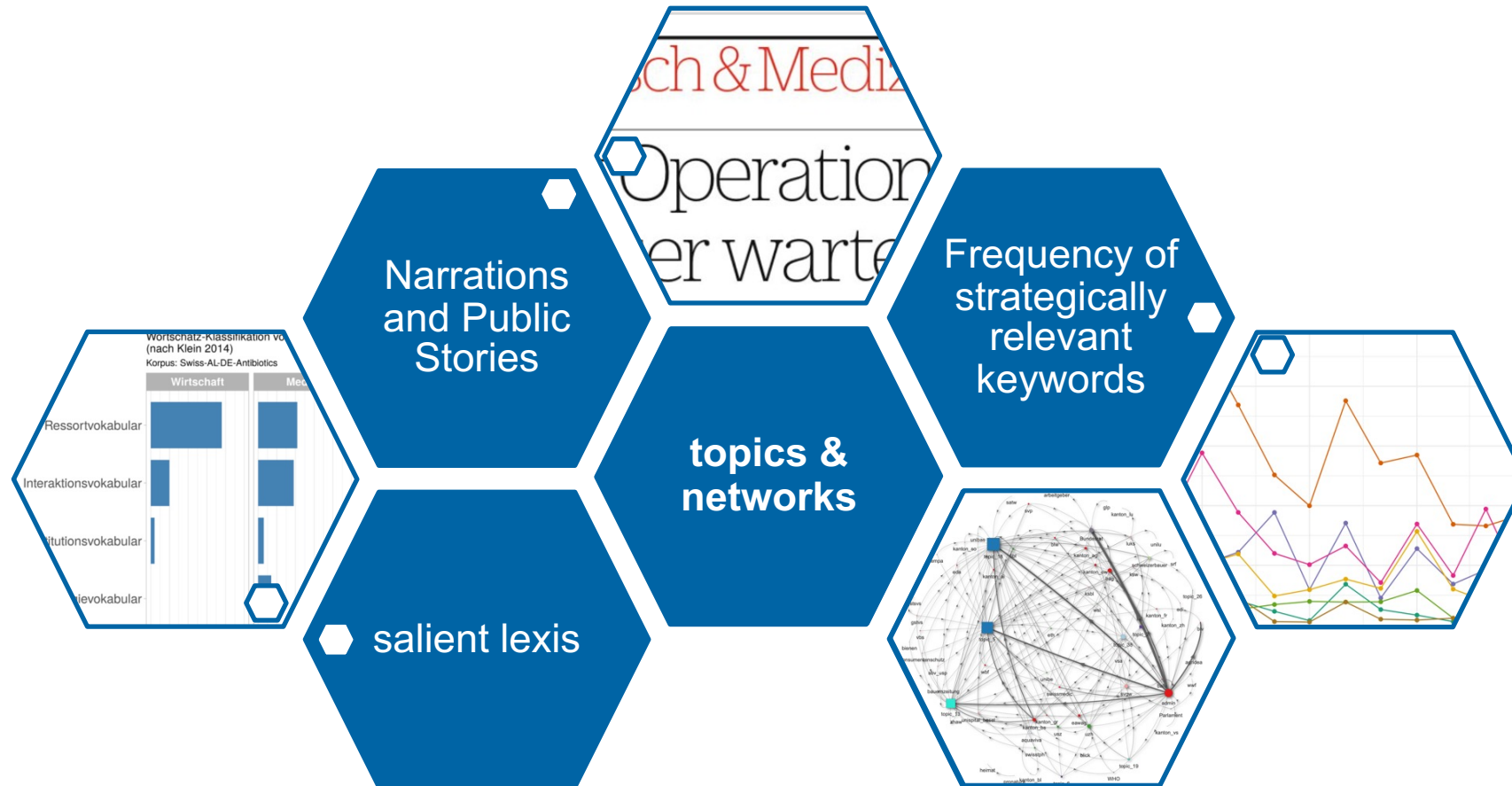
Data: Swiss-AL

Swiss-AL-Antibiotics

- Corpus containing
 1. Journalistic media
 2. texts published on websites from **150 relevant stakeholders**
 - Politics (e.g., federal offices & cantons)
 - Industry (e.g., hospitals, professional associations, pharmaceutical associations)
 - Science (e.g, universities)
- **German corpus:** 40 mio. words / 63k texts (2018-2021)
- **French corpus:** 3.15 mio. words / 4.7k texts (2018-2021)
- richly linguistically annotated

Methods & Results

Methods & Results



Methods & Results

Research Questions

1. Which topics characterize the discourse on antibiotic resistance?
2. Which topics dominate the discourse on antibiotic resistance?
3. Who is talking about which topics to what extent?
4. How do the answers to questions 1, 2 and 3 change over time?

Methods & Results

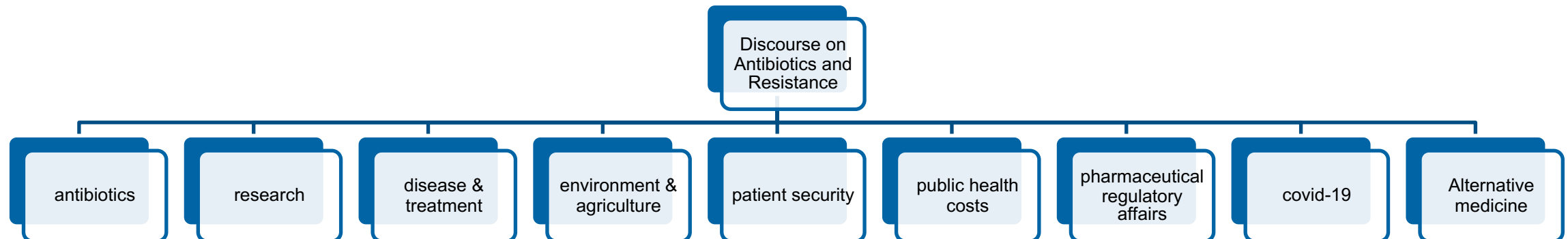
1. Which topics characterize the discourse on antibiotic resistance?

- Topic Modeling: algorithmic identification of thematic structure in large collection of texts
- Topic: probability distribution over the vocabulary in a corpus (bag-of-words approach)
- Latent dirichlet allocation (LDA)
- database: large, thematically unspecific corpus
- Qualitative labeling of topics and identification of topics relevant to discourse on antibiotic resistance

Methods & Results

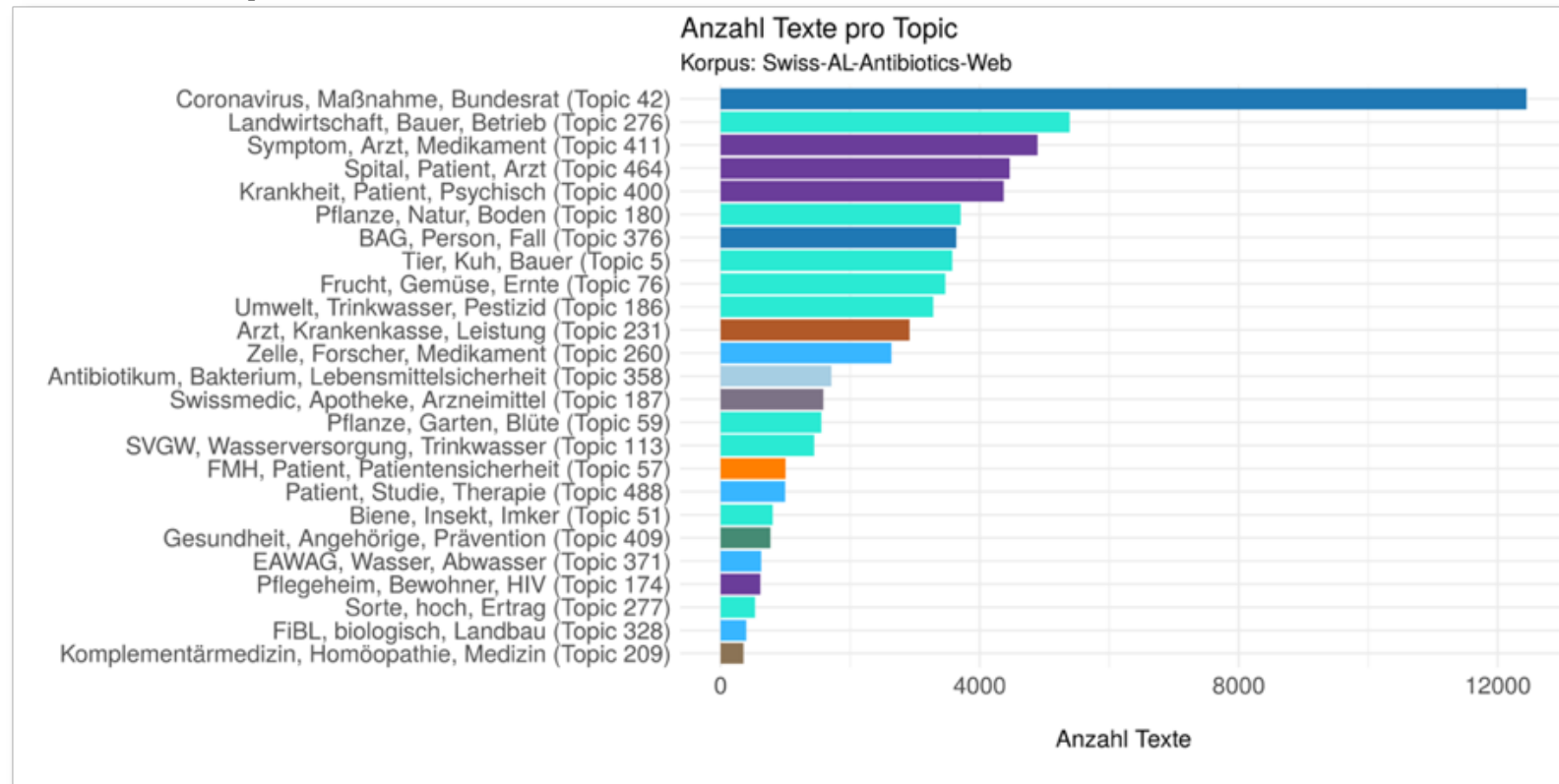
Which topics characterize the discourse on antibiotic resistance?

- Results of the manual clustering of topics



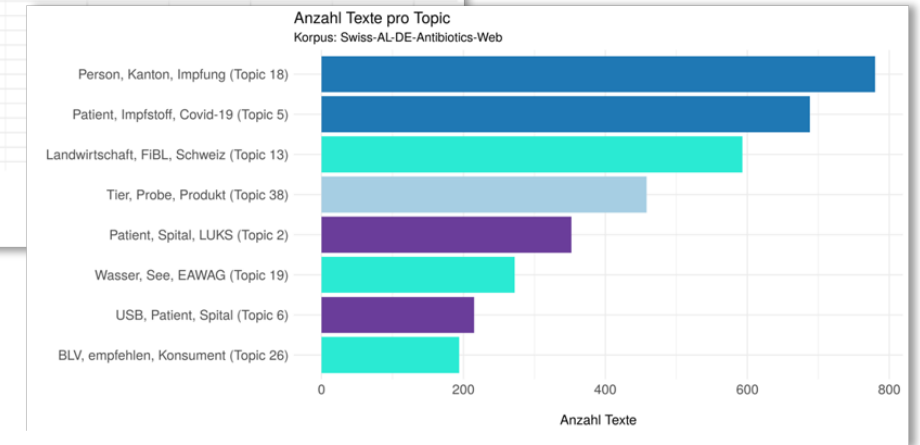
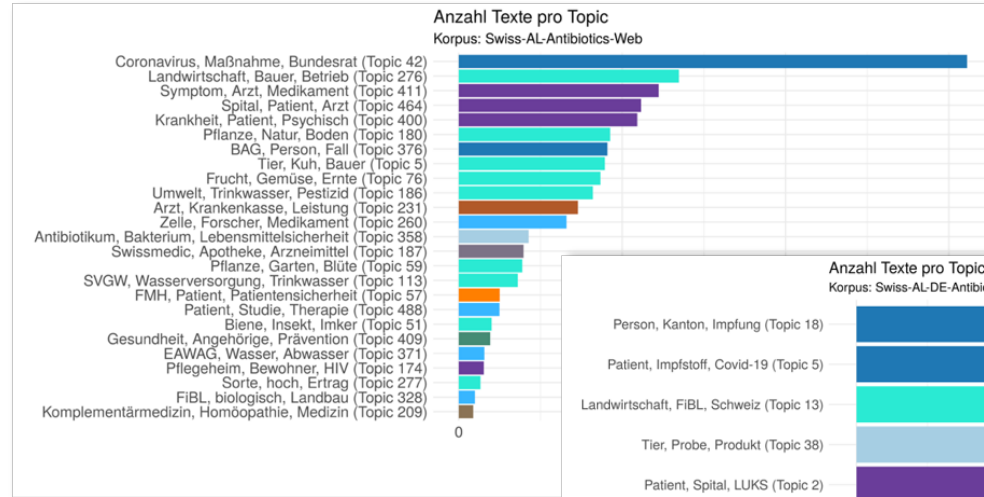
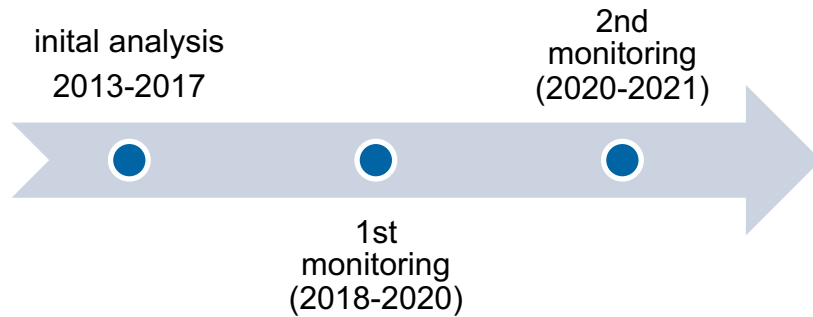
Methods & Results

Which topics dominate the discourse on antibiotic resistance?



Methods & Results

Tracking Perspective



environment & agriculture is a constant thematic hub

formative events: (1) Popular initiatives on clean drinking water and pesticide ban (2) Covid-19

disease & treatment loses the status as a thematic hub

antibiotic core themes play a peripherale role

Discussion

Discussion

So what?

- Why is a discourse analysis on antibiotic resistance valuable for the FOPH?
 - Communication about antibiotic resistance takes place in networks of stakeholders and patterned language use
 - To communicate successfully, a change of perspective is necessary

WHERE am I in this network?

WHO else is in this network?

WHAT is talked about in this networks?

HOW do I have to position myself in this network?

...

- Methods developed in NLP are an opportunity of applied discourse analysis



Thank you!

Who we are:

Digital Discourse Lab

[Web: https://www.zhaw.ch/de/linguistik/dienstleistung/digital-discourse-lab/](https://www.zhaw.ch/de/linguistik/dienstleistung/digital-discourse-lab/)

Mail: digitaldiscourselab.linguistik@zhaw.ch

Twitter: @lab_discourse



Prof. Dr. Peter Stücheli-Herlach
(Organisationskommunikation &
Öffentlichkeit)
stue@zhaw.ch



Prof. Dr. Philipp Dreesen
(Digital Linguistics)
dree@zhaw.ch



Dr. Julia Krasselt
(Digital Linguistics)
krss@zhaw.ch