Risk Perception of COVID-19 in the German Internet Media and its epidemiological consequences during first wave of infections - exploration of possible research topic

Andrzej Jarynowski (✉ ajarynowski@gmail.com)

IBI

Research Article

Keywords: SARS-CoV-2, traditional media, social media, COVID-19, Germany

Posted Date: November 25th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-114934/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Risk Perception of COVID-19 in the German Internet Media and its epidemiological consequences during first wave of infections - exploration of possible research topic

Andrzej Jarynowski^{1,2}

1. Interdisciplinary Research Institute, Wroclaw, Poland

2. System Modeling Group, Institute of Veterinary Epidemiology and Biostatistics, Freie Universität Berlin, Germany

Abstract

Due to the spread of SARS-CoV-2 virus infection and COVID-19 disease, there is an urgent need to analyze epidemic perception in Germany. This would enable authorities for preparation of specific actions minimizing public health and economic risks. The aim of this article is to singal possible research activities for future research.

The aim of this study is to quantitatively investigate perception of COVID-19 in German Media (Twitter, Google, Youtube and selection of news articles) in the Internet by infodemiological approach. We proposed quantitative Media analysis as Retrospective and for future Prospective observatory analysis of secondary data. We attempt to analyze main discourses via natural language processing tools (such as topic modelling and sentiment analysis), multilayer and temporal network analysis of accounts/words/topics and time series analysis.

There were just a few previous works quantitatively linking Internet activities and risk perception of infectious diseases in Germany. Traditional and social media do not only reflect reality, but also create it. German authorities, having a reliable analysis of the perception of the problem, could optimally prepare and manage the social dimension of the epidemic.

The analysis of electronic media makes it possible to analyze the problem perception in Germany and early detect possible behavioral changes (e.g. panic) associated with the epidemic, which is crucial for a targeted response and tailored containment scenarios to minimize public health risks. Mistrust of governmental measures implementation has fuelled Querdenken movement - an unlikely alliance of far-right and left-wing, as well as conspiracy theorists.

Being aware of many shortcomings of computational/digital epidemiology and its exploratory approach, it provides us with an opportunity to analyze a huge amount of digital footprint data at low cost and in a short time.

Introduction to Internet research in infectious disease epidemiology

Internet research allows for quick analysis of Big Data. Due to the urgent need of empirical knowledge of the effect of COVID-19, this is the best trade-off between quality and time (Lopez, et al., 2020). We are sure that a mixed qualitative and quantitative approach provides better quality of research, but such research needs months or years to be prepared and conducted. Unfortunately, the virus does not give us such an amount of time. As one can provide collected data to the public, there will be a chance for deeper analysis later on for any other researchers.

Up to our knowledge there were just a few previous studies quantitatively linking the Internet activities and risk perception of infectious disease in Germany (Samaras, et al., 2020; Sudhakar, et al., 2014). Thus the present study is filling this gap with a quantitative "data driven" exploratory approach (Jarynowski, Buda, Paradowski, 2019).

Measuring risk perception could allow to estimate the probability of panic during a disease outbreak which can be mainly mediated by the following latent variables (Oh, et al., 2020):

- fear (e.g. fear of the unknown, or a sense of threat to oneself and loved ones);
- anger (e.g. anger at the condition of health care, mistakes of those in power, or negative side effects of physical distancing).

The spread of information and opinions (product life cycle) can also be like an epidemic (Jarynowski, Jankowski, Zbieg, 2015; Christakis, Fowler, 2008) starting with the phase of growing interest (so-called "early adoption"), continues to the phase of general interest (so-called "majority") to eventually popularity loss (so-called "laggards stage"). We will start with this

as a possible null model of social dynamics and extend or generalize it taking into account our observations for Germany and other European countries. The obtained results will potentially allow us to optimally manage such extraordinary emergency events as epidemics in the future and provide insights on the societal resilience mechanisms.

The purpose of quantitative media analysis in connection with the COVID-19 pandemic should be related mainly to:

- ability to adapt and change the lifestyle (eg. in the context of physical distance and adaptation of standard isolation rules).
- risk perception and communication (eg. fear and anxiety). Investigation of risk perception, in particular, panic and anxiety/fear, effectiveness of communication and information campaigns (cf. results of sentiment analysis);
- contact / hygiene behavior patterns and acceptance of contact tracing (see case of tracing App by Annan, et al. 2020)
- Investigation of reaction to mitigation scenarios, in particular dynamics of protest movements (as backlash phenomenon), interactions, conflicts, and polarization between key actors and information cascades.

Possible theories, working hypothesis and research questions

Social and traditional media can provide information and disinformation about the virus globally at unprecedented speed (traditional media even faster than social (Schultz, et al., 2011; Jarynowski, Wojta-Kempa, Belik, 2020; Jarynowski, et al. 2020)), fueling panic and creating so-called infodemics. Infodemiology is a new field of research supporting traditional surveillance and health monitoring (Effenberger, et al., 2020). As a result of the proposed infodemiological study we will prepare the analysis of COVID-19 awareness and risk perception in the Internet media in Germany, reflecting the evolution of social dynamics and its interaction with the actual epidemic situation.

Term "digital footprints" meaning entries in the Internet/social media across the entire spectrum of social dimensions – e.g., interpersonal/institutional relationships, the activities of social movements, or the ideological climate in a given community. Secondary data (here Internet sources) comprises various sectors of social life. The analysis of this material from selected categories of agents should be carried out within the social field theory (Diani, 2015) and Actor Network Theory (Latour, 1996). Those theories assume that the social behaviour constitutes the decisions of the actors that form a network of cooperation or conflict.

In Germany, there were protests in many cities for example against restrictions against COVID-19. For instance, demonstrations against racism or religious meetings took place without satisfying sanitary regimes. Moreover, people used the Internet to communicate before gathering. Often the protesters seem to be wildly misinformed about the virus have no trust in anti-covid measures being implemented. So media monitoring could be a method for preparing resources and safety cautions before such a protest could take place.

We target possible questions to:

- dynamics of interest in events related to the coronavirus in the world and in Germany;
- people's behavior towards the epidemic and each other;
- estimation by German population of threats related to virus transmission and its consequences for the country;
- information needs, fears and concerns, media coverage, compliance to preventive measures, conflicts in the public and political sphere, effectiveness of information campaigns and nudging;
- operation of various federal and land institutions (Government, Ministry of Health, etc.) and municipal (hospitals, public offices, president of cities, etc.);
- conspiracy theories, like fake news about the virus and foreign influence spreading in social networks;
- common dynamics of information and COVID-19 in Germany and the world.

The pandemic-related spread of misinformation and fear in the Internet is a real threat. We are dealing here with a time race with a spread of COVID-19 and spread of social moods. It seems that development of a sophisticated algorithm capable to prevent this threat is a challenging undertaking (Lipsitch, et al., 2020; Cinelli, et al., 2020). Our solution is quick, straightforward and probably cost-effective as well. We have distinguished 2 main modifiable variables: Information Campaigns and Spread of Information (misinformation and rumors) to be extracted from “digital footprints”.

Effective pharmaceutical mitigation strategies such as therapeutic treatment or a vaccine will not be available until the very probable secondary wave of infections in Autumn 2020 (Ferguson, et al., 2020; Jarynowski, et al., 2020). The only mitigation methods so far are contact reduction (e.g. isolation / quarantine, restrictions for travel or mass assembly) and the reduction of the probability of infection transmission (e.g. standard precautions such as hand hygiene, immunomodulation like sleeping well or use of protective equipment) and infection period reducing (e.g. tracking contacts and targeted testing). These strategies depend massively on risk perception in the population – which can be measurable at a relatively high precision with our approach (Lopez, et al., 2020). Moreover, according to the WHO Europe President (Kluge, 2020) “Authorities need to listen to their publics and adapt accordingly, in real-time” and our study is exactly fitting in this frame.

Similar studies on risk perception were launched already in the UK (<https://www.academia.edu/s/c7277878e4>) and USA (<https://around.uoregon.edu/content/study-will-look-perceived-risk-new-coronavirus-real-time>), as well as for Asian languages, however each society is different and results from Anglo-American culture could not be simply transferred to other countries (Rosiński et al., 2019). There are some qualitative studies for non-English speaking European Union countries (Lohiniva et al., 2020), but until the beginning of May 2020 (at least up to our knowledge and keywords search in Pubmed/medline with more than 150 non-EU and non-English oriented studies) there is not a single quantitative article similar to this approach. Moreover, European Commission, which is monitoring media attention on COVID-19 in the EU, use English social media material only for their publicly available reports (https://ec.europa.eu/jrc/sites/jrcsh/files/emm_covid-19_media_surveillance_-_30_april_2020.pdf). We plan to provide much more information on our responsive webpage and our data will be available for download by public health officers, other researchers as well as interested politicians and journalists. The EU have also launched projects (https://ec.europa.eu/info/sites/info/files/research_and_innovation/research_by_area/documents/ec_rtd_cv-projects.pdf) concerning modeling approaches and social surveys. More generally, our risk perception study will be useful complementary to other epidemiological studies such as simultaneous surveillance studies, epidemiologic field investigations, and case series (Lipsitch, et al., 2020).

Non-pharmaceutical interventions have been accompanied by anti-lockdown protesters, which in Germany (Jarynowski, Semonov, Belik, 2020) is a strange mix of pro-ecological left-wing liberals, conspiracy theorists, far-right extremists and ordinary citizens. Coronavirus measures were not the only reason for demonstrations and protests at this time. For instance, demonstrations against racism or BLM or religious meetings took place without compliance to sanitary norms, however they did not reach such a high attention as protest around Querdenken movement.

Possible Research Design

There are plenty of reasons for using Internet (traditional and social) media for monitoring the latest content on a broad topic of disease spread. Infodemiology – “science of distribution and determinants of information in an electronic medium, specifically the Internet, or in a population, with the ultimate aim to inform public health and public policy” according to the father of this discipline - Prof. Eysenbach (Eysenbach, 2009), with the primary aim of surveillance. Infodemiological analysis of epidemiological data is important to increase situational awareness (risk perception) and design suitable interventions.

Our approach targets a wide range of general population with a relatively high coverage of the Internet users with quite a significant audience variability across platforms. It's important to mention that our analysis operates on subjective populational perception and there is no direct translation between scientific evidence (e.g. the question of the effectiveness of protective masks in infection prevention) and colloquial knowledge describing infectious diseases and fear of acquiring an infection [Fig 2, 5]. Text analytics and processing could describe sociolinguistic picture of the society. Application of natural language processing techniques (NLP), such as topic modeling technique Latent Dirichlet Allocation and Machine Learning techniques suitable for text, sequential data and image analysis in combination with deep learning techniques could be applied if quality of the data will prove suitable. The labeling for e.g. sentiment analysis could be done by for qualitative analysis too.

The virus outbreak is accompanied by adherence to preventing behaviors influencing in its turn the epidemic spread, and the Internet is the main mediating mechanism. This study aims to investigate social perception of coronavirus in the Internet media during the epidemic. The number of internet users in Germany in 2020 is 65 millions and it covers 78% of the population. During COVID-19 pandemic, the Internet is even more important due to lockdown. The Internet is full of content, however there is no chance to analyze all of them. We choose social media and traditional media as representatives of various repertoires of users.

Target / Study Population,

There are 64,600,000 Internet users in Germany with a number of German speaking Internet users being 61,370,000 (according to Google Ads). Surveys on German Internet users (Statista, 2017) showed that the largest share of active Internet users is on Twitter (approx. 70%), the average share – on Facebook (approx. 50%) and the low share – on YouTube (approx. 15%). The passive representativeness of the Internet is relatively high, but active (own content creation) is biased towards younger age groups and women.

Twitter has about 11 million users in total and almost 2 millions are using Twitter daily (Kontor, 2020). Using Twitter as a sampling tool for the whole society will be efficient for age-groups between 15-40. We choose such keywords as coronavirus (and others which could later come on into use as Internet users are changing hashtags) with selection criteria for a language being German.

EventRegistry is crawling alone from Germany a few hundred thousand news webpages daily. We choose EventRegistry as a traditional media search engine because it has a large range of online magazines representing various political sites. In addition, it gives priority to the digital versions of other broadcasting channels, including television, radio or newspapers. Between February 10 and March 11 over 20 thousands representative articles were selected (the non-systematic sampling method was applied). Readers of selected articles can be representative for the population of 20-60 age cohort (APressInst, 2017). There are also alternative searching tools such as <https://medisys.newsbrief.eu/medisys/homeedition/pl/home.html> or <https://emm.newsbrief.eu> but none of them allows for intensive crawling.

Youtube has 77%-share among German Internet users and is responsible for 30% of total streaming (Klicksafe, 2019). YouTube is overrepresented by teenagers with an affinity index > 200 in this group (Klicksafe, 2019). We recommend to select videos on the subject of anti-covid protests.

Additionally we supplement our analysis quantitatively by Google Trends and qualitatively by Facebook posts (the most popular social network in Germany in terms of number of users - 32 million) and comments of articles. Despite the highest population penetration and the largest ranges, Facebook does not allow automated analysis, so it cannot be included in our analysis.

Telegram is an important communication channel for conspiracy theories and movements against the government measures.

Possible Data collection and methodology

For Twitter, one can collect (tweets and their metadata such as: Time – time of posting (with second); Tweet/Status - when a status message is shared on Twitter, Retweet - when a Tweet is re-shared by another specific user (and how many times this Tweet has been retweeted); Like - when a Tweet receives a 'like' from a specific user (and how many times this Tweet has been liked); Reply - If the represented Tweet is a reply (and how many times this Tweet has been replied to); user- The user who posted this Tweet; geography - Represents the geographic location (coordinates/place) of this Tweet as reported by the user; Quoting - when the Tweet is a quote Tweet (and how many times this Tweet has been quoted). Favorite - when a Tweet favorited by a specific user (and how many times this Tweet has been favorited).

Additionally one can parse Twitter to get a list of followers and followings as well as friends of a given account starting at least for most active accounts.

One can analyze retweet networks as proxies for known spreading of ideas from person to person in a social network (Christakis, Fowler, 2007) from a point of view of social influence, homophily and external field. One can correlate accounts and local societies information preferences (in terms of words used) taking into account age, language, and geo factors where possible (as both tweets and news can be geotagged). It will allow us to observe differences between federal states in Germany.

Each platform has its own bias, however Twitter is the 4th most popular (Kontor, 2020) social network in Germany and EventRegistry filters most important article webpages.

Ultimately, data analysis can be also biased due to involvement of media platform content presenting algorithms in the discourse. E.g. technology giants like Google, Twitter, Facebook are supposed to implement fact verification algorithms to filter out false information. Being aware of this, computational techniques of social sciences (Jarynowski, Buda, Paradowski, 2019), despite some shortcomings and their exploratory nature, provide us with the opportunity to analyze a huge amount of digital footprint data at low cost and in a short time.

To minimize bias we propose to apply triangulations approach of the Internet research (Digital Footprint) in the dimensions of

- techniques (quantitative such as data mining, network analysis and qualitative such as discourse analysis);
- data (Social Media/Content such as Twitter and Youtube, and Mainstream Media as digital news agency);
- researchers (Social Scientists, Data Analysts and Epidemiologists);
- theories (Social Field Theory, Action Network Theory, Theory of Risk Perception and Theory of Collective Action).

Fig.1. Media sources. Possible layers of investigation on Twitter (twitter.com), Google trends analysis (trends.google.com), EventRegistry information retrieval from news agencies (EventRegistry.org)

Statistical analysis is a crucial point in this kind of study. We propose to apply:

- the number and nature of social media events such as information queries (time series analysis);
- sentiment and conceptual fields analysis;
- social network analysis;

- topic modeling technique;
- influence of an external field.

Metadata includes attributes such as user, time location etc. One can extract/retrieve important information from the text too and possibly images. The following tasks will be done in R/python: computational analyses of the data, NLP and data mining (such as sentiment, conceptual fields analysis, semantics topic modeling technique), Social Network Analysis (such as analysis of the properties of the interaction networks and their topology, normalisation, cluster analysis, paths analysis, community detection, causality) multidimensional analysis such as PCA, LDA, identification of predictive variables, principal component analysis, computation of correlations and regression, machine learning models.

Open science ubuntu philosophy

Aggregated publicly available data will be mainly reported, but some publicly available content or Users/actors could be also individually investigated. However, reported controversial cases of social network and perception analysis (Sapiezynski, et al., 2019; Carpenter, 2009) occurred if feedback was given to a subject, which is not a case in our study. There are no known additional risks to those we plan to incorporate in this study.

Balanced metrics should be helpful in navigating the risk perception through reducing uncertainty.

Preliminary Results Google

A large part of the German population has heard about coronaviruses for the first time in January 2020. Most of the coronaviruses are harmless and only the emergence of a novel strain from Wuhan gave the word "Coronavirus" a new meaning. Interest in the Coronavirus in Germany appeared relatively late and is relatively low than in other countries [Fig. 2].

Fig. 2. The search intensity of the topic "Coronavirus" in Google for Poland, Italy, Spain, USA, UK, Germany and Sweden (15.01-29.04.20) generated using the Google Trends tool.

Media Intelligence helps to spot trends, to analyze any topic over a longer time span, to create sociolinguistic landscapes. In the corresponding semantic networks we recognize the most common co-occurring phrases in the search engine together with the word "Coronavirus". It tells us what are the first associations between COVID-19 and information needs [Fig. 3, 4].

individuals. Moreover, indirect messages of “trying to avoid” social gatherings may be interpreted differently, thereby putting some groups and areas at a higher risk of contagion [Fig. 4]. Therefore the investigation of understanding of the official state decisions, as well as feelings and attitudes towards the current COVID-19 situation is necessary.

Fig. 5. The intensity of queries with the phrases “hand washing”, “protective mask”, “hand disinfection”, “Vitamin C”, “cough” in German Google (15.01-03.05.2020) generated using the Google Trend tool.

We also illustrate the intensity of popularity of some concepts, within their social meaning in time as precautionary measures such as hand washing or protective masks [Fig. 5]. Other studies suggest that people search for professional solutions rather than simple, practical and effective ones such as keeping hand hygiene (Jarynowski, Wojta-Kempa, Belik, 2020). peak search for symptoms as cough and PPE as protective mask hand disinfection” was first time in end of February the same time when first outbreak in Heinsberg county and Ischgl ski resort were described in media (Szmuda, et al., 2020). Next wave of interest in protective mask appeared when some measure were announced in end of March.

Preliminary Results Twitter

For Twitter data we attempt to deploy the temporal multilayer network analysis (using metadata on quote, retweet, mention, reply, and following), because networks can be constructed based on the mentioned metadata types. To illustrate the application of Social Network Analysis (Wasserman, et al., 1994) methods to the Twitter data, we have built a network with vertices representing Twitter accounts and edges representing retweets [Fig. 6]. Such a network could reveal various connections (social impact, trust, friendship, etc.) between accounts being social actors and the characteristics of the actors (political affiliation, views, etc.). Propagation of information on a network can have a variety of factors e.g. emotional value (Jarynowski, Jankowski, Zbieg, 2015) or information/misinformation type (Pierri, et al., 2020). An unsupervised weighted Louvain algorithm (Blondel, et al., 2008) for community analysis revealed a polarization implying that governmental campaigns fall into the information bubble or echo chambers (Baumann, et al., 2020).

Table 1. 20 the most influential accounts in the German retweeting network with a hashtag #Coronavirus (14.04-06.05.2020) according to weighted degree centrality.

rank	account	weighted degree	rank	account	weighted degree
1	faznet	2474	11	PEI_Germany	637
2	RegSprecher	1413	12	heutejournal	631
3	SZ	1396	13	FAZ_Politik	618
4	lipsandskin	983	14	ZeitW	519

5	MalteKaufmann	974	15	berlinerzeitung	515
6	BAG_OFSP_UFSP	836	16	HerrNaumann	496
7	polenz_r	819	17	BR24	477
8	Daniel_Roi_AfD	811	18	MatthiasMeisner	472
9	AnneWillTalk	666	19	lawyerberlin	434
10	DLFNachrichten	656	20	OliverGeorgi	397

One can indicate influential actors on German Twitter according to various metrics such as network centrality measures [Tab. 1].

Fig. 6 Retweet networks for language de 19.03-07.06.2020. Selection of accounts $100 < \text{tweets} < 1000$ (around 1000 nodes). Then building a retweeting network by selecting around nodes (only giant component) with highest in-degree centrality. Colors according to community detection.

Preliminary Results News Articles

We selected Event Registry as a content aggregator because it contains a large database of Internet sources [Tab. 2]. Distribution the topics in first approach keywords [Fig. 7] and according to how they are changing with time.

Fig. 10 Theoretical attempt to describe main aspects of a German social system in Pandemic time. Rectangles are subpopulations, arrows mean main interactions, shadow variable are modifiable factors, signs represents positive or negative influence, crossing bars means interaction which are the most likely to disappear with time

To this end, we measure the relevant activity in the Internet news and social media. The revealed insights on social dynamics will enable us to devise urgently needed proper epidemic containment and information management strategies as well as to raise preparedness for future pandemics.

We depose more than 900 000 Twitter posts (Key inclusion criteria: language: de #coronavirus) are available for 19.03-07.06.2020 (some holes are still there) with columns tweet_id, date,user,is_retwitter, retweet_id, quote:

CSV <https://box.fu-berlin.de/s/K8Qp4mossdnjQ6y>

Data Frame (R) <https://box.fu-berlin.de/s/2GeJcwyC3MXGpMb>

References

- Amann J, Sleigh J, Vayena E. Digital contact-tracing during the Covid-19 pandemic: an analysis of newspaper coverage in Germany, Austria, and Switzerland. *medRxiv*. 2020 Jan 1.
- APressInst 2017, Print vs. digital subscribers: Demographic differences and paths to subscription, <https://www.americanpressinstitute.org/publications/reports/survey-research/print-vs-digital/>
- Baumann F, Lorenz-Spreen P, Sokolov IM, Starnini M. Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*. 2020 Jan 27;124(4):048301.
- Betsch C, Wieler LH, Habersaat K. Monitoring behavioural insights related to COVID-19. *The Lancet*. 2020 Apr 18;395(10232):1255-6.
- BfR 2020, Corona Monitor, <https://www.bfr.bund.de/cm/343/200407-bfr-corona-monitor.pdf>
- Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*. 2008 Oct 9;2008(10):P10008.
- Carpenter, M.A. Ethical considerations with social network analysis. in: *An Executive's Primer on the Strategy of Social Networks*. Business Expert Press 2009, 315.
- Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med*. 2007;357(4):370-379. doi:10.1056/NEJMsa066082
- Cinelli M, Quattrocioni W, Galeazzi A, Valensise CM, Brugnoti E, Schmidt AL, Zola P, Zollo F, Scala A. The covid-19 social media infodemic. *arXiv preprint arXiv:2003.05004*. 2020 Mar 10.

Diani, M. . *The Cement of Civil Society*. Cambridge: Cambridge University Press 2015.

Effenberger M, Kronbichler A, Shin JI, Mayer G, Tilg H, Perco P. Association of the COVID-19 pandemic with Internet Search Volumes: A Google Trends™ Analysis. *International Journal of Infectious Diseases*. 2020 Apr 17.

Eysenbach G. Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. *J Med Internet Res*. 2009;11(1):e11.

Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunubá Z, Cuomo-Dannenburg G, Dighe A. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team 2020.

Jarynowski A, Jankowski J, Zbieg A. Natural vs artificial viral spread within the online community. *E-methodology*. 2015;2:71–78. doi:10.15503/emet2015.71.78

Jarynowski A, Paradowski MB, Buda A. Modelling communities and populations: an introduction to computational social science. *Stud Metodol*. 2019;39:117-139. DOI: 10.14746/sm.2019.39.5

Jarynowski A, Buda A, Płatek D, Belik V. African Swine Fever Awareness in the Internet Media in Poland—exploratory review. *E-methodology* 6.6 (2019): 100-115.doi.org/10.15503/emet2019.100.115

Jarynowski A, Wójta-Kempa M, Płatek D, Belik V. Social Values Are Significant Factors in Control of COVID-19 Pandemic—Preliminary Results. *Preprints* 2020 <https://www.preprints.org/manuscript/202005.0036/v1>

Jarynowski A, Wójta-Kempa M, Płatek D, Czopek K. Attempt to understand public health relevant social dimensions of COVID-19 outbreak in Poland. *Society Register* 2020, 4(3), 7-44.

Jarynowski A, Wójta-Kempa M, Belik V.. Perception of “coronavirus” on the Polish Internet until arrival of SARS-CoV-2 in Poland, 2020 *Nursing and Public Health* 2020;10(2).
http://www.pzp.umed.wroc.pl/ahead-of-print/120054_en.pdf

Jarynowski A, Wójta-Kempa M, Belik V. Trends in Interest in COVID-19 on Polish Internet, *Polish Epidemiological Review* 74 (2), 106-123, 2020

Jarynowski A, Semenov A, Belik V. Protest perspective against COVID-19 risk mitigation strategies on the German Internet. *LNCS*, 2020

KickSafe 2019, Zahlen und Fakten - YouTube-Nutzung in Deutschland
<https://www.klicksafe.de/themen/kommunizieren/youtube/zahlen-und-fakten-youtube-nutzung-in-deutschland/>

Kluge H. Behavioural insights are valuable to inform the planning of appropriate pandemic response measures. 14.05.2020 <http://www.euro.who.int/en/media-centre/sections/statements/2020/statement-behavioural-insights-are-valuable-to-inform-the-planning-of-appropriate-pandemic-response-measures>

Kontor 2020, SOCIAL MEDIA 2020: AKTUELLE NUTZERZAHLEN, <https://www.kontor4.de/beitrag/aktuelle-social-media-nutzerzahlen.html>

Latour, B. On actor-network theory: A few clarifications. *Soziale welt* 1996, 369-381.

Lipsitch M, Swerdlow DL, Finelli L. Defining the epidemiology of Covid-19—studies needed. *New England Journal of Medicine*. 2020 Mar 26;382(13):1194-6.

Lohiniva, AI, Sane J, Sibenberg K, Puumalainen T, Salminen M. "Understanding coronavirus disease (COVID-19) risk perceptions among the public to enhance risk communication efforts: a practical approach for outbreaks, Finland, February 2020." *Eurosurveillance* 25, no. 13 (2020): 2000317.

Lopez CE, Vasu M, Gallemore C. Understanding the perception of COVID-19 policies by mining a multilanguage Twitter dataset. *arXiv preprint arXiv:2003.10359*. 2020 Mar 23.

Oh SH, Lee SY, Han C. The Effects of Social Media Use on Preventive Behaviors during Infectious Disease Outbreaks: The Mediating Role of Self-relevant Emotions and Public Risk Perception. *Health communication*. 2020 Feb 17:1-0.

Pierri F, Artoni A, Ceri S. HoaxItaly: A collection of Italian disinformation and fact-checking stories shared on Twitter in 2019. *arXiv Prepr arXiv200110926*. 2020. doi:10.7910/DVN/PGVDHX

Samaras L, García-Barriocanal E, Sicilia MA. Syndromic surveillance using web data: a systematic review. In: *Innovation in Health Informatics*. Elsevier; 2020. p. 39–77.

Sapieżyński, P., Stopczyński, A., Lassen, D.D. & Lehmann, S. (2019). Interaction data from the Copenhagen Networks Study. *Scientific Data* 6

Schultz F, Utz S, Göritz A. Is the medium the message? Perceptions of and reactions to crisis communication via twitter, blogs and traditional media. *Public relations review* 2011;37(1):20–27.

Statista 2017. Distribution of active and passive social media users in Germany 2016, by platform;
<https://www.statista.com/statistics/792357/social-media-active-passive-users-germany>

Sudhakar V Nuti, Brian Wayda, Isuru Ranasinghe, Sisi Wang, Rachel P Dreyer, Serene I Chen, and Karthik Murugiah. The use of google trends in health care research: a systematic review. *PloS one*, 9(10), 2014.

Szmuda T, et al., Assessing the Awareness of the Novel Coronavirus (COVID-19) in Germany: A Google Trends Analysis, 2020, preprint

Wasserman S, Faust K. *Social network analysis: Methods and applications*. Cambridge university press; 1994 Nov 25.

Figures

Figure 1

Media sources. Possible layers of investigation on Twitter (twitter.com), Google trends analysis (trends.google.com), EventRegistry information retrieval from news agencies (EventRegistry.org)

Figure 2

The search intensity of the topic "Coronavirus" in Google for Poland, Italy, Spain, USA, UK, Germany and Sweden (15.01-29.04.20) generated using the Google Trends tool.

Figure 3

Semantic web of the noun "coronavirus" in German Google search engine : LEFT (question) (snapshot on 11.06.2020). RIGHT (prepositions) snapshot on 07.05.2020).

Figure 4

Semantic web (prepositions) of the noun LEFT "Covidioten"/ RIGHT "Querdenken" in German Google search engine (snapshot on 25.08.2020)

Figure 5

The intensity of queries with the phrases "hand washing", "protective mask", "hand disinfection", "Vitamin C", "cough" in German Google (15.01-03.05.2020) generated using the Google Trend tool.

Figure 6

Retweet networks for language de 19.03-07.06.2020. Selection of accounts $100 < \text{tweets} < 1000$ (around 1000 nodes). Then building a retweeting network by selecting around nodes (only giant component) with highest in-degree centrality. Colors according to community detection.

Figure 7

Topics cloud from articles 08.04-07.05.2020 in German selected by EventRegistry.

Figure 8

News media topics distribution provided by EventRegistry.org for protest around 01.08 [Left] and around 29.08 [Right] demonstrations.

Theoretical attempt to describe main aspects of a German social system in Pandemic time. Rectangles are subpopulations, arrows mean main interactions, shadow variable are modifiable factors, signs represents positive or negative influence, crossing bars means interaction which are the most likely to disappear with time