



Project Report **VINFO**

We investigated the language and rhetoric used by alternative media channels affiliated with the Querdenken movement in Germany during the COVID-19 pandemic. With our results we hope to contribute combating the spread of disinformation, increase public trust in evidence-based public health measures, and foster informed decision-making.

Preface by the Supervisors	116
Journalistic part	118
Scientific part	120
Self-reflection	130
Process description	132

Team	Pia Gutsmedl Max Hampel Seong-Min Jun Cheng Pan Valentin Pauli Alexander Sobieska Rosa Josephine Weidenspointner
Tutors	Paul Sieber Junianna Zatsarnaja
Supervisors	Dr. Eleni Georganta Dr. Kristina Schick

Preface by the Supervisors

Dr. Eleni Georganta and Dr. Kristina Schick

Understanding how misinformation is created is particularly relevant in today's digital age where information can spread rapidly and widely, having significant impact on individuals, communities, and societies. Our group Team VINFO decided to explore this topic and empirically investigate the language and discourse utilized by the German anti-COVID movement "Querdenken." Based on the misleading information spread during the COVID pandemic, they focused on the representation of the German anti-COVID movement Querdenken in their prominent news sites and online blogs. Their empirical evidence can provide insights into how conspiracy theories and misinformation spread, making suggestions to help individuals become more critical consumers of information, identify and counteract false or misleading claims, and ultimately make more informed decisions.

Team VINFO has decided to use mixed methodology and investigate which linguistic features and strategies are used by Querdenken media and which evidence is brought into the discourse when reporting their perspective on the pandemic. Using a quantitative approach, they first performed a systematic quantitative analysis to explore the extent to which scientific terminology was used in Querdenken media. Second, applying a qualitative analysis, they evaluated how evidence was used to support the movement's claims.

As supervisors, we had the privilege of guiding this interdisciplinary group of young students who were highly motivated, committed and self-organized. Supervising this group was a great experience and all of us learned a lot from this project.



Supervisor insights

For me, **Eleni Georganta**, as a team researcher, it was an absolute delight to supervise a group of such motivated young students. Right from the beginning of the project, they showed great inspiration and determination, and their enthusiasm never dwindled throughout. Watching them grow both academically and as individuals was very rewarding. It was inspiring to see them put in so much effort towards their research, and seeing their hard work pay off, especially with the acceptance of presenting their findings at an international conference.

What is your research interest or motivation for science?

My research interest is centered on understanding the dynamics that emerge within teams. Whether it's conflict, innovative ideas, or motivation among individuals who don't know each other, my aim is to identify the essential mechanisms that promote suc-

cessful and effective collaboration. I believe that all team members can grow when these mechanisms are in place, and my research focuses on using evidence-based approaches to shape the future of teamwork for organizations.

What special experience from your studies/career would you like to share with the scholars?

Good research is only possible if you are passionate about the questions you are seeking to answer. It's essential to keep exploring and investigating the phenomena that you find most intriguing. Challenges and obstacles may arise, but every experience is an opportunity to learn and grow. Embrace the difficulties and always push yourself to overcome them. Remember that research is about discovering something new and contributing to the field, so always strive to ask and answer the questions that matter most to you.

Supervisor insights

I can only agree with Eleni. It was a great pleasure for me, **Kristina Schick**, to experience the spirit and the great commitment of the group to the topic in particular and to answering scientific questions in general. The work of the group was characterized by a great independence and a passion for knowledge which is second to none. Their project results, especially the opportunity to present their findings to the scientific community beyond the TUMJA community, are the fruits of their work over the past two years.

What is your research interest or motivation for science?

My research interest is to foster medical communication skills. With my research, I would like to contribute to the long-term improvement of medical care and to gain insights into how communicative competencies can be further fostered and improved

in the long term. For this purpose, I use different methods and concepts: On the one hand, I investigate how attitudes affect performance in patient conversations and how different learning settings, like simulation and workplace learning, can affect competence development.

What special experience from your studies/career would you like to share with the scholars?

In the research I have done on the issues of doctor-patient communication and professional identity formation, I have seen over and over again how important it is to speak the same language and what difficulties arise when people do not understand each other or do not want to understand each other. With your work, you contribute to reducing the communication gap with the Querdenken community and to raise awareness of the importance of phrasing and its interpretation. ■

This Title Can Manipulate You

"This vaccination must be an IQ test. You can bring arguments as you want, it's like talking to a wall ... unbelievable." This paragraph, written by a user named Juliane, can be found in the comments column of the website "corona-blog.net," which describes itself as "news from citizens for citizens." The author continues: "In any case, no one who has had the stuff given to them needs to whine all over me about any damage. If you're too lazy or too stupid to do the research, you'll just have to live with it."

The specific article that provoked this reaction is about SM-102: an ingredient of the Moderna vaccine, which, according to the sub-heading, can allegedly pose a risk to fertility and unborn children, and is also carcinogenic. Juliane is not the only one to vent her anger in the comments about this: while Christiane simply demands that the "pharmaceutical giants should be sued," Isabel suspects that the vaccine will make cells glow in order to "[...] scan people with it, for example at the airport or in the shopping mall." Birk, under another article referencing alleged deaths connected to mRNA vaccines, refers to biblical sources instead of his own thoughts:

"[...] To all politicians who present these deadly vaccinations as good, let this be said: 'Woe to those who call evil good and good evil, who put darkness for light and light for darkness, who put bitter for sweet and sweet for bitter!' (Isaiah 5:20)."

Emotions are running high. But why? What dark secrets has the "Corona Blog" revealed about the pharmaceutical industry? If you look at the content of the article, you will find one main source for the claim that the substance SM-102 is supposed to be toxic: A safety data sheet from Cayman Chemicals. The article points out that the following warning can be found there: "For research purposes only, not for human or veterinary use." As a layman, this is enough to send shivers down your spine, as a chemical whose use in humans is explicitly prohibited here is being added to a widely administered vaccine. In addition, according to the safety instructions, this chemical is said to be carcinogenic, teratogenic

and toxic to the liver – the list is long and threatening. However, if you look at the document with a bit of background knowledge, you will see that the dangers do not come from the substance SM-102 at all, but from the solvent, more precisely chloroform, in which it is dissolved. Apart from the fact that chloroform is a solvent and therefore not present in the final vaccine, it is not even explicitly mentioned that this SM-102 chloroform solution is used to manufacture the Moderna vaccine. These are several small but very important differences, but they reveal the central claim of this article to be based on misunderstanding. So how can we explain the fact that many readers react in such an outraged way to the article's "revelations," even though they are not revelations at all?

Part of this can probably be attributed to the blog's position as part of the "alternative media." These mostly see themselves as a corrective to the mainstream media and thus attract readerships that are highly suspicious of it. Linked to this is the phenomenon of confirmation bias, which describes a tendency to value information as correct if it reinforces existing positions. Since a dislike of vaccines led most Corona Blog readers to seek out the website in the first place, they will also be more willing to acknowledge its alleged revelations as such. Another such cognitive bias is the Dunning-Kruger effect, a phenomenon in which people with low skills, expertise, or experience in a particular type of task or area of knowledge tend to overestimate their competence or abilities. In this case, ignorance and superficial knowledge regarding vaccines leads both the author and the readers to falsely see themselves as experts in the field and, as a result of this distorted perception, to express themselves as such. In the words of Juliane, "If you're too lazy or too stupid to do the research, you'll just have to live with it."

The language of the text can also be a major factor influencing the apparent credibility and persuasiveness of the "Corona Blog." After all, how a statement is received depends not only on the content, but also – and to the same, if not greater, extent – on the way it is conveyed. Language is, as Wilhelm von Humboldt said, the

"ultimate medium of thinking and of one's world view." The list of things we can use language for is endless. Among other things, we can use language to communicate our thinking and world view to others, but we can also use language to influence, or worse, manipulate the thinking and world view of others.

The ways in which language can be shaped take a variety of forms. For example, so-called "othering," a creation of group identities, aims to antagonize vaccination advocates: "Us" versus "them". This, combined with an emotionally charged vocabulary that is negative in tone and at the same time enriched with technical terms from science, creates an almost culinary harmony: The bitter outrage is made palatable to the reader by the sweet and multifaceted taste of scientific legitimacy, which docks directly onto the Dunning-Kruger taste buds. Such rhetorical morsels frequently appear within the text of the Corona Blog, as exemplified in the following paragraph, which attempts to put the article's alleged revelations into perspective:

"In view of these points, one can only be speechless. Does anyone really still believe that the Corona vaccination is about health? Why inject millions of healthy people with such poison – without even being able to guess at the long-term effects?"

The fact that these linguistic devices often occur in texts from alternative media such as the Corona Blog is not merely an assumption, as our research project shows. We were able to prove quantitatively that language with both an emotive tone and peppered with scientific terms is used significantly more often here than in established media. But the question still remains: Do these scientific statements have any substance, or are they ultimately just smoke and mirrors intended to cloud the reader's view? We have examined this on a random basis for 40 other articles to get an overview of how often sources are misrepresented.

Language, as the foundation of our communication, is omnipresent and its influence is enormous. If linguistic means are used skillfully, a text can carry a very high power of persuasion or manipulation. Of course, this influence of language comes with dangers, which can be seen in propaganda, among other things. Enemy images can quickly arise from othering, which often leads to racism and discrimination. The influence of language on our opinion is not something bad per se, we just have to be aware of the dangers to be able to prevent them. At the same time, we can create something good by strengthening communication in the right places and making it more persuasive, for example in the mediation and resolution of conflicts or in the field of science communication. Then, in the end, everyone would be a little smarter and the misunderstandings reported at the beginning of this text would not have arisen in the first place.

What would be the link to Juliane now? Let's try to see the world from her point of view: Frustrated by the Corona policy of the German government, she looks on the internet for alternatives to mainstream media, which she doesn't trust anymore. There, she comes across an article that not only reinforces her skepticism about mRNA vaccines (keyword: confirmation bias), but also expresses a dissatisfaction about the status quo that one cannot find in "conventional" media. At the same time, a rhetorical aura of scientific seriousness hangs over the text, elevating it above the level of contentless polemics and leaving her with the feeling that she might be onto something big. It seems that if you understand these distinctly human mechanisms that make you susceptible to dubious information, it is easier to understand the thought processes of the people that fall victim to it. After all, these mechanisms are universal: No one is free of cognitive distortions, and no one can claim to see the world objectively and free of judgment. A small part of Juliane is in each of us – perhaps we should be aware of this so that we notice early enough before it comes to the surface. ■

Decoding the discourse

Abstract

This study investigates the language and rhetoric used by alternative media outlets affiliated with the Querdenken movement in Germany during the COVID-19 pandemic. We analyzed articles from Querdenken, tabloid, and newspaper outlets using dictionary-based analyses to examine emotional tone, health-related language, and usage of scientific terminology. Further, we conducted a qualitative analysis to assess the originality and congruence of sources cited. Results show that Querdenken outlets use a more scientific language, exhibit a more negative emotional tone, and have a higher usage of health-related language compared to other outlets. Additionally, we found a significant mismatch between sources and claims in the interpretation of information. Our study highlights the important role of alternative media in shaping public discourse and opinion and emphasizes the need for improved communication of scientific information and regulation of alternative media outlets.

Introduction

The COVID-19 pandemic has been and continues to be one of the biggest challenges our world has faced in recent memory. At the time of writing, over 760 million confirmed cases and over 6.9 million deaths have been reported to WHO (World Health Organization, 2023). While its impacts seem to be predominantly health-related, its societal effects should not be overlooked. The toll that prevention measures such as lockdowns or school closings took on the economic and social life of many citizens led to them voicing their concerns and displeasures about their situation in various ways, be it on the street or on conventional and social media (Douglas, 2021). One group that emerged out of this growing discomfort is the so-called “Querdenken” movement, a loosely organized German-speaking collective made up of people from all sociological and political walks of life whose attitudes range from skepticism to extremism and radicalism (Marko, 2022). They are united by their concern about governmental and pharmaceutical interventions meant to curb the spread of the pandemic (Nachtwey et al., 2020). It is this movement that sparked controversy through actions such as demonstrative “strolls” through cities during lockdown (Finkbeiner, 2022), the formation of a political party (“DieBasis”) explicitly positioned in opposition to COVID-19 measures (Niedermayer, 2023), and their fervent activism and vaccine skepticism online.

For many movements that claim to be opposed to the mainstream, alternative media plays an important role. Alternative media serves as a “self-perceived corrective” (Holt et al., 2019) to what is seen as the current political or journalistic mainstream. This is hardly a clear-cut distinction, as there exists an overlap between the worlds of alternative and mainstream media (Harcup, 2005), but it does serve as an important descriptor of a collection of news sources that cater to those with serious distrust of mainstream organizations. Members of the Querdenken community can be described as such (Nachtwey et al., 2020). This distrust creates a demand for alternative COVID-19 narratives, leading to widespread proliferation of online outlets within Querdenken groups. These outlets are used as sources of evidence for claims on vaccine efficacy, virus danger, and adverse effects of measures. We want to examine these outlets to understand their communication methods and the use of outside sources to legitimize their views.

There is widespread concern about a rising prevalence of disinformation (Sousa-Silva, 2022), something the World Health Organization has coined the “infodemic” (World Health Organization, 2020) – a perception that has led to an increase in research on the topic. Especially now in the context of the COVID-19 pandemic, disinformation such as conspiracy theories thrives more and more as people feel threatened, uncertain, and insecure (Douglas, 2021). Alternative media has been identified as a frequent source of disinformation (Frischlich et al., 2020; Grinberg et al., 2019) and existing research has examined its organizational dimensions (Figenschou & Ihlebæk, 2019; Harcup, 2005) as well as its content (Haller & Holt, 2019; Nygaard, 2019). Since conspiratorial thinking is well established within the Querdenken community (Nachtwey et al., 2020), it is also worth mentioning the similar body of research pertaining to conspiracy theories and theorists (Douglas et al., 2019; Douglas & Sutton, 2018; Fong et al., 2021). Within these studies, it has been discovered that even brief exposure to conspiratorial content can have concerning effects, such as lower social and civic engagement and a greater likelihood of engaging in science rejection (Jolley & Douglas, 2014; Lewandowsky et al., 2013; Lewandowsky & Oberauer, 2016; van der Linden, 2015). Therefore, it is urgent to analyze them more closely to find ways to counter them effectively (Sunstein & Vermeule, 2009). It is important to note

that most of this research addresses anglophone communities. A major difference between German- and English-speaking anti-mainstream movements seems to be a strong predisposition to right-wing populism in English communities (Kerr & Wilson, 2021) that is not as present in the Querdenken movement, as most of their members report having voted for left-liberal parties in the last Bundestag election before the pandemic (Nachtwey et al., 2020). Thus, this movement seems to have certain unique attributes that warrant a specific analysis of its properties.

Existing research has mainly employed content analysis (Boberg et al., 2020; Quandt et al., 2020). However, not only the content, but also how it is conveyed, more precisely the language of the Querdenken movement, should be examined in more detail. Drawing on the framework of Critical Discourse Analysis (CDA), which underlines the role of language in the construction of knowledge and societal norms (Baker & McGlashan, 2020), we want to concentrate on the language used within alternative media outlets that are used to support Querdenken arguments. Past linguistic examinations of this ecosystem have been relatively small-scale and/or have focused more on the language of readers than on the outlets themselves (Douglas & Sutton, 2018; Marko, 2022; Thiele, 2022; Zollo et al., 2015).

In a qualitative analysis, Marko found that in a COVID-19 conspiracy group, linguistic features that are typical for extremist groups, such as ideological in-group and out-group presentation and the agentless passive voice which aims to evoke fear, are used. Other linguistic techniques include the misrepresentation of scientific knowledge and the use of colloquial language, among other things, to establish a connection between members of the group and to distance themselves from the government and the rest of the population, whom they perceive as enemies and oppressors.

In this study, connecting this previous, qualitative work with computational methods, we wish to answer two research questions: Which linguistic features and strategies are used by Querdenken media in the discourse to legitimize their views on the pandemic? Which evidence is brought into the discourse when reporting on their perspective on the pandemic?

To examine the language of alternative media outlets connected to Querdenken (henceforth referred to as “Querdenken media”), we performed a dictionary-based analysis of their language on a dataset of 25,934 articles using a proprietary dictionary for scientific terms as well as the Linguistic Inquiry and Word Count dictionary (LIWC) (Pennebaker et al., 1999) to examine emotional tone and, as a completely novel approach, scientificity of language, respectively, in comparison to datasets of mainstream news and tabloid news (3,241 articles in total). Furthermore, we conducted a case study on a small subset of our Querdenken media articles to examine their usage of sources qualitatively.

Methods

Data

Data were gathered for three different media sets: mainstream media, science journalism publications, and outlets connected to Querdenken. To create the Querdenken media set, a web scraper was used to crawl the websites of Report24, Uncut-News, Rubikon, Transition News, and Corona Blog, and capture all pertinent data, including the title, author, publication date, and complete text of each item. The generated dataset consists of 25,934 articles with a relationship to COVID-19, the Querdenken movement, or a far-right online news network in Germany. These articles are either directly or indirectly related to COVID-19. This was accomplished by including articles that were unmistakably about COVID-19, like those from Corona-Blog, those having the words “corona,” “pandemic,” or “COVID,” or those with a content column showing their emphasis on the COVID-19 pandemic.

For the mainstream media sample, two outlets were included, Der Tagesspiegel and Bild, which were scraped using the academic access to LexisNexis. A constructed week sampling approach was utilized to select a sample of articles from March 2020 to December 2021 for both publications. Constructed week sampling is a widely used sampling technique in content analysis, which involves the creation of a random sample of weeks over a given time period (Luke et al., 2011). The resulting dataset, which is based on the same filtering method as described in the preceding paragraph, consists of 3,241 newspaper and tabloid items that were either directly or indirectly related to the COVID-19 pandemic in Germany.

Zeit Wissen and Spektrum der Wissenschaft are the two science journalism publications that were used in this study. Zeit Wissen is a popular science publication in Germany that focuses on subjects like technology, medicine, and the natural world. In contrast, Spektrum der Wissenschaft is a monthly magazine that covers recent advances in science and technology. Its 12,000-page sample spans the years 2008 through 2022. No additional filtering was done on the scraped articles because the main goal of this sample is to create a frequency table for scientific terminology.

Ethical Considerations

This research project addressed ethical considerations regarding the sensitivity of the available data. To prevent the spread of potentially harmful information, we decided not to make the dataset publicly available. In addition, we avoided disclosing any detailed identifying information to safeguard the identity and privacy of the people featured in the publications.

Methodology

Dictionary Creation

A dictionary was designed for a systematic and quantitative analysis of scientific language used by Querdenken media, and to provide insight into the extent to which they use scientific terminology in their discourse.

In doing this, the words included in the dictionary were obtained from word frequency tables of articles from the field of science journalism taken from samples of the media Spektrum der Wissenschaft and Zeit Wissen. All words in our sample with a frequency greater than 100 were evaluated independently by five judges and classified into one of the following categories or rejected: (1) professions and institutions associated with sciences, (2) research-related terms, (3) life sciences terminology, and (4) terminology from other sciences. The five judges reviewed the lists they had compiled for each publication together and eliminated any duplicate submissions. The final result was a joint dictionary combining the lists from both publications.

Quantitative Analysis

In this study, we conducted a quantitative analysis of the language used in the COVID-19-related articles collected from five Querdenken-affiliated websites (Report24, Uncut-News, Rubikon, Transition News, and Corona Blog), as well as articles from a newspaper

(Der Tagesspiegel) and a tabloid (Bild), spanning from March 2020 to December 2021.

Given that the number of articles per month varied across the different media outlets, we decided to average the values of LIWC and dictionary categories per month and outlet type (Querdenken media, tabloid, and newspaper). We then performed statistical analyses using linear mixed models with fixed effects for the month, the publication type (tabloid, newspaper, Querdenken outlet), and the random effect for an outlet (Uncut-News, Corona Blog, Rubikon, Tagesspiegel, Bild, etc.) for the LIWC categories and our self-developed dictionary categories. Post hoc, we also computed contrasts (ANOVAs) to clarify the differences between the publication types, independent of the temporal trend.

Qualitative Analysis

Furthermore, a qualitative analysis of the claims made by Querdenken media outlets (Corona-Blog, Report24, Rubikon, Transition News, and Uncut-News) was conducted to evaluate how cited sources were used to support the movement's claims. To ensure a sufficient degree of objectivity, a codebook consisting of five categories and providing instructions for the evaluation process was created. A minimum of 50 claims was analyzed for each Querdenken medium, including 50 claims from Corona-Blog, 51 claims from Report24 and Rubikon, 50 claims from Transition News, and 52 claims from Uncut-News.

Articles were chosen from the Querdenken media dataset for study by the assessors using a random number generator. Only articles with a connection to COVID-19 or SARS-CoV-2 were included. Articles with the following focus were excluded: letters to the editor, excerpts from discussion forums, interviews, or articles that linked exclusively to a video. This ensured that the cited sources were analyzed in the context of the author's argumentation.

All article claims were listed and numbered consecutively. A maximum of ten claims per article was analyzed. If the number of claims was greater than 10, ten claims were randomly selected from the article to avoid bias. Each claim was evaluated based on three criteria: whether it was a translation or copy of an existing article, whether sources were cited, and whether the sources were freely available without any restrictions like a paywall or subscription.

When a claim had a source, that source was examined, and if the material cited any other references, those references were examined as well. The last reference accessible in this chain of citations was used as the main source for the qualitative analysis. Five criteria were used to determine source-citation congruence: whether the article and source's content were linked (topic), whether there was a logical link between the article and source (logic), whether all relevant information was taken from the source (leave), whether no non-source content was added (add), and whether the information from the source was interpreted correctly in the sense of the author of the source (interpretation).

Throughout the analysis, the assessors took notes on the article and the source in a separate document to justify their evaluations. This approach ensured a rigorous and systematic evaluation of the claims made by Querdenken media outlets, providing valuable insights into the use of evidence in the movement's discourse on the pandemic.

For the qualitative analysis of the claims made by Querdenken media outlets, descriptive statistics were displayed for the different categories per outlet and in total. The five categories used to evaluate the source-citation congruence included topic, logic, leave, add, and interpretation. Furthermore, data on degrees of separation, whether a source was given for the claim, and the type of source (article, blog, governmental source, etc.) were also collected.

Measures

In this study, we employed a self-developed dictionary for the German language to measure the use of scientific language in texts, which we believe is a novel approach to understanding the discourse surrounding the COVID-19 pandemic. This approach is particularly relevant, as previous research suggests that Querdenken-affiliated media may use pseudo-scientific statements to support their perspectives on the pandemic, positioning themselves as enlightened, while characterizing those who disagree with them as blindly following the mainstream narrative (Marko, 2022).

Our dictionary consists of four categories: (1) professions and institutions associated with sciences, (2) research-related terms, (3) life sciences terminology, and (4) terminology from other sciences. To compute the values for each of these categories, we used the Linguistic Inquiry and Word Count (LIWC) software to analyze each

article in our dataset of science journalism media, to be named Spektrum der Wissenschaft and Zeit Wissen. The number of entries per category and examples from each can be found in Table 1.

Category	Number of entries	Examples
Professions and institutions associated with sciences	28	“biologen,” “dr.,” “forschungsinstitute”
Research-related terms	68	“analyze,” “beweis,” “experiment”
Life sciences terminology	93	“dna,” “enzym” “immunsystem”
Terminology from other sciences	53	“atom,” “exoplaneten,” “frequenz”

Table 1: Number of entries per category and examples

To assess the internal consistency of each of our self-developed dictionaries, we used a subset of our science journalism sample dataset containing around 3,000 pages not used for the dictionary creation. We calculated both Cronbach's alpha and the Kuder-Richardson Formula 20 (KR-20) for each dictionary category. Each word included in the lexicon was measured as a percentage of total words per text (Cronbach's α) or, alternatively, in a binary "present versus absent" manner (Kuder-Richardson Formula 20; Kuder & Richardson, 1937). The raw, unstandardized Cronbach's alpha and KR-20 values for each dictionary category are presented in Table 2.

It is important to note that the traditional Cronbach's alpha method, calculated from relative word frequencies, tends to underestimate reliability in language categories due to the highly variable base rates of word usage within any given category, which may explain the moderate to high but not perfect values for Cronbach's alpha observed in some categories (Pennebaker et al., 2007). However, the Kuder-Richardson Formula 20 is generally considered a better approximation of each category's true internal consistency and reports a value close to 1 for the life sciences terms and research-related terms categories, indicating a very high internal consistency. The professions and other sciences terminology categories also

showed acceptable internal consistency with KR-20 values of 0.32 and 0.66, respectively. The Cronbach's alpha values were moderate to high for all categories, ranging from 0.52 to 0.87.

LIWC

The Linguistic Inquiry and Word Count (LIWC; Boyd et al., 2022; Pennebaker et al., 2007) is one of the most widely used computational methods to convert text to psychological constructs. LIWC is a text analysis program consisting of over 80 dictionaries for linguistic, psychological, and topical categories indicating various social, cognitive, and affective processes (for a comprehensive list of the dictionaries, see Table 1 in Pennebaker et al., 2007). To analyze a text, LIWC calculates the percentage of words in the text that match a dictionary word, out of the total number of words in the text. For example, the positive emotion dictionary includes 408 stems such as “amaz*”, “excit*”, etc. In a six-word text “I love going to the dentist,” the output by LIWC is 16.6% for the positive emotion dictionary (i.e. one positive emotion word “love” divided by six total words in the text), 16.6% for the health dictionary (i.e. one health word, “dentist”), and 0% for the negative emotion dictionary. Past empirical studies on the validity of LIWC have found that it is reliably able to detect meaning from text in a wide variety of contexts, as well as detect emotional states, intentions, motivations, thinking styles, and individual differences (Pennebaker et al., 2015). We used the latest available LIWC dictionary for German (LIWC 2015), more specifically we used the categories emotional tone and health, generating values for each article in our sample.

The emotional tone category in LIWC refers to the extent to which language expresses emotions, such as joy, sadness, anger, or anxiety. As a combined dictionary, its value expresses how positive or negative a text’s emotional valence is: Texts with a score higher than 50 usually are understood to be more positive in their tone, texts with

a score lower than 50 are understood to be more negative in their emotional tone (Pennebaker et al., 2007). This category is particularly relevant to the study of discourse around the COVID-19 pandemic, as emotions are known to play a significant role in shaping public opinion and attitudes toward health issues (Hase & Engelke, 2022).

The health category in LIWC refers to the use of language related to health and well-being, such as terms related to illness, medication, and medical procedures. In the context of our study, analyzing the health category allows us to examine the extent to which media outlets are focused on discussing the actual health implications of the pandemic. We included this category as an addition to our self-developed dictionary.

Results

All data and graphics are available in the supplementary material at <https://www.ja.tum.de/en/ja/projects/2022/vinfo/>.

Quantitative Results

LIWC: Emotional Tone

We fit a linear mixed-effects model with emotional tone as the response variable, month and type as fixed effects, and outlet as a random effect. The model revealed a significant effect of type on emotional tone, indicating that Querdenken outlets had a significantly lower mean emotional tone (Estimate = -13.388, SE = 4.691, p = 0.0437) than newspapers. No significant effect was found for the month of publication (Estimate = 0.002431, SE = 0.002537, p = 0.3400). The estimate for the effect of outlet type on emotional tone remained significant after adjusting for multiple comparisons using the Tukey method (Querdenken outlets vs. newspapers: z = -2.854, p = 0.00864). In addition, we found a significant difference in mean emotional tone between tabloids and Querdenken outlets (z = 4.830, p = 4.1e-06); only the difference between tabloids

Metric	Professions and institutions associated with science	Research-related terms	Life sciences terminology	Terminology from other sciences
Cronbach’s Alpha (Raw)	0.52	0.87	0.78	0.53
Kuder-Richardson Formula 20	0.32	0.79	0.84	0.66

Table 2. Internal consistency metrics for each dictionary category

and newspapers ($z = 1.534$, $p = 0.12501$) delivered no significant results. All analyses were carried out in R version 3.0.2 using the lme4 package (version 1.0-5).

The results of the linear mixed-effects model revealed that the emotional tone of articles was significantly predicted by the type of outlet. Specifically, articles published in Querdenken outlets had a significantly more negative tone than articles published in newspapers ($b = -13.39$, $t = -2.85$, $p < 0.05$), while articles published in tabloids had a significantly more positive tone than articles published in newspapers ($b = 9.27$, $t = 1.53$, $p > 0.05$). The month of publication did not significantly predict emotional tone ($b = 0.00$, $t = 0.96$, $p > 0.05$). The strongest predictor of emotional tone was outlet type, with Querdenken outlets consistently exhibiting a more negative tone than the other types of outlets.

LIWC: Health

The linear mixed-effects model with health as the response variable, month and type as fixed effects, and outlet as a random effect, showed that the type of outlet significantly predicted the mean health score of articles. Specifically, articles published in Querdenken outlets had a significantly higher mean health score than articles published in newspapers (Estimate = 0.00..., SE = 0.00..., $p = 0.019$), while articles published in tabloids did not significantly differ from those published in newspapers (Estimate = 0.00..., SE = 0.00..., $p = 0.492$). The month of publication did not significantly predict the mean health score (Estimate = 0.003..., SE = 0.002..., $p = 0.643$). The strongest predictor of health score was outlet type, with Querdenken outlets consistently exhibiting a higher mean health score than the other types of outlets.

The estimate for the effect of outlet type on the variable of health remained significant after adjusting for multiple comparisons using the Tukey method (Querdenken outlets vs. newspapers: $z = 3.818$, $p = 0.0004$). In addition, we found no significant difference in health-related language usage between tabloids and newspapers ($z = 0.157$, $p = 0.4489$), while the difference between tabloids and Querdenken outlets was significant ($z = -2.847$, $p = 0.0088$). The results of the linear mixed-effects model revealed that there were significant differences in health-related language usage between the types of outlets. Querdenken outlets had a significantly higher mean usage of health-related language than newspapers, while tabloids did not significantly differ from newspapers.

Dictionary: Professions and Institutions

A linear mixed-effects model was fitted to the data with the dictionary variable for the usage of professions and institutions associated with sciences as the response variable, month and type as fixed effects, and outlet as a random effect. The results show that the number of professions and institutions mentioned in articles was significantly influenced by the month of publication (Estimate = -0.0002... SE = -0.00001..., $p = 0.00862$), but not by the type of outlet (Querdenken outlets: Estimate = 0.0002... SE = 0.00001..., $p = 0.08852$; tabloids: Estimate = -0.0007... SE = -0.00001..., $p = 0.63386$). However, the type of outlet for Querdenken outlets can be seen as approaching significance.

Post-hoc testing using the Tukey method revealed no significant differences in the number of professions and institutions mentioned between newspapers and Querdenken outlets or between newspapers and tabloids. However, the difference between Querdenken outlets and newspapers approached significance ($z = 2.232$, $p = 0.0767$). Overall, these findings indicate that the month of publication is a significant predictor of the number of professions or institutions mentioned in articles, but the type of outlet does not have a significant effect.

Dictionary: Research-Related Terms

We fitted a linear mixed-effects model with research-related terms as the response variable, month and type as fixed effects, and outlet as a random effect. The model revealed a significant effect of type on the use of research-related terms, indicating that Querdenken outlets had significantly higher mean usage of research-related terms (Estimate = 0.252, SE = 0.105, $p = 0.0477$) than newspapers. Additionally, there was a significant effect of month on the use of research-related terms (Estimate = -0.0001722, SE = 0.0000572, $p = 0.0032$). No significant difference was found between tabloids and newspapers (Estimate = 0.002298, SE = 0.1349, $p = 0.9864$). After adjusting for multiple comparisons using the Tukey method, the estimate for the effect of outlet type on the use of research-related terms remained significant (Querdenken outlets vs. newspapers: $z = 2.411$, $p = 0.0477$). Moreover, a significant difference in the mean usage of research-related terms was found between tabloids and Querdenken outlets ($z = -2.389$, $p = 0.0477$), but no significant difference was found between tabloids and newspapers.

The results of the linear mixed-effects model revealed that the usage of research-related terms in articles was significantly predicted by the type of outlet and the month of publication. Specifically, articles published in Querdenken outlets had significantly higher usage of research-related terms than articles published in newspapers ($b = 0.252$, $t = 2.411$, $p < 0.05$). In contrast, no significant difference was found between articles published in tabloids and newspapers ($b = 0.0023$, $t = 0.017$, $p > 0.05$). The month of publication significantly predicted the usage of research-related terms ($b = -0.0001722$, $t = -3.008$, $p < 0.01$). The strongest predictor of research-related term usage was outlet type, with Querdenken outlets consistently exhibiting a higher usage of research-related terms than the other types of outlets.

Dictionary: Life Sciences Terminology

The linear mixed-effects model with the usage of life sciences terminology as the response variable, month and type as fixed effects, and outlet as a random effect, showed that neither type of outlet or month significantly predicted their mean usage ($p > 0.6$).

Dictionary: Terminology from Other Sciences

The linear mixed-effects model with the usage of terminology from other sciences as the response variable, month and type as fixed effects, and outlet as a random effect, showed that neither type of outlet or month significantly predicted their mean usage ($p > 0.6$).

Qualitative Results

In this study, a qualitative analysis was conducted on five Querdenken outlets, namely Corona-Blog, Report24, Rubikon, Transition News, and Uncut-News. The study examined the extent to which the claims made by these outlets were original or copied from other sources, the types of sources used to support these claims, the number of clicks required to reach the original sources, and the extent to which the sources cited were congruent with the claims made.

The analysis found that Corona-Blog had the lowest percentage of claims that were translations or copies of other articles (18%), while Uncut-News had the highest percentage (86.54%). The other outlets had intermediate levels of copied content, ranging from 25.49% to 52%. This finding suggests that there is considerable variation in the originality of the claims made by these outlets.

The source types used to support the claims made by Querdenken outlets also varied. Authorities, papers or studies, and newspaper articles were the most common types of sources cited by Corona-Blog and Rubikon, while newspaper articles, blogs, and press releases were the most common types cited by Report24. Transition News predominantly cited newspaper articles, but other source types were also present in the dataset. Uncut-News mainly cited papers or studies and newspaper articles, with other source types also present.

The number of clicks required to reach the original sources also varied among the outlets. For most outlets, the median number of clicks required was 1, indicating that the sources were relatively easy to access. However, Uncut-News had a median of 2 clicks, indicating that the sources were less accessible.

In terms of source-citation congruence, the study found that the percentage of claims with no thematic link between the article and the source was generally low across all outlets. However, the percentage of claims with no logical link ranged from 8% (Transition News) to 36.96% (Corona-Blog). Moreover, 22.58% (Report24) to 71.43% (Uncut-News) of claims added additional information not found in their original source, and 20% (Transition News) to 71.43% (Uncut-News) of claims did not interpret the information correctly in the sense of the author of the source. These findings suggest that the outlets varied in the extent to which they accurately represented the sources cited.

Overall, the mean source-citation congruence (SCC) score ranged from 2.81 to 3.66, with a merged dataset having a mean SCC score of 3.39. The mean number of clicks was 1, the median number of clicks was 1.33 with a range from 1 to 7 clicks necessary. This indicates that while the outlets differed in the extent to which they accurately represented their sources, the overall level of source-citation congruence was moderate. Furthermore, given the rather low mean and median number of clicks needed, this shows that original sources were rather easily accessible.

Discussion

The COVID-19 pandemic has had significant societal impacts, with citizens expressing their concerns through movements like the "Querdenken" movement in Germany. These movements often rely on alternative media sources to substantiate their claims and cater

to individuals who distrust mainstream organizations. However, the prevalence of disinformation in alternative media has become a growing concern, prompting research on its content, organizational dimensions, and effects on conspiracy thinking. Despite this, limited attention has been given to the language and rhetoric employed by the Querdenken movement. This study aims to fill this gap in the literature by conducting a comprehensive analysis of the linguistic features and strategies utilized by Querdenken media outlets to legitimize their views on the COVID-19 pandemic.

Existing research suggests that political sources play a prominent role in constructing pandemic news, indicating a strong reliance on mainstream media coverage of political elites (Mellado et al., 2021). However, it remains unclear how those who feel unrepresented by or disagree with the mainstream media or political establishment construct their own coverage of the pandemic. Preliminary research on the Querdenken movement indicates that seemingly scientific evidence is frequently employed by their media outlets to support their own lines of thinking (Marko, 2022). However, a thorough examination of the language and rhetoric used by the movement is lacking, and this study aims to address this gap by providing a detailed analysis of their language and sources of evidence.

The findings from the quantitative analysis in our study reveal important insights into the linguistic strategies employed by Querdenken media outlets. Querdenken media outlets exhibit a significantly lower emotional tone compared to newspapers, suggesting a focus on presenting objective information rather than appealing to readers' emotions. Tabloids, on the other hand, display a significantly more positive tone, potentially indicating a bias toward promoting a specific perspective or agenda. These results are comparable to previous studies such as Douglas (2021) indicating that a key driver behind pandemic news coverage is negative emotions. Querdenken media outlets also demonstrate a significantly higher mean health score than newspapers, suggesting a heightened emphasis on health-related issues and their implications. However, there were no significant differences observed between tabloids and newspapers in terms of health focus.

The analysis of research-related terms indicates that Querdenken media outlets employ these terms more frequently than newspapers, potentially indicating a stronger emphasis on scientific lan-

guage and evidence-based claims. However, the framing of the pandemic changed over time, leading to a shift from a research and science-focused narrative to a more political one. This shift may have influenced the usage of research-related terms in media coverage, highlighting the dynamic nature of media discourse during crises. In a similar approach, Pan & Meng (2016) examined the role of media framing during different stages of a flu pandemic. Taking a crisis management perspective, the authors analyzed news coverage to understand how media frames influenced public perceptions and responses to the health crisis, emphasizing the importance of effective crisis communication strategies in managing public health emergencies.

The qualitative analysis provides additional insights into the types of sources used by Querdenken media outlets. The study finds that the percentage of original content varies widely across outlets, with some relying heavily on external sources while others generate the majority of their own content. Authorities and papers or studies are the most frequently cited sources in Querdenken media outlets, indicating an attempt to provide credibility and legitimacy to their claims. However, the source-citation congruence score reveals a generally poor alignment between claims and their sources, suggesting a lack of consistency and potential misrepresentation of information.

Overall, the findings suggest that Querdenken media outlets employ specific linguistic features and strategies to legitimize their views on the pandemic. These outlets appear to prioritize health-related issues and use scientific language to create an illusion of scientific legitimacy. However, misrepresentation of scientific findings, interpretation of information, and discrepancies between sources and claims were observed. This raises concerns about the manipulation and misinterpretation of scientific content to further their agenda.

The results of this study have significant implications for crisis communication and public health. They highlight the important role that alternative media outlets, particularly those affiliated with the Querdenken movement, play in shaping public opinion and discourse surrounding the COVID-19 pandemic. The study underscores the need to improve the communication of scientific information, ensuring its accurate interpretation and implications in the wider context. This includes the use of clear and accessi-

ble language, considering cultural context and social factors, and working with local communities and stakeholders.

To combat the spread of disinformation and conspiracy theories, policymakers should consider implementing measures to regulate alternative media outlets and promote media literacy education. These efforts are crucial in enabling individuals to distinguish between credible and non-credible sources of information, empowering them to make informed decisions about their health and well-being. Furthermore, as our qualitative analysis shows, a substantial amount of coverage in Querdenken media was translated and not original, underlining the need for international cooperation and guidelines.

It is important to acknowledge the limitations of this study. The findings are based on data from Germany, and caution should be exercised when generalizing the results to other countries with different political cultures and systems. The study focuses exclusively on the Querdenken movement and its affiliated alternative

media outlets, and future research should explore other protest movements and compare their language use. Additionally, while the study identifies linguistic features and strategies, it does not directly test their perceived legitimacy or impact on public opinion. Further research is needed to investigate the psychological effects of these linguistic features and strategies on individuals' beliefs and behaviors.

In conclusion, this study provides valuable insights into the language and rhetoric employed by Querdenken media outlets. The findings highlight the urgent need for policymakers and health professionals to address the issue of disinformation and conspiracy theories in the context of the COVID-19 pandemic. By improving the communication of scientific information, implementing measures to regulate alternative media outlets, and investing in media literacy education, we can combat the spread of disinformation, increase public trust in evidence-based public health measures, and foster informed decision-making. ■

References

- Baker, P., & McGlashan, M. (2020). Critical discourse analysis. In *The Routledge Handbook of English Language and Digital Humanities*. Routledge.
- Boberg, S., Quandt, T., Schatto-Eckrodt, T., & Frischlich, L. (2020). Pandemic Populism: Facebook Pages of Alternative News Media and the Corona Crisis - A Computational Content Analysis. *ArXiv*. <https://www.semanticscholar.org/paper/Pandemic-Populism%3A-Facebook-Pages-of-Alternative-A-Boberg-Quandt/08b-769b0017cd515b2f3fa802fb460cfb2208a63>
- Boyd, R. L., Ashokkunar, A., Seraj, S., & Pennebaker, J. W. (2022). The Development and Psychometric Properties of LIWC-22.
- Douglas, K. M. (2021). COVID-19 conspiracy theories. *Group Processes & Intergroup Relations*, 24(2), 270–275. <https://doi.org/10.1177/1368430220982068>
- Douglas, K. M., & Sutton, R. M. (2018). Why conspiracy theories matter: A social psychological analysis. *European Review of Social Psychology*, 29(1), 256–298. <https://doi.org/10.1080/10463283.2018.1537428>
- Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology*, 40(S1), 3–35. <https://doi.org/10.1111/pops.12568>
- Figenschou, T. U., & Ihlebæk, K. A. (2019). Challenging Journalistic Authority: Media criticism in far-right alternative media. *Journalism Studies*, 20(9), 1221–1237. <https://doi.org/10.1080/1461670X.2018.1500868>
- Finkbeiner, F. (2022). Corona-Protteste, Verschwörungsmythen und Antisemitismus. In K. Trittel, P. Scharf, & FoDEx: Forschungs-und Dokumentationsstelle zur Analyse politischer und religiöser Extremismen in Niedersachsen (Hrsg.), *Demokratie-Dialog: Werkstattbericht FoDEx* (S. 51–59). Göttingen University Press. <https://doi.org/10.17875/gup2022-1944>
- Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & van der Linden, S. (2021). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, 24(4), 606–623. <https://doi.org/10.1177/1368430220987596>
- Frischlich, L., Klapproth, J., & Brinkschulte, F. (2020). Between Mainstream and Alternative – Co-orientation in Right-Wing Populist Alternative News Media. In C. Grimme, M. Preuss, F. W. Takes, & A. Waldherr (Hrsg.), *Disinformation in Open Online Media* (S. 150–167). Springer International Publishing. https://doi.org/10.1007/978-3-030-39627-5_12
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>
- Haller, A., & Holt, K. (2019). Paradoxical populism: How PEGIDA relates to mainstream and alternative media. *Information, Communication & Society*, 22(12), 1665–1680. <https://doi.org/10.1080/1369118X.2018.1449882>
- Harcup, T. (2005). "I'm Doing this to Change the World": Journalism in alternative and mainstream media. *Journalism Studies*, 6(3), 361–374. <https://doi.org/10.1080/14616700500132016>
- Hase, V., & Engelke, K. M. (2022). Emotions in Crisis Coverage: How UK News Media Used Fear Appeals to Report on the Coronavirus Crisis. *Journalism and Media*, 3(4), Article 4. <https://doi.org/10.3390/journalmedia3040042>

- Holt, K., Ustad Figenschou, T., & Frischlich, L. (2019). Key Dimensions of Alternative News Media. *Digital Journalism*, 7(7), 860–869. <https://doi.org/10.1080/21670811.2019.1625715>
- Jolley, D., & Douglas, K. M. (2014). The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions. *PLoS ONE*, 9(2), e89177. <https://doi.org/10.1371/journal.pone.0089177>
- Kerr, J. R., & Wilson, M. S. (2021). Right-wing authoritarianism and social dominance orientation predict rejection of science and scientists. *Group Processes & Intergroup Relations*, 24(4), 550–567. <https://doi.org/10.1177/1368430221992126>
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160. <https://doi.org/10.1007/BF02288391>
- Lewandowsky, S., Gignac, G. E., & Oberauer, K. (2013). The Role of Conspiracist Ideation and Worldviews in Predicting Rejection of Science. *PLoS ONE*, 8(10), e75637. <https://doi.org/10.1371/journal.pone.0075637>
- Lewandowsky, S., & Oberauer, K. (2016). Motivated Rejection of Science. *Current Directions in Psychological Science*, 25(4), 217–222. <https://doi.org/10.1177/0963721416654436>
- Luke, D. A., Caburnay, C. A., & Cohen, E. L. (2011). How Much Is Enough? New Recommendations for Using Constructed Week Sampling in Newspaper Content Analysis of Health Stories. *Communication Methods and Measures*, 5(1), 76–91. <https://doi.org/10.1080/19312458.2010.547823>
- Marko, K. (2022). Extremist language in anti-COVID-19 conspiracy discourse on Facebook. *Critical Discourse Studies*, 0(0), 1–20. <https://doi.org/10.1080/17405904.2022.2110134>
- Mellado, C., Hallin, D., Cárcamo, L., Alfaro, R., Jackson, D., Humanes, M. L., Márquez-Ramírez, M., Mick, J., Mothes, C., I-Hsuan LIN, C., Lee, M., Alfaro, A., Isbej, J., & Ramos, A. (2021). Sourcing Pandemic News: A Cross-National Computational Analysis of Mainstream Media Coverage of COVID-19 on Facebook, Twitter, and Instagram. *Digital Journalism*, 9(9), 1261–1285. <https://doi.org/10.1080/21670811.2021.1942114>
- Nachtwey, O., Schäfer, R., & Frei, N. (2020). Politische Soziologie der Corona-Proteste. *SocArXiv*. <https://doi.org/10.31235/osf.io/zyp3f>
- Niedermayer, O. (2023, Januar 25). Basisdemokratische Partei Deutschland. Bundeszentrale für politische Bildung: bpb.de. <https://www.bpb.de/themen/parteien/wer-steht-zur-wahl/berlin-2023/517362/basisdemokratische-partei-deutschland/>
- Nygaard, S. (2019). The Appearance of Objectivity: How Immigration-Critical Alternative Media Report the News. *Journalism Practice*, 13(10), 1147–1163. <https://doi.org/10.1080/17512786.2019.1577697>
- Pan, P.-L., & Meng, J. (2016). Media Frames across Stages of Health Crisis: A Crisis Management Approach to News Coverage of Flu Pandemic. *Journal of Contingencies and Crisis Management*, 24(2), 95–106. <https://doi.org/10.1111/1468-5973.12105>
- Pennebaker, J., Boyd, R., Jordan, K., & Blackburn, K. (2015). The Development and Psychometric Properties of LIWC2015. <https://doi.org/10.15781/T29G6Z>
- Pennebaker, J., Chung, C., Ireland, M., Gonzales, A., & Booth, R. (2007). The Development and Psychometric Properties of LIWC2007.
- Pennebaker, J., Francis, M., & Booth, R. (1999). Linguistic inquiry and word count (LIWC).
- Quandt, T., Boberg, S., Schatto-Eckrodt, T., & Frischlich, L. (2020). Pandemic News: Facebook Pages of Mainstream News Media and the Coronavirus Crisis -- A Computational Content Analysis (arXiv:2005.13290). *arXiv*. <https://doi.org/10.48550/arXiv.2005.13290>
- Sousa-Silva, R. (2022). Fighting the Fake: A Forensic Linguistic Analysis to Fake News Detection. *International Journal for the Semiotics of Law - Revue Internationale de Sémiotique Juridique*, 35(6), 2409–2433. <https://doi.org/10.1007/s11196-022-09901-w>
- Sunstein, C. R., & Vermeule, A. (2009). Conspiracy Theories: Causes and Cures*. *Journal of Political Philosophy*, 17(2), 202–227. <https://doi.org/10.1111/j.1467-9760.2008.00325.x>
- Thiele, D. (2022). Pandemic Populism? How Covid-19 Triggered Populist Facebook User Comments in Germany and Austria. *Politics and Governance*, 10(1), 185–196. <https://doi.org/10.17645/pag.v10i1.4712>
- van der Linden, S. (2015). The conspiracy-effect: Exposure to conspiracy theories (about global warming) decreases pro-social behavior and science acceptance. *Personality and Individual Differences*, 87, 171–173. <https://doi.org/10.1016/j.paid.2015.07.045>
- World Health Organization. (2020). Novel Coronavirus (2019-nCoV): Situation report, 13. World Health Organization. <https://apps.who.int/iris/handle/10665/330778>
- World Health Organization. (2023, März 21). WHO Coronavirus (COVID-19) Dashboard. <https://covid19.who.int>
- Zollo, F., Novak, P. K., Del Vicario, M., Bessi, A., Mozetič, I., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Emotional Dynamics in the Age of Misinformation. *PLOS ONE*, 10(9), e0138740. <https://doi.org/10.1371/journal.pone.0138740>

Self-reflection

Looking back on the past 18 months, we faced many challenges leading to valuable insights and lessons learned. Lessons that were important for the personal growth of each one of us but which might also help future teams from #class23, #class24, and so on.

But first things first: we, as team VINFO, are incredibly thankful for the help of our supervisors Dr. Kristina Schick and Dr. Eleni Georganta, as well as our tutors Paul Sieber and Junianna Zatsarnaja. They provided us with close guidance, especially Paul Sieber and Junianna Zatsarnaja, who visited nearly every Jour fixe and assisted us with their critical thinking. It was very important for us to hear the opinion of a third party, which is also capable of raising questions and pointing out areas where we still had work and brainstorming to do. This allowed us to stay on track and keep a level of consistency so that we managed to produce results that we are all very proud of. Also, we are very thankful to TUMJA, in particular Peter Finger and Dennis Lehmann, for funding our trip to Copenhagen for the 9th International Conference on Computational Social Science where we will have the opportunity to present our methods and findings.

Our individual expectations were similar, but it took us a lot of time at the beginning to figure out our project objectives. We researched all sorts of directions in the field of the “Querdenken” movement,

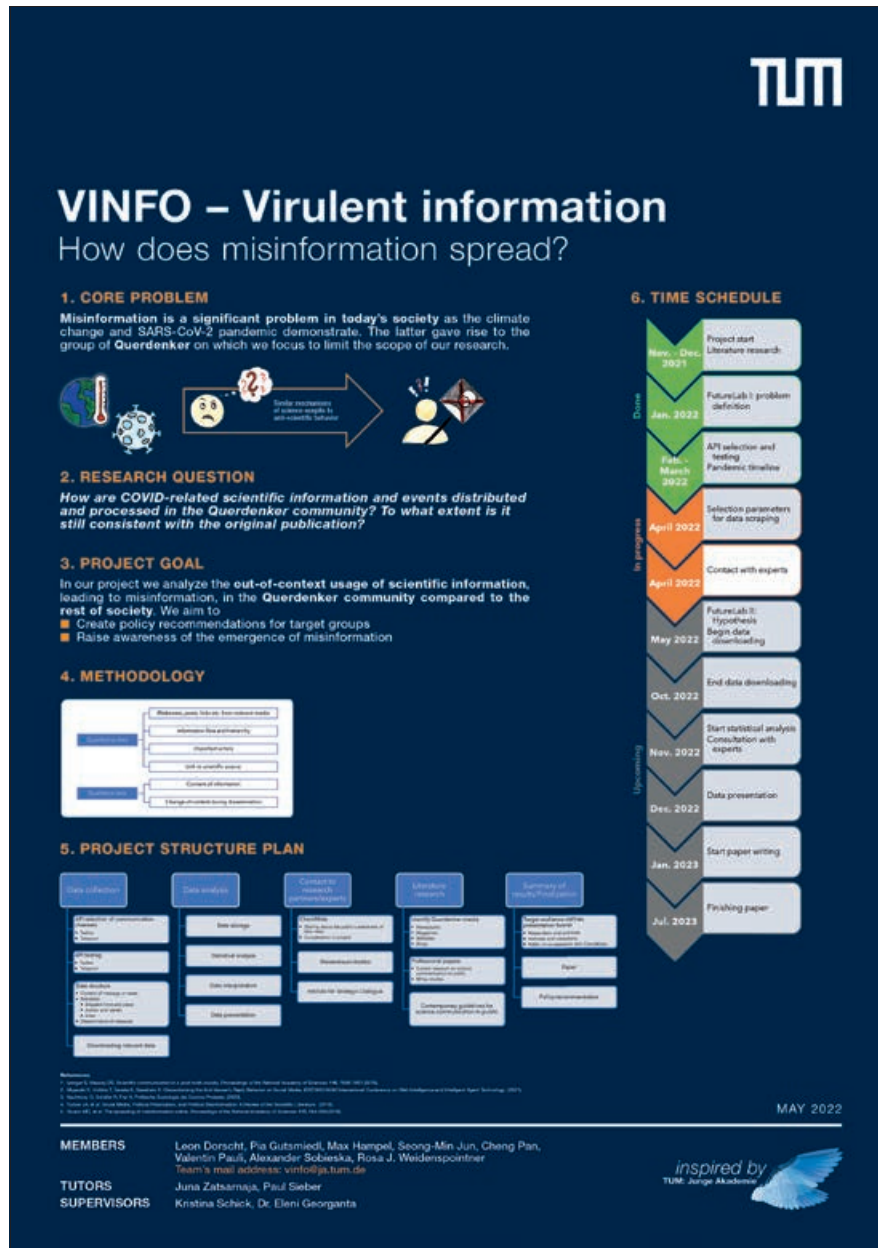
so we had many different ideas and leads. For us, the identification of a topic was not a linear process but rather iterative. To give an example: we started working on the idea of analyzing the language of the “Querdenken” movement and structures on social media or messengers like Twitter and Telegram. We even coded the APIs to obtain the data. A few months in, we decided to stop pursuing this lead due to unsatisfactory results. Consequently, we then developed the idea of analyzing articles coming from various websites associated with the “Querdenken” movement – the idea we then stuck to until the end.

At the beginning, we were hoping for a linear process in establishing our topic. However, looking back at that phase of our project, the months that we worked on the first idea weren't wasted at all but were necessary – in fact, essential – for us to settle on a topic that truly interested us and was, at the same time, feasible. In addition to that, there were a lot of lessons we learned from reflecting on how we worked as a team. We learned how to communicate effectively, how we should assign tasks, what we should talk about in plenum, and when it is better to divide and conquer. So as a tip for all the following classes: Before thinking about a topic for too long, it might be helpful to start working on it first and iteratively come back to reassess your objectives. This phase helped us a lot in terms of teamwork even though we didn't stick to our initial focus.

A further big learning outcome of our work together has been our collaborative team dynamics. We all have different study backgrounds, personalities, strengths, and weaknesses. As a result, we had different rhythms concerning our university schedules. Throughout the whole project, there were phases where some teammates couldn't commit as much time as others, and team members even lived in other time zones. Thus, we as a team needed to plan well ahead and at the same time stay flexible enough to ensure that everybody could manage to participate and engage. Thus, we changed the time and weekday of our Jour fixe several times, dividing tasks and giving enough freedom to the respective sub-teams to work out a schedule by themselves while maintaining the deadlines. We applied this concept throughout the whole period as it was easier to work as a smaller working group to achieve some milestones. For example, we assigned teams to work on data collection, development of tools, qualitative and quantitative analyses, writing texts, managing external relationships, etc. This felt very natural, too, as the tasks fell to the strengths of the team members who gained a certain level of expertise in the fields of their studies. So having all of those people coming from different backgrounds meant that we had to plan and communicate well to stay in rhythm but it also meant that everybody could focus on what they were good at and what they wanted to work on.

However, this self-reflection would be rather boring if we only pointed out the good things, right? So, what were the challenges throughout the project that we struggled with? A challenge was time commitment during intensive periods of studying, working, etc., and consequently also keeping up with deadlines. As we are a relatively large team, consisting of 7 team members, it was easy to stay in the background and let other people do the work. Especially during the early stages, when our weekly meetings had to be held via Zoom due to COVID-19, it was more difficult for us as a team to meet the deadlines due to a less engaging format. Connecting that with exam periods was the perfect mixture for forgetting project work. What helped us a lot with this problem were the seminar weekends as they served as intensive brainstorming sessions where we could all come together for two days and push our project back on track. Also, we organized a team weekend for working on our project and enjoying team activities.

Overall, we feel well-prepared for future scientific projects in interdisciplinary teams, as well as for further specializations during a Master's or Ph.D. program. We value the contacts we made, both scientific and friendly, and will stay in touch. ■



POSTER 1:

After the Kick-Off event and several brainstorming sessions where we combined the ideas from our different disciplinary backgrounds, we started with the literature research on Querdenken platforms and their language. We established our research questions, project goals, and project structure plan. At first, we performed API (Application Programming Interface) selection and testing to get access to data from Twitter and Telegram. Later we rejected this idea and just looked at Querdenken websites. At the beginning of 2022, we created a timeline with important dates during the pandemic, e.g. new variants, the first vaccine, and others. On the poster, you can see all the information about our project structure.

POSTER 2:

In the summer, after getting a lot of feedback internally from TUMJA and, during the Symposium of class 21, for the first time externally from the public, we specified two research questions: To what extent is COVID-19-related information presented in a scientific manner in Querdenken online journalism? To what extent is this information aligned with information from scientific publications? We finished the download of several thousand articles from newspapers, tabloids, and blogs. ■

TUM

VINFO – Virulent information

How does misinformation spread?

1. WHAT HAPPENED SO FAR?

- More concrete definition of our project goals:
 - >> Analysis of positive and negative tone
 - >> Analysis of scientific language
 - >> Analysis of text similarity to primary literature/articles
- Formulation of these goals into a research plan, concretization of the research question
- Downloading of most of the necessary text data for the construction of the library and for testing
- Figuring out solutions for web scraping to semi-automatically fetch text from Querdenker websites
- Regular "four fix" meetings to increase team member engagement

2. RESEARCH QUESTION

To what extent is COVID-19-related information presented in a scientific manner in Querdenker online journalism? To what extent is this information aligned with information from scientific publications?

3. METHODOLOGY

Step 1: Choosing a Dictionary for Scientific Texts	Research: Zot Wozen (~1000 articles, 1970-2020) Scraping the Internet: ~120 articles, 2011-2020	Qualitative Text Processing: Building with Expert System	Validation: Zot Wozen (~1000 articles, 1970-2020) Scraping the Internet (~120 articles, 2011-2020)	Result: Universal Dictionary Measuring Scientific Texts in News Publications
Step 2: Measuring Scientific Texts & Emotionality	Research: Sentimental effects (positive, negative, neutral) Control Group: Tobias (2018, 2019) & Neumann (2019, 2020)	Quantitative Analysis: Measuring Emotionality (EMV) & Scientific Tone (Step 1) in Querdenker & Control Group Using Querdenker online articles are tone-negative that positive in their sentiment tone and Querdenker online articles are positive in their use of scientific language compared to control articles and scientific newspaper articles		
Step 3: Text Similarity Analysis of Selected Articles	Training Data: Identifying Original Publications Measured in Querdenker Articles	Quantitative Text Similarity: Using Querdenker online articles as a model for original newspaper articles and original newspaper articles as a model for original newspaper articles Qualitative Text Similarity: Using Querdenker online articles as a model for original newspaper articles and original newspaper articles as a model for original newspaper articles		

Bibliography

1. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.
2. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.
3. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.
4. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.
5. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.
6. Borchers, M. (2018). A. The growing of scientific texts. *Proceedings of the 12th International Conference on Text Mining*, 1-10.

4. TIME SCHEDULE

Jan. 2020	• Problem definition
Feb. 2020	• (Re-)Defining Hypotheses for our research question
Mar. 2020	• Establish contact to experts • Pick relevant media • Start of dictionary construction
Apr. 2020	• End of dictionary construction • Test dictionary • Construct qualitative methodology
May 2020	• Download newspaper articles
June 2020	• Pick software tools to analyze the data
July 2020	• Start analyzing • Consultation with experts
Aug. 2020	• Data presentation
Sept. 2020	• Start paper writing
Oct. 2020	• Finish the research paper

OUR TEAM

MEMBERS

Leon Dorasch, Pia Gutsmiedt, Max Hampel, Seong-Min Jun, Chang Pan, Valentin Pauli, Alexander Sobieska, Rosa J. Wiedenspoortner
Team's mail address: vinfo@ia.tum.de

TUTORS


Jana Zatsarnaja, Paul Sieber

SUPERVISORS

Dr. Kristina Schick, Dr. Elani Georgantia

SEPTEMBER 2022

inspired by
TUM: Judge Akademie




VINFO – Virulent information

Understanding the language of misinformation

1. WHAT IS OUR PROJECT ALL ABOUT?

Misinformation is a widely recognized problem in today's society. We aim to better understand misinformation by analyzing a part of it which might be vital to its persuasiveness: Its language. For this, we aim to look at COVID-19-related online journalism from publications belonging to the German "Querdenken" movement, analyzing the scientificity and emotionality of their language using dictionary approaches. Furthermore, we want to look more closely at a subset of the articles to trace back their claims to their sources and evaluate how they process information.




2. PROJECT PROGRESS AND NEXT STEPS

Task	Progress
Raw data collection • Article scraping, downloading	Completed
Scientific dictionary construction • Generation of frequency tables, word picking	Completed
Dictionary analysis of collected data • Scientificity, emotionality	In progress
Statistical analysis	Future milestone
Codebook construction • Guidelines for performing source-citation congruence analysis	Completed
Source traceback & SCC analysis • Manual traceback and analysis of source-citation congruence (SCC)	In progress
Statistical analysis	Future milestone
Evaluation of methodology by TUM Stat • Seeking expert opinion on methodology and analysis	Completed
Presentation of results • Scientific report, publishing a scientific paper	Future milestone

■ Quantitative analysis
 ✓ Completed
 ■ Qualitative analysis
 ⌚ In progress
 ■ General milestones
 → Future milestone

References:
 1. J. A. Han, "Misinformation: A Review of the Literature," *Journal of the American Medical Association*, vol. 323, no. 1, pp. 1-10, 2020.
 2. J. A. Han, "Misinformation: A Review of the Literature," *Journal of the American Medical Association*, vol. 323, no. 1, pp. 1-10, 2020.
 3. J. A. Han, "Misinformation: A Review of the Literature," *Journal of the American Medical Association*, vol. 323, no. 1, pp. 1-10, 2020.
 4. J. A. Han, "Misinformation: A Review of the Literature," *Journal of the American Medical Association*, vol. 323, no. 1, pp. 1-10, 2020.
 5. J. A. Han, "Misinformation: A Review of the Literature," *Journal of the American Medical Association*, vol. 323, no. 1, pp. 1-10, 2020.

OUR TEAM




MEMBERS
Pia Gutsmedt, Max Hampel, Seong-Min Jun, Cheng Pan, Valentin Pauli, Alexander Sobieska, Rosa J. Weidenspointner
Team's mail address: vininfo@tum.de

TUTORS
Jana Zehsarnaja, Paul Sieber

SUPERVISORS
Kristina Schick, Dr. Eleni Georganta

JANUARY 2023

inspired by
TUM Junge Akademie 

POSTER 3:

Between the second and third posters, we achieved great progress. Especially during a weekend that we spent together in Füssen, we had lots of time to work on our project, develop new ideas, and strengthen the bonds in our group. For the methodology of our quantitative analysis, we constructed a dictionary that contains scientific terms to measure scientific tone. We familiarized ourselves with the methodology of the well-established LIWC-dictionary. For the qualitative analysis, we designed a code book to trace back randomized Querdenken articles and evaluate how they cite sources/evidence. We consulted "TUM|Stat" for the statistical analysis. On the poster, you can see the next steps beginning in January 2023. ■



VINFO – Virulent information

Decoding the Discourse: Analyzing the Linguistic Features and Strategies behind the Querdenken Movement's COVID-19 Narrative

SUMMARY

We investigated the language and rhetoric used by alternative media channels affiliated with the Querdenken movement in Germany during the COVID-19 pandemic. For quantitative analysis, we used a self-developed dictionary to examine emotionally, health-related language and usage of scientific terminology. For qualitative analysis, we analyzed the originality and congruence of sources cited.

IMPACT & SUSTAINABILITY

- >> Our results fills the research gap in literature by providing a comprehensive analysis of the linguistic features utilized by Querdenken media outlets.
- >> We will present our results at the Symposium and the International Conference on Computational Social Science in Copenhagen in July 2023.
- >> With our results we hope to contribute combating the spread of disinformation, increase public trust in evidence-based public health measures, and foster informed decision-making.

RESEARCH LIFE CYCLE



RESULTS AND OUTCOME

Data shows that Querdenken outlets use more scientific language, exhibit a more negative tone and make higher usage of health-related language compared to other outlets. Also, we found a significant mismatch between claims in the Querdenken outlets and their sources.

References:
1. Besser, T. E. (2019). The role of social media in the spread of misinformation. *Journal of the American Medical Association*, 321(12), 1155-1156.
2. Besser, T. E., & Smith, J. (2019). The role of social media in the spread of misinformation. *Journal of the American Medical Association*, 321(12), 1155-1156.
3. Besser, T. E., & Smith, J. (2019). The role of social media in the spread of misinformation. *Journal of the American Medical Association*, 321(12), 1155-1156.
4. Besser, T. E., & Smith, J. (2019). The role of social media in the spread of misinformation. *Journal of the American Medical Association*, 321(12), 1155-1156.
5. Besser, T. E., & Smith, J. (2019). The role of social media in the spread of misinformation. *Journal of the American Medical Association*, 321(12), 1155-1156.



OUR TEAM

MEMBERS

Pia Gutschmidt, Max Hampel, Seong-Min Jun, Cheng Pan, Valentin Pauli, Alexander Sobieska, Rosa J. Weidenspointner
Team's mail address: vinfo@tum.de

TUTORS

Jana Zatsarnaja, Paul Sieber

SUPERVISORS

Kristina Schick, Dr. Eleni Georganta



MAY 2023

POSTER 4:

In the last part of our project, we tested the tone and word usage of Querdenken and tabloid articles with LIWC and our dictionary and performed the source-citation congruence analysis and statistical tests. We achieved amazing results that are shown on our poster. We were very focused on writing our research report, including the journalistic and scientific parts. Our journalistic part was reviewed by Thomas Fromm, a journalist from "Süddeutsche Zeitung" which was helpful and insightful. We submitted an abstract of our findings to the "9th International Conference on Computational Social Science" and we were very happy to be accepted and will give a presentation in Copenhagen in July 2023. ■