

Final project report

Algorithmic content selection in Switzerland – a study of Google and YouTube

Mykola Makhortykh*, Maryna Sydorova*, Aleksandra Urman**, Franziska

Keller*, Silke Adam*

* Institute of Communication and Media Studies, University of Bern

** Social Computing Group, University of Zurich

March 2023

Report authorship statement

Funding acquisition: Mykola Makhortykh, Silke Adam

Project administration: Mykola Makhortykh, Silke Adam

Research design: Mykola Makhortykh, Silke Adam, Aleksandra Urman, Franziska Keller, Maryna Sydorova

Audit infrastructure design and deployment: Maryna Sydorova, Mykola Makhortykh

Data preprocessing and preparation: Maryna Sydorova, Mykola Makhortykh

Coder training for manual content analysis: Franziska Keller

Data analysis: Mykola Makhortykh, Aleksandra Urman

Data visualisation: Mykola Makhortykh, Maryna Sydorova

Report writing: Mykola Makhortykh, Aleksandra Urman, Maryna Sydorova, Franziska Keller

Report editing: Mykola Makhortykh, Silke Adam, Aleksandra Urman

Funding statement

The project has been funded by the Swiss Federal Office of Communication (OFCOM). Additional research time of Mykola Makhortykh has been funded by the Alfred Landecker Foundation.

Table of contents

Management Summary.....	3
Introduction.....	5
Research Questions.....	8
Related Work.....	9
General State of Algorithm Audit Research.....	9
Algorithm Audit Research for Google Search	10
Algorithm Audit Research for YouTube Search.....	13
Algorithm Audit Research for YouTube Recommendations.....	14
Methodology.....	16
General Description of the Data Collection Approach.....	16
Implementation of Data Collection.....	17
Implementation of Data Analysis.....	24
Limitations of the Study.....	27
Results.....	28
Descriptive Analysis of Google Search Audit Results.....	29
Descriptive Analysis of YouTube Search Audit Results.....	45
Descriptive Analysis of YouTube Recommendation Audit Results.....	61
Regression Analysis – Bringing it All Together.....	69
Conclusions.....	73
Summary of Findings.....	73
Implications and Policy Recommendations.....	76
References.....	78
Appendix A1: Codebook for Manual Content Labelling.....	88

Management Summary

Algorithmic content selection systems, such as Google search or YouTube recommendations, shape today's media ecosystem. By automatically filtering and prioritising information in response to user queries, these systems help individuals to navigate the abundance of available digital content. However, algorithms powering these systems can be subjected to the erratic performance that may result in them facilitating the spread of factually incorrect or misleading information, which can undermine trust in democratic institutions or convince individuals to make decisions harmful to their personal or collective well-being.

To examine the impact of algorithmic systems on exposure to misinformation in Switzerland, the project looked at Google and YouTube, two monopolist platforms in the search and video-hosting markets. It investigated what sources are prioritised by Google and YouTube in relation to COVID and the Holocaust, two topics which are often targeted with misinformation. The project analysed how algorithms' performance is influenced by 1) search query formulation (i.e. whether queries expressing interest in misinformation result in more misinformation exposure); 2) algorithmic personas (i.e. whether visiting websites with specific political leanings before searching leads to more misinformation exposure); and 3) time (i.e. whether exposure to misinformation changes over time).

The project used a virtual agent-based algorithm audit, which is a research technique that simulates human behaviour to generate inputs for the algorithmic system and then record outputs. It consisted of three waves of data collection in May and June 2022. In the case of Google Search and YouTube search, it collected the first 10 search results in response to 28 queries in German: 14 for the Holocaust and 14 for COVID, with half of the queries expressing interest in misinformation (e.g. "corona-impfung mikrochip") and half not expressing such interest (e.g. "corona-impfung"). In the case of the YouTube recommendation system, the project examined a set of recommendations for videos with an explicit interest in Holocaust/COVID misinformation and without it. To analyse the results, the trained team of assistants manually examined the full content of Google Search results and the first three minutes of videos from YouTube Search and recommendation system outputs.

The findings of the project suggest that both Google and YouTube tend to prioritise content coming from journalistic websites and channels in response to COVID- and Holocaust-related queries. Such content also usually comes from media outlets from outside Switzerland (usually Germany), whereas Swiss journalistic outlets appear in top outputs less commonly (with the exception of YouTube recommendation outputs). Similarly, educational sources prioritised by algorithms usually originate outside of Switzerland, unlike government/administrative sources which tend to be Switzerland-based.

Relatively few Google and YouTube outputs support misinformation, in particular in relation to the Holocaust; however, for COVID, such outputs appear more often (up to 40% of outputs for the queries “corona-diktatur” and “corona-lüge” for Google Search). This observation indicates that it is more difficult for algorithms to deal with developing misinformation topics. Misinformation was more present in Google Search compared with YouTube, despite YouTube having a reputation as a platform facilitating exposure to false information through its algorithms. It is important to note, however, that our analysis is more likely to identify misinformation on Google than on YouTube due to YouTube content analysis focusing only on the first three minutes of each video.

In terms of specific factors influencing exposure to misinformation, the project found little influence of algorithmic personas, thus indicating that websites visited before searching are not necessarily taken into consideration by the algorithms consistently. The factors which matter more are the selection of the search queries (with queries with an explicit interest in misinformation resulting in more exposure to it) and the time at which the input was generated, as shown by the variation in the visibility of misinformation-related content.

Together, these findings demonstrate that under certain circumstances (in particular, when users signal interest in it), algorithmic systems can facilitate the exposure of Swiss citizens to false information. They stress the importance of increasing awareness of algorithmic misinformation in the Swiss population, the need to establish monitoring infrastructure to track how algorithms deal with misinformation over time, and fostering normative discussion on what the role of algorithms should be in the Swiss context.

Introduction

Today's digital media ecosystem is shaped by algorithmic content selection systems. Examples of such systems range from search engines, such as Google Search (Google Search, n.d.) or YouTube search (YouTube, n.d.a), which retrieve and rank information in response to user queries, to recommender systems, such as YouTube video recommendations (YouTube, n.d.b) or NZZ Companion App (Leuener, 2017), which provide content suggestions based on the user's earlier behaviour, to conversational agents, such as ChatGPT (Ramponi, 2022), which use algorithms to select information for generating text-based responses to user requests. The growing reliance of users on these systems is attributed to the abundance of digital content – sometimes referred to as 'information overload' (Bawden & Robinson, 2009) – which makes content difficult to navigate and consume due to excessive volume. Under these circumstances, there is a growing need for algorithmic systems capable of automatically filtering and prioritising digital content to make it accessible to humans.

Despite the importance of algorithmic content selection systems for informing individuals on topics ranging from healthcare (Elkin et al., 2020; Ghezzi et al., 2020; Li et al., 2022; Makhortykh et al., 2020) and politics (Juneja et al., 2023; Pradel, 2021; Trielli & Diakopoulos, 2022; Urman et al., 2022a) to gender equality (Kay et al., 2015; Otterbacher et al., 2017; Ulloa et al., 2022b; Urman & Makhortykh, 2022) and collective memory (Makhortykh et al., 2021a, 2022a; Zavadski & Toepfl, 2019), the performance of these systems raises many concerns. A number of studies have demonstrated how these systems can be subjected to non-systematic errors, such as the retrieval of information irrelevant to the user's request (Makhortykh et al., 2021a), or systematic bias which leads to consistent prioritisation of outputs promoting specific (and often distorted) viewpoints, for instance, that women are less capable workers than men (Kay et al., 2015) or nudging system users towards consuming more extreme content (Chen et al., 2021). These forms of malperformance have profound societal implications by limiting the diversity of information diets, amplifying societal stereotypes (Vlasceanu & Amodio, 2022) and potentially contributing to individual or group radicalisation (Ribeiro et al., 2020).

While the degree to which different content selection systems perform wrongly and whether such malperformance might have the above-mentioned societal implications remains debated (e.g. Bruns, 2019; Brown et al., 2022; Möller et al., 2019; Zuiderveen Borgesius et al., 2016), these systems are certainly capable of facilitating the spread of misinformation. Understood as factually incorrect or misleading information which is spread without a clear intent to deceive the audience (Pamment, 2020), misinformation poses multiple challenges for individuals and societies. The spread of misinformation can undermine trust in democratic institutions (Ognyanova et al., 2020) and convince individuals to make decisions which can be harmful to the individual and to collective well-being (e.g. in the case of 'anti-vax' attitudes; Featherstone & Zhang, 2020). Misinformation can also enable possibilities for manipulating public opinion and justifying violence (for instance, in the case of the 2021 'Capitol riots' in the USA; Riley, 2022) and facilitate the radicalisation of society by transforming ordinary matters into the issues of existential security, often by utilising distorted historical narratives for articulating past injustices (e.g. as in the case of Russia's war in Ukraine; Gaufman, 2015; Makhortykh, 2018).

While the possibility of algorithmic content selection systems amplifying the spread of false and/or distorted information on topics ranging from gender (Kay et al., 2015; Noble, 2018; Urman & Makhortykh, 2022) and elections (Chen & Wang, 2022; Juneja et al., 2023; Makhortykh et al., 2022b) to historical atrocities (Guhl & Davey, 2020; Makhortykh et al., 2022a) and disproven conspiracy theories (Hussein et al., 2020; Urman et al., 2022b) has been demonstrated in the contexts of the USA, Russia and Germany, in the case of Switzerland little research has been done on the topic. There is also lack of clarity concerning how content selection systems perform in relation to different types of misinformation (e.g. more established versus emerging forms of factually incorrect claims) and differences in performance between individual systems, especially with some platforms, such as YouTube, relying on multiple selection systems (e.g. an internal platform search system and a content recommendation system) for curating content. Finally, there is still limited understanding of the degree to which the performance of algorithmic content selection systems in the context of misinformation is influenced by specific factors, in particular potential changes in the relevance of content over time and the web history of users (i.e. websites visited before using the system), and

the degree to which these factors affect algorithmic personalisation of content selection (i.e. the customisation of system outputs for individual users).

To address these gaps, we conducted a series of algorithm audits to examine how algorithmic content selection systems used by two platforms – Google and YouTube – have dealt with misinformation in relation to the COVID pandemic and the Holocaust in the Swiss context. Our decision to focus on Google and YouTube is attributable to these platforms being monopolists in the Swiss search and video-hosting markets. As of 2023, Google is responsible for 91% of the search market share in Switzerland (Statscounter, 2023), whereas YouTube in 2022 was used by 86% of the Swiss population (Datareportal, 2022), which puts it ahead of other video hosting services (e.g. TikTok) as well as social media platforms (e.g. Facebook or Twitter). Consequently, we assume that the performance of content selection systems used by Google and YouTube is of importance for Switzerland due to their use by a large part of the Swiss population.

In the course of the audits, we focused on the performance of three content selection systems: Google Search, the YouTube search system as well as the YouTube recommendation system. These systems enable the key functionalities of the two platforms – the retrieval of content in response to user queries (Google and YouTube search algorithms) and the provision of content suggestions based on a user's earlier interactions with the platform (YouTube recommender algorithms). We examined whether the performance of the three systems is influenced by web history constituted by the earlier browsing behaviour of users, namely visits to the websites of Swiss politicians and media outlets with different political leanings, and how the performance changes over time. Additionally, we compared the performance of the content selection systems when these systems receive explicit signals that users might be interested in misinformation content – for instance, by using queries associated with known misinformation claims (e.g. that gas chambers did not exist in Holocaust camps or COVID is harmless) or by starting to watch YouTube content containing misinformation (e.g. Holocaust denialism) – and when such signals are absent.

The rest of the report is organised as follows: first, we introduce the research questions which we aim to answer. Then, we discuss the general state of research on auditing algorithmic content selection systems, followed by a discussion of existing scholarship dealing with auditing Google and

YouTube algorithms, with a particular emphasis on possible exposure to misinformation enabled by these algorithms' (mal)performance. We then introduce the methodology used to conduct the algorithm audits for Google Search, YouTube search, and YouTube recommender and explain our data analysis approaches. Following this, we present our findings concerning the impact of algorithmic personas, user queries, and the factor of time on the composition of sources and misinformation stances prioritised by the algorithmic systems. We end with a short summary of our findings together with a discussion of their implications and policy recommendations.

Research Questions

RQ1: How do the selection of information sources and stances on COVID- and Holocaust-related misinformation vary for Google Search, YouTube search, and YouTube recommendation outputs?

RQ2: What is the impact of web history (modelled via algorithmic personas) on the selection of information sources and stances on COVID- and Holocaust-related misinformation for Google Search, YouTube search, and YouTube recommendation outputs?

RQ3: What is the impact of user queries on the selection of information sources and stances on COVID- and Holocaust-related misinformation for Google Search, YouTube search, and YouTube recommendation outputs?

RQ4: How does the selection of information sources and stances on COVID- and Holocaust-related misinformation change over time for Google Search, YouTube search, and YouTube recommendations?

Related Work

General State of Algorithm Audit Research

The growing reliance of users on algorithmic content selection systems and the profound implications this has for information exposure and agenda-setting in today's media ecosystem has prompted the establishment of a new field of research dealing with algorithm audits. Defined as a research methodology aiming to investigate the functionality and impact of decision-making algorithms

(Mittelstadt, 2016), algorithm auditing has been increasingly applied to study how algorithmic systems of content selection help users navigate online platforms and distribute different types of digital content.

The primary focus of algorithm audits in the context of platforms has been “problematic behavior” (Bandy, 2021) of the audited system and the ways systems deal with “problematic content” (Yesilada & Lewandowsky, 2022). Examples of problematic behaviour in algorithmic content selection systems summarised by Bandy (2021) can range from discrimination (e.g. in the form of treating individuals disparately in respect to their individual characteristics; Mikians et al., 2012; Chen et al., 2016) to exploitation (e.g. in the form of inappropriate use of external content by individual platforms; e.g. Vincent et al., 2019). In our study, however, we specifically focus on another form of problematic behaviour of algorithmic systems which deals with the distortion of reality by promoting false information (for some examples, see Bradshaw, 2019; Makhortykh et al., 2022b; Urman et al., 2022b; Williams & Carley, 2023).

To detect unwanted behaviour by algorithmic systems used by the platforms, several algorithm audit approaches have been introduced. These approaches can be grouped into four broad categories (for a detailed review, see Bandy, 2021): 1) code audits, where researchers study the programming code powering the algorithmic content selection system audited (e.g. Weber & Kosterich, 2018); 2) direct scraping audits, where researchers collect outputs of the algorithmic system either via manual queries or the respective application programming interface (API) (e.g. Kay et al., 2015; Noble, 2018; Paramita et al., 2021); 3) virtual agent audits (also known as sock puppet or carrier puppet audits), where researchers use software simulating human behaviour to generate inputs for the algorithmic system in a controlled environment (e.g. Hussein et al., 2020; Mikians et al., 2012; Urman et al., 2022a); and 4) crowdsourcing audits, where system outputs are collected with the help of crowd-workers (e.g. Courtois et al., 2018; Puschmann, 2019).

These approaches have been applied to examine the performance of the algorithmic content selection systems used by different types of platforms. Two common subjects of algorithm audit studies are web search engines, in particular Google (Kulshrestha et al., 2019; Makhortykh et al., 2021b; Puschmann, 2019; Steiner et al., 2022; Urman et al., 2021, 2022c) and YouTube

recommender systems (Chen et al., 2021; Hussein et al., 2020; Ribeiro et al., 2020). However, recently, more studies have appeared looking at other platforms, such as Twitter (Bandy & Diakopoulos, 2021), Facebook (Kuznetsova & Makhortykh, 2023), and Amazon (Juneja & Mitra, 2021).

Algorithm audit studies have played an important role in detecting problematic behaviour by algorithms. However, the field has had to cope with multiple challenges which impede research on algorithmic content selection systems. Two key challenges in this context concern the highly non-transparent and complex nature of systems used by the platforms (e.g. Galindo & Garcia-Marco, 2017; Leerssen, 2020), where system outputs can be determined by multiple user- and system-side factors, and changes in the functionality of algorithmic systems (e.g. Cornia et al., 2018; Goodrow, 2021), which pose both methodological and regulatory challenges for evaluating and monitoring the performance of these systems over time.

Algorithm Audit Research for Google Search

The functionality of Google Search is based on three processes: 1) collecting information from websites by using web crawlers; 2) analysing this information to index websites in the Google database; and 3) retrieving information from the database and ranking it in response to the user query to guide the search user towards the most relevant results (Google Search Central, n.d.a). The relevance of search results is decided by the Google algorithm and is based on multiple factors, which can be broadly divided into user-based ones, such as the language and the exact formulation of the query entered by the user, and system-based ones, such as what the algorithm treats as the most relevant information sources for a particular query at the given time and location and how the algorithm interprets the earlier history of user searches (Google Search, n.d.). System-based factors might also include the larger web history of the user (i.e. website visits outside ones made through Google Search), albeit there is no decisive confirmation of it affecting the ranking of search results.

The degree to which individual factors influence the composition of Google Search results and enable search personalisation for individual users has been long debated in academic scholarship (e.g. Hannak et al., 2013; Kliman-Silver et al., 2015; Norocel & Lewandowski, 2023).

The debate has been complicated by frequent changes in Google Search functionality, which has been subjected to major changes over its history (for examples, see Google Search Central, n.d.b) and which have affected both its overall functionality and the treatment of specific subjects (for instance, information about the COVID pandemic; Google Search Central, n.d.c).

Despite these complications, researchers have identified several factors which have a substantial impact on the composition of Google Search results. Makhortykh et al. (2021a, 2022a) observed substantial differences in the thematic composition of content retrieved by Google in Cyrillic and Latin scripts for information on the Holocaust and the Holodomor in Germany. Hannak et al. (2013) identified the importance of the factor of time for the composition of Google Search results, while Kliman-Silver et al. (2015) established the effects that user location – inferred through their IP address – have on the composition of results they receive. Finally, Puschmann (2018) found that being logged in a Google account has little implications for personalisation of search outputs.

Another factor which influences the composition of Google Search results is serendipity. Several studies (e.g. Haim et al., 2017; Makhortykh et al., 2020) have demonstrated that Google can randomise its outputs for users conducting searches at the same time and under the same conditions (e.g. location or operating system type). Such randomisation can have substantial implications for the quality of information outputs, as in the case of the presence/absence of suicide prevention hotline numbers, which were provided in response to some suicide-related search queries but not others (Haim et al., 2017).

A number of studies have also looked at how the above-mentioned factors can influence the distribution of false information via Google Search. Noble (2018) highlighted how Google spreads false information about specific gender/racial groups in U.S. content. In a study examining search outputs in Sweden and Germany in relation to migration, Norocel and Lewandowski (2023) demonstrated how the selection of more extreme queries (e.g. “rapefugees”) led to exposure to misinformation, as Google prioritised information coming from extreme-right fringe websites (e.g. pi-news.net) for such queries. Finally, Bradshaw (2019) highlighted a number of instances when Google prioritised “junk news” content (i.e. content promoting conspiracy theories and hyperpartisan narratives) in its search outputs.

Despite these shortcomings, Google is known to put substantial effort into countering the potential spread of misinformation via its algorithm that often pays off, in particular when compared with other search engines. Urman et al. (2022b), in a study of search outputs dealing with conspiracy theories, found that among the four search engines examined, Google included the least amount of content promoting misinformation. Trielli and Diakopoulous (2022) and Unkel and Haim (2021) also observed the tendency of Google to prioritise mainstream information sources over more niche/partisan ones, in particular in the context of politics. Such “mainstreaming effect” (Trielli & Diakopoulous, 2022) can neutralise some user-side factors (e.g. the choice of search queries with distinct political leaning; Tripodi, 2022) which otherwise could result in retrieval of more niche or false information.

Algorithm Audit Research for YouTube Search

For YouTube, there are two content selection systems which help users navigate the platform’s content: an internal search system and a recommender system. The internal search is essential for users to be able to find relevant materials in the massive volume of content hosted by the platform. Compared with Google Search, YouTube search is easier to implement due to its not requiring the collection and analysis of information from external websites and instead relying on information existing within a single platform. According to YouTube (n.d.a), the search takes into consideration three main components: relevance (i.e. how well information about the video corresponds to the user query), engagement (i.e. how much other users engage with the video) and quality (i.e. whether the video comes from a trustworthy and authoritative channel). Additionally, YouTube search can take into consideration individual search and watch history to personalise search suggestions.

Research has been primarily focused on the YouTube recommender system, whereas the search system has remained quite under-investigated (for a few exceptions, see Hussein et al., 2020; Juneja et al., 2023). The focus in the case of the YouTube recommendation system has been on its potential for exposing individuals to misinformation, in particular in the context of ‘alt-right’ conspiracy theories. Specifically, there are concerns about YouTube recommendations performing as a “filter bubble” (i.e. by capturing the user in the information bubble based on one’s interests or ideological

preferences; Bryant, 2020) or a “rabbit hole” (i.e. by nudging the users away from rich information environment and towards an extreme echo chamber; Brown et al., 2022). Both of these phenomena can amplify the individual’s exposure to misinformation and facilitate one’s radicalisation. Consequently, concerns about them have contributed to YouTube acquiring the reputation of “great radicaliser” (Tufekci, 2018).

There are several reasons why YouTube search has caused substantially less concerns than YouTube recommendations. Unlike the YouTube recommendation algorithm, which is a notoriously complex system and makes decisions in a rather obscure way. YouTube search takes explicit input from users in the form of queries and retrieves content which the user is explicitly interested in. In addition to being more transparent, the degree of personalisation of YouTube search has also been debated, in particular as earlier research (e.g. Hussein et al., 2020; Juneja et al., 2023) has demonstrated that some factors expected to influence YouTube search performance do not necessarily have a substantial impact, thus resulting in limited personalisation of YouTube search. Specifically, Hussein et al. (2020) showed that neither the geographic location nor demographic data for user accounts influenced exposure to misinformation when searching for content related to conspiracy theories on YouTube.

Despite limited personalisation of its outputs, YouTube search still raises concerns about the possibility of it facilitating exposure of its users to misinformation. Li et al. (2022) examined the performance of YouTube search in the context of information about COVID vaccines and found that approximately 11% of the top 150 search outputs contained non-factual information. Similarly, Donzelli et al. (2018), in a study on Italophone content discussing the relationship between vaccination and autism, observed the tendency of YouTube search to prioritise information promoting anti-vaccination claims, which appeared three times more often than pro-vaccination ones.

By contrast, Hussein et al. (2020) found that while for some conspiratorial queries (e.g. chemtrails) YouTube search retrieves more information supporting misinformation, for many of these queries it prioritises items debunking misinformation. In a recent study, Juneja et al. (2023) found

that in relation to 2020 US election misinformation, YouTube search tends to prioritise content debunking misinformation.

Algorithm Audit Research for YouTube Recommendations

The core functionality of the YouTube recommender system is argued to be about helping users “find the videos they want to watch and that will give them value” (Goodrow, 2021). Unlike YouTube search, which retrieves information in response to explicit queries from the user, YouTube recommendations rely on implicit signals to suggest content users might be interested in. With this aim, the system takes into consideration different signals coming from users (e.g. clicking, watching, and responses to user surveys; Goodrow, 2021).

In contrast to Google Search and YouTube search, where the degree of algorithmic personalisation remains debated, with the YouTube recommender system there is consensus about its enabling strong customisation of outputs for individual users. However, the role of individual factors is often unclear, in particular considering the changes made to the system over the years. An important part of these changes has concerned the alleged tendency of YouTube to increase exposure to extreme or false content, which was mentioned in the previous section (for the in-depth review of studies on the subject, see Yesilada & Lewandowsky, 2022).

The key factor influencing the selection of YouTube recommendations has been shown to be user interactions with YouTube content, as demonstrated by a number of studies (e.g. Hamidy, 2022; Srba et al., 2023). Specifically, the type of video watched by the user previously plays a prominent role in follow-up recommendations (Makrmann & Grimme, 2021). The location of the user has been argued to have a lesser effect than content interactions (Hamidy, 2022). Additionally, while the number of subscriptions to a channel now plays a less significant role than during the earlier days of the recommender system, this still matters in terms of recommendation selection (Makrmann & Grimme, 2021).

Another factor which influences YouTube recommendations is serendipity. Similar to Google Search, where content selection can be randomised, the YouTube recommender system aims to

add more serendipity and diversity, in particular the further the user moves from the point of entry into the system (e.g. the initial seed video; Kaiser et al., 2021)

A number of studies (e.g. Ribeiro et al., 2020; Tufekci, 2018) have argued that YouTube recommendations amplify extreme views and facilitate the spread of misinformation. Tufekci (2018) argued that the YouTube recommender has a tendency to generate recommendations “in a manner that appears to constantly up the stakes” and that potentially increases the likelihood of it exposing users to extreme or false claims, thus lending YouTube the reputation of a “great radicaliser”. Ribeiro et al. (2020) used simulations of random walks for YouTube channels and videos in the context of ‘alt-right’ content (which often tends to promote fringe or false statements) and found that the algorithm potentially prioritises such content. Juneja et al. (2023) used crowdsourcing to highlight that engagement with videos supporting misinformation increases the likelihood of YouTube recommending content supporting misinformation to the users.

At the same time, some studies (e.g. Hosseinmardi et al., 2021; Kaiser et al., 2021; Ledwich & Zaitsev, 2020) have argued that the above-mentioned issues with the YouTube recommendation algorithm are not as pronounced. Hosseinmardi et al. (2021) used digital trace data to examine YouTube content consumption in the U.S. and found no direct evidence of YouTube recommendations systematically amplifying engagement with far-right content. Based on a combination of YouTube API data and web scraping, Ledwich & Zaitsev (2020) suggest that YouTube recommendations actually discourage users from engaging with radicalising content. Kaiser et al. (2021) also used web scraping data to collect YouTube recommendations and did not identify misinformation filter bubbles in the context of content dealing with Zika virus, despite observing misinformation-related recommendations.

Methodology

General Description of the Data Collection Approach

To collect data for the study, we conducted a series of virtual agent-based algorithm audits (Ulloa et al., 2022). This type of algorithm audit relies on software (i.e. virtual agents) that simulates user browsing behaviour (e.g. opening web pages or entering search queries) to generate inputs for the

algorithmic system and then record its outputs (e.g. in the form of an HTML page with search results). The benefit of virtual agent-based audits is that this approach allows controlling not only for different factors influencing system personalisation (e.g. customisation of web search outputs for individual users; Hannak et al., 2013), but also the potential randomisation of system outputs (e.g. variation of web search outputs due to ongoing A/B testing aiming to identify the optimal output ranking; Makhortykh et al., 2020).

To control for specific personalisation factors, virtual agents are usually synchronised (i.e. to control for the effect of the time at which system inputs, in the form of agent actions, are generated) and deployed in a controlled environment (e.g. a network of virtual machines using the same IP range, the same type of operating system (OS) and the same browsing software with no earlier history of interactions). Both time and user environment specifications can potentially affect system outputs (e.g. the selection of YouTube recommendations) and are rather difficult to control in auditing approaches relying on system inputs generated by human users (e.g. due to potential differences in browsing software and differences in browsing history). Consequently, virtual agent-based auditing is particularly useful for isolating the effects of both user- (e.g. the way in which search queries are formulated) and system-driven personalisation (i.e. the impact of web history on the outputs of the algorithmic content selection system). At the same time, the possibility of simultaneously deploying a large number of agents generating the same system inputs allows controlling for randomisation that might otherwise make the audit results less consistent (e.g. by making it difficult to determine whether the differences in system outputs are due to effects of personalisation or just randomisation).

Implementation of Data Collection

Data collection infrastructure and implementation of algorithmic personas. To conduct virtual agent-based algorithm audits and collect data for the study, we built a cloud-based research infrastructure using Google Compute Engine, an online service provided by Google. We used this infrastructure to deploy 24 virtual machines hosting two virtual agents each (i.e. 48 agents altogether) and store content collected by the agents (i.e. HTMLs of web pages containing Google and YouTube search results together with YouTube pages for each recommended video visited by the agents in

the course of the YouTube recommender audit). Each agent was deployed in either Firefox or Chrome web browser using Selenium (a set of software tools utilised for browser automation and commonly used for virtual agent-based algorithm audits) to simulate human browsing activities.

The virtual machines were made from scratch with the same hardware and software specifications for each type of audit to minimise differences between agents which could potentially affect the outputs of the audited systems. For Google Search audits, we used e2-standard-2 machines with two central processing units (CPUs) and 8 GB of random-access memory (RAM). For YouTube search and recommendation audits, we had to use virtual machines with more resources due to higher RAM requirements for opening YouTube pages in the browsers; hence, we utilised e2-standard-4 machines with four CPUs and 16 GB RAM. All machines used the Debian distribution of Linux as an OS.

Because of our particular interest in the effects of algorithmic personalisation, we created five algorithmic personas following the research design introduced by Haim et al. (2018). Algorithmic personas are user profiles modelled via sequences of actions (for instance, visits to specific websites) which precede the generation of system inputs (e.g. the search activity) and can influence the system outputs (e.g. search results; Haim et al, 2018; Mikians et al., 2012). We were particularly interested in the impact of politics-related web history on the Google and YouTube algorithmic systems, so we modelled personas based on visits to the websites of media organisations and political actors with different political leanings in Switzerland.

To investigate such an impact, we made agents visit a selection of websites before conducting each search action or initiating the recommendation chain routine (see below). To uncover the effects of web histories of different political leanings and following the assumption that algorithmic content selection systems might use the political leaning of users as a signal for selecting the outputs, we created five personas labelled “right”, “centre-right”, “centre-left”, “left”, and “control”. Each persona (except the control one) visited five websites associated with the respective political leaning: one of a media outlet, one of a political party and personal websites of three Swiss National Council members. In the case of media outlets and parties, the decision regarding political leaning was made by two experts in the Swiss political and media landscape, whereas in the case of Swiss

National Council members the decisions were made based on individual Swiss politicians' political scores (Hermann & Krähenbühl, 2020). The selection of the websites per persona is summarised in Table 1.

Table 1. *Composition of Algorithmic Personas for Audits*

Persona	Media outlet	Political party	Swiss National Council members
Left	WOZ (https://www.woz.ch/)	JUSO (https://juso.ch/de/)	Christian Dandrès (https://www.christiandandres.ch), Léonore Porchet (https://leonoreporchet.ch/intro/), Ada Marra (https://adamarra.blogspot.com/)
Right	Weltwoche (https://weltwoche.ch/)	SVP (https://www.svp.ch/partei/)	Werner Salzmänn (https://www.werner-salzmänn.ch/), Pirmin Schwander (https://pirmin-schwander.ch/), Hansjörg Knecht (https://www.hansjoerg-knecht.ch/)
Centre-left	Tagesanzeiger (https://www.tagesanzeiger.ch/)	SP (https://www.sp-ps.ch/de)	Elisabeth Baume-Schneider (https://elisabeth-baume-schneider.ch/), Marina Carobbio Guscetti (https://www.marinacarobbio.ch/), Paul Rechsteiner (https://paulrechsteiner.ch/)
Centre-right	NZZ (https://www.nzz.ch/)	FDP (https://www.fdp.ch/willkommen)	Hans Wicki (https://www.wickihans.ch/), Thomas Hefti (https://thomas-hefti.ch/), Damian Müller (https://www.damian-mueller.ch/)
Control	None	None	None

The 48 agents were divided in the following way between the five personas discussed earlier: 10 agents were assigned to each of the right, left, centre-left and centre-right personas, and 8 agents were assigned to the control persona. The control person was intended to be used as a baseline against which the impact of other algorithmic personas can be measured.

Google Search and YouTube search audits. To audit Google Search and YouTube search, we used a selection of 28 search queries related to the Holocaust and COVID as both topics are common targets of misinformation (for COVID, see Brennen et al., 2020; Cuan-Baltazar et al., 2020; Gabarron et al., 2021; for the Holocaust, see Allington, 2017; González-Aguilar & Makhortykh, 2022; Guhl, & Davey, 2020). Furthermore, the comparison of system performance in relation to COVID and the Holocaust allowed us to examine whether algorithmic systems are more capable of dealing with more established forms of misinformation (e.g. Holocaust denialism) as contrasted to more recent/emerging forms (e.g. COVID misinformation).

To select the queries, we used the following procedure: first, we identified common forms of misinformation in relation to the COVID-19 pandemic and the Holocaust by using authoritative resources. For the Holocaust, we relied on the International Holocaust Remembrance Alliance working paper (IHRA, 2021) discussing different forms of Holocaust distortion. For COVID, we used the Alliance for Science article (Lynas, 2020) examining different forms of COVID-related conspiracy theories.

Then, we manually produced a set of queries which were discussed by the authors of the report to shortlist the options which were viewed as particularly fitting for the study. While producing the list of queries, we aimed to collect queries directly focusing on the specific misinformation narrative (e.g. “corona vaccination microchips”) and queries dealing with a similar topic, but without an explicit focus on misinformation (e.g. “corona vaccination”).

The shortlisted options were tested in different formulations by using Google Trends to identify the formulations which attracted more attention from users (e.g. it was one of reasons we used queries with the word “corona” and not “COVID”, because for German language queries with “corona” were more often used). We assumed that the more frequently used queries would be worth focusing on due to their potentially larger impact on Swiss society. The final list of queries consisted

of 28 queries (presented in Table 2; first we present German versions of queries which were used for data collection, followed by English translations in square brackets): 14 for the Holocaust and 14 for COVID. For each topic, 7 queries exhibited interest in misinformation and 7 did not.

Table 2. *Queries Used for Google and YouTube Search Audits*

COVID queries [focused on misinfo]	COVID queries [not focused on misinfo]	Holocaust queries [focused on misinfo]	Holocaust queries [not focused on misinfo]
5g corona [5g corona]	corona [corona]	holocaustopfer übertrieben [exaggerated Holocaust victims]	6 millionen opfer [6 million victims]
corona biowaffe [corona bioweapon]	woher kommt das coronavirus [from where does the coronavirus originate]	holocaust deutsche nicht wissen [Germans did not know about the Holocaust]	holocaust erinnerung [holocaust memory]
corona labor unfall [corona lab accident]	mensch tier corona [human animal corona]	juden profitieren holocaust [Jews benefit from the Holocaust]	juden holocaust opfer [jews Holocaust victims]
corona-impfung mikrochip [corona vaccination microchip]	corona-impfung [corona vaccination]	hitler hatte recht [Hitler was right]	nie wieder [never again]
corona harmlos [corona harmless]	long covid [long covid]	gaskammer-lüge [gas chamber lie]	konzentrationslager [concentration camp]
corona-diktatur [corona dictatorship]	corona-fallzahlen [corona case numbers]	holocaust-lüge [Holocaust lie]	holocaust [Holocaust]
corona-lüge [corona lie]	corona-massnahmen [corona measures]	bomben-holocaust [Holocaust by bombing]	judenverfolgung [persecution of the Jews]

For the Google and YouTube audits, our agents were programmed to open the main page of Google Search (i.e. google.com) and YouTube (i.e. youtube.com) and then enter search queries one by one in the search field. Before each query, each agent conducted the corresponding algorithmic

persona routine (see above); after each query, the agent saved the HTML page with the search results and then closed the browser. The searches were conducted in anonymous mode to prevent the history of browsing being recorded and potentially affecting subsequent system outputs.

YouTube recommendation audit. In the case of YouTube recommendation audits, we relied on a different design. We selected four YouTube videos: 1) “Der Volkslehrer beim Prozess gegen Alfred Schaefer” (<https://www.youtube.com/watch?v=xx5H1ZYXiFo>); 2) “MASS-VOLL! - Wir kämpfen für unser Recht zu leben!” (<https://www.youtube.com/watch?v=axy5ZgtP3JI>); 3) “Die bizarre Welt der Holocaustleugner | SPIEGEL TV” (https://www.youtube.com/watch?v=UJkBvn8j_WY); and 4) “Wie radikale Schweizer gegen Corona-Regeln rebellieren auslandsjournal” (<https://www.youtube.com/watch?v=9GwWkXHeVW8>). The first two videos focused on misinformation about the Holocaust and COVID without explicitly debunking it, respectively, whereas the latter two focused on the debunking of misinformation about the topics. The videos were selected by using the queries used for the Google/YouTube search audits and then examining the results to find videos mentioning misinformation and debunking misinformation on similar issues. This design was informed by our interest in whether the YouTube recommendation algorithm still follows the principle of the “rabbit hole” (e.g. Brown et al., 2022) – i.e., a tendency to amplify user exposure to misinformation-related content under the condition that the user starts watching such content on the platform.

For the audit, we programmed the agents to conduct the usual persona-modelling routine (i.e. visiting five websites). After it, they opened one of the four seed videos, and watched it for 45 seconds and then shifted to the top recommended video. The procedure (i.e. 45 seconds of watching followed by the shift to the top recommendation) was repeated 50 times, so each agent made 50 recommendation steps from the initial video. The process was repeated for each of the four seed videos.

To examine the effect of time on the performance of the algorithmic content selection systems, we conducted six waves of audits for each of the systems (i.e. Google Search, YouTube search and YouTube recommendations). We were particularly interested in the degree of variation in system outputs across a short period (e.g. several hours) and more lengthy periods (e.g. several

weeks). We therefore planned two waves per day with one wave taking place in the morning (e.g. wave 1) and another one in the evening (e.g. wave 2). The dates of the data collection waves are provided in Table 3.

Table 3. *Data Collection Waves for Algorithmic Audits*

	Google search	YouTube search	YouTube recommendation
1st wave	15 May [morning]	16 May [morning]	16 May [morning]
2nd wave	15 May [evening]	16 May [evening]	16 May [evening]
3rd wave	4 June [morning]	5 June [morning]	5 June [morning]
4th wave	4 June [evening]	5 June [evening]	5 June [evening]
5th wave	25 June [morning]	26 June [morning]	26 June [morning]
6th wave	25 June [evening]	26 June [evening]	26 June [evening]

We collected 76,251 top 10 organic Google Search results (the uneven number is due to occasional variation in the number of organic search results appearing on the page). For YouTube, we collected 76,100 top 10 organic search results. Finally, for YouTube recommendation chains, we collected 57,035 recommendation pages visited by the agents.

Implementation of Data Analysis

We extracted all unique URLs in the top 10 organic Google and YouTube search results for all six waves. We thus gathered 436 unique URLs for Google Search and 660 unique URLs for YouTube search. For the YouTube recommendation system, we extracted all unique URLs for recommendation steps 1–20 for waves 2, 3 and 5 due to the substantially larger volume of unique outputs collected (i.e. 23,219 as contrasted by a few hundreds for Google Search and YouTube search). The final sample for YouTube recommendations consisted of 4,939 URLs.

To investigate the outputs of algorithmic content selection systems used by Google and YouTube, we designed a codebook consisting of the six variables described below (for more

information about the variables, see the codebook in the Appendix A1). The selection of options for individual variables (e.g. COVID-related MISINFORMATION) followed the deductive principle.

- (1) ACCESS: whether the system output (i.e. its URL) can be accessed;
- (2) RELATED: whether the output contains information about COVID, the Holocaust or neither;
- (3) SOURCE: the type of the information source to which the URL leads. Possible options include reference websites (e.g. online encyclopaedias such as Wikipedia), journalistic media (e.g. NZZ), academic or educational websites (e.g. University of Bern website), social media (e.g. Facebook; YouTube content was not included in this category - instead, individual channels were examined and labelled using one of other SOURCE categories), HAC (hyperpartisan, alternative or conspiratorial) media, healthcare-related websites, government or administration websites (e.g. Bundesrat website), commercial websites (e.g. Amazon), other websites (e.g. individual blogs or personal YouTube channels);
- (4) COVID-RELATED MISINFORMATION: whether system outputs contain one or more of the following claims: COVID is caused by 5G networks, COVID is a bioweapon, COVID was created in a lab, Big Pharma or another group of economic elites are conspiring to use COVID for economic gain, the severity of COVID is lower than claimed, the government is using the pandemic to install a dictatorship, COVID is a hoax, COVID vaccines are used to insert microchips, and other types of COVID-related misinformation (this category included misinformation narratives other than the ones listed before);
- (5) HOLOCAUST-RELATED MISINFORMATION: whether system outputs contain one or more of the following claims: the number of Holocaust victims is inflated, Germans/Hitler/Nazi were unaware of the Holocaust, Jews benefit from the Holocaust, Hitler/Nazi did more good than bad, gas chambers did not exist, the Holocaust is a lie or a hoax, the Holocaust was justified by the existential threat from Jews to Germany/Christian Europe, the Holocaust in its scale/severity is similar to other historical events (e.g. the bombing of Dresden by the Allies) or present phenomena (e.g. abortion), and other types of Holocaust-related

misinformation (this category included misinformation narratives other than the ones listed before); and

- STANCE (6): whether the output promotes misinformation, shares misinformation without debunking it, explicitly debunks the misinformation, does not mention misinformation, or refers to misinformation in a satirical manner.

After compiling the codebook, we trained seven students to use it for labelling Google and YouTube search and YouTube recommendation outputs. The students attended a 4-hour training during which we discussed the codebook and example websites and videos. Then, the students independently labelled additional examples and training material, for which they received feedback. Finally, they had to pass a test based on two gold standards (i.e. the sets of URLs consensus-coded by at least two of the researchers conducting the study). The first gold standard consisted of 50 Google Search results and 50 YouTube videos and was used for evaluating the performance of the first four students trained. The second gold standard consisted of 41 YouTube videos and was used for evaluating the performance of three additional students hired specifically to label YouTube search and recommendation content.

Of the seven students, six achieved scores higher than .8 (for the intercoder agreement percentage) and .79 (for Brennan–Prediger’s kappa) for all variables when compared with the gold standard. Between individual students, the intercoder reliability scores were above .74 both for the intercoder agreement percentage and Brennan–Prediger’s kappa. These six students were then tasked with labelling the collected search and recommendation outputs.

The results of content labelling by the students were used for producing descriptive analyses of outputs of the algorithmic content selection systems used by Google and YouTube. We also conducted a multinomial logistic regression analysis for Google Search and YouTube search outputs to statistically test the relationship between specific factors (see below) and the degree to which individual agents were exposed to content with different stances on misinformation. The regression analysis was not conducted for the outputs of the YouTube recommendation algorithm due to the highly unbalanced composition of these outputs in terms of stance on misinformation, which made it

unfeasible to run a regression analysis (for more information, see the corresponding Findings section below)

The reference category for the regression analysis was “misinformation debunked”, and the results for the other categories – i.e. “misinformation simply mentioned”, “misinformation promoted”, and “no misinformation mentioned” – are presented in relation to it. The model included predictors for query type, browser type, persona, and wave. The query type predictor had four levels: “COVID [not focused on misinfo]”, “Holocaust [focused on misinfo]”, “Holocaust [not focused on misinfo]”, and reference level “COVID [focused on misinfo]”. The browser predictor had two levels: “Firefox” and “Chrome” (reference level). The persona predictor had five levels: “right”, “left”, “centre-right”, “control” and “centre-left” (reference level). The wave predictor had six levels: “Wave 1” (reference level), “Wave 2”, “Wave 3”, “Wave 4”, “Wave 5”, and “Wave 6”.

The coefficients for each predictor represent the change in the log-odds of the outcome compared to the reference level. For example, the coefficient of 1.86 for “COVID [not focused on misinfo]” in the “misinformation simply mentioned” row indicates that the log-odds of a “misinformation simply mentioned” observation are 1.86 units higher compared to the reference level “misinformation debunked” for the “COVID [not focused on misinfo]” query type.

Limitations of the Study

Before presenting our findings, it is important to discuss several limitations and their possible implications for the study’s results. The first concerns the approach chosen for data collection. Despite the numerous benefits of virtual agent-based algorithm audits, which were outlined earlier, this approach is based on (often simplified) modelling of user behaviour by researchers. In particular, the inputs to the algorithmic systems (e.g. search queries in the case of Google and YouTube search) as well as the context in which they are produced might differ from those of actual users. Consequently, there is also a possibility that system outputs generated via this approach might also be different from the ones obtained by the actual users. However, based on existing research, we do not have reasons to expect the outputs to be completely different for the human and virtual agents.

A related limitation concerns the way in which search personas are modelled. For the current study, we relied on a set of rather simple short sequences of browsing actions for modelling individual personas. Furthermore, these actions were performed outside the platforms where respective algorithmic systems (e.g. Google Search or YouTube search) are deployed, which further decreases their potential for generating signals which can be used for personalisation purposes. While the generated observations are still valuable for understanding the effects of algorithmic personalisation, future research will benefit from testing the impact of more complicated sequences of browsing actions as well actions taking place directly within the audited platforms.

One more limitation relates to the analysis of YouTube content collected in the course of the audits. Due to the substantial length (e.g. 1 hour or longer) of many of the videos retrieved in the course of the audits, for their manual analysis we had to focus on the first 3 minutes of each video to make it feasible with the resources available. While it is reasonable to assume that the beginning of a video is the part likely to be viewed by the most users, our analysis of YouTube content does not capture potential misinformation claims present beyond the first 3 minutes. Consequently, the amount of misinformation identified in the case of YouTube search and recommendation algorithm audits might be deflated compared with that for Google. One potential approach for future research is the use of automated content analysis; however, this might be difficult to implement considering that the automated detection of misinformation remains a notoriously difficult computation task.

The final limitation concerns our focus on the possibility of exposure to misinformation via algorithmic content selection system outputs, but not necessarily the impact of this exposure on how individual citizens in Switzerland perceive misinformation about the Holocaust and COVID. The effects of exposure to information via such systems (e.g. search engines) have been suggested to be rather substantial - for instance, in the context of elections, ranking of search outputs has been argued to be able to shift preferences of undecided voters by 20% (Epstein & Robertson, 2015). While the scope of this effect was criticised for being exaggerated (see, for instance, Zweig, 2017), it is still reasonable to expect that exposure to information via algorithmic systems have implications for the public sphere (e.g. Lev-On, 2008; Helberger, 2019). However, investigating the exact

implications of exposure to misinformation for Swiss citizens via algorithmic content selection systems remains a crucial task for future research.

Results

In this section, we present the results of the audits conducted for the study. It is structured as follows: first, we present descriptive analysis results for each of the three algorithmic content selection system audits (i.e. Google Search, YouTube search, and YouTube recommendation systems). To address our RQs, in each subsection we first discuss how prioritisation of information sources for individual systems is influenced by algorithmic personas (i.e. system-side factors), user queries (i.e. user-side factors) and whether the variation changes over time. Then, the descriptive analysis is repeated for prioritisation of specific stances on misinformation (e.g. whether system outputs promote, mention or debunk misinformation). The descriptive analysis results are followed by the results of the regression analysis, where we statistically test the significance of different factors for the likelihood of users of different content selection systems being exposed to information promoting, debunking or mentioning misinformation in relation to COVID and the Holocaust.

Descriptive Analysis of Google Search Audit Results

Prioritisation of Information Sources by Google Search

General Composition of Search Outputs. Figure 1 highlights the tendency of Google Search to prioritise journalistic sources, which constituted 43% of outputs for COVID queries and 30% of outputs for Holocaust queries. These outputs were primarily coming from the German media companies and outlets, such as Südwestrundfunk (the most common source of journalistic outputs), Mitteldeutscher Rundfunk, Das Erste, and ZDF. The only two Swiss media outlets among the 10 most commonly prioritised journalistic sources (altogether these 10 outlets constituted 52% of all Google Search results coming from journalistic media) were NZZ and Swissinfo.

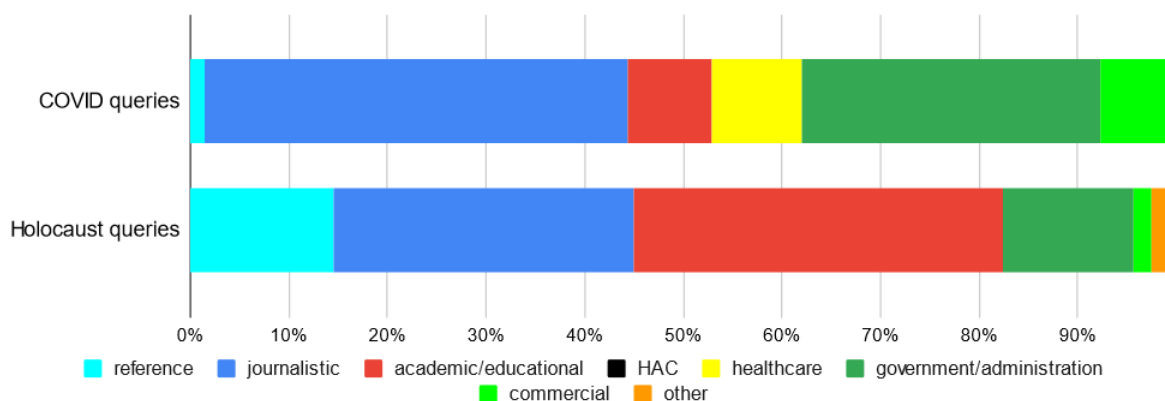


Figure 1. *Aggregated distribution of information sources for all agents across all six waves for COVID and Holocaust queries for Google Search*

In the case of Holocaust queries, however, the most common (i.e. 37%) type of outputs was related to academic/educational sources, which were relatively rare for COVID queries (i.e. 8%). These sources were primarily international (e.g. websites of Anne Frank House or United States Holocaust Memorial Museum) or Germany-based (e.g. websites of Deutsches Historisches Museum or Planet Schule); among the 10 most commonly prioritised sources in this category, we found no Switzerland-based institutions. By contrast, for COVID queries Google prioritised government/administration websites, in particular websites related to the Bundesamt für Gesundheit (e.g. bag-coronavirus.ch, www.covid19.admin.ch) and websites of individual canton administrations (e.g. of Bern and Zurich). In this case, most commonly prioritised sources were all coming from Switzerland with a single exception of the website of Robert Koch-Institut.

In addition to this variation in the top categories of sources, we found a number of differences between Holocaust and COVID queries concerning less common source types. Unlike YouTube search and recommendation outputs, where we observed relatively little diversity in terms of source composition, Google Search results also included content coming from reference websites (in particular Wikipedia and Statista) and commercial websites (in particular Amazon and Orell Füssli). In the case of COVID queries, 9% of search outputs came from healthcare websites, such as Helios Gesundheit and München Klinik (both Germany) or Luzerner Höhenklinik Montana (Switzerland).

HAC media (e.g. website of NWA Schweiz) and ‘other’ websites (e.g. website of Haus kirchlicher Dienste) appeared in Google search results very rarely (primarily, in response to Holocaust queries).

Impact of Algorithmic Personas. Figures 2–3 show that the impact of algorithmic personas on Google Search outputs is negligible. While there are a few instances where different personas received different outputs (for instance right, left and centre-left personas encountering more outputs from journalistic websites for individual waves for COVID queries and the centre-right persona receiving more such outputs for Holocaust queries), the variation remained rather inconsistent.

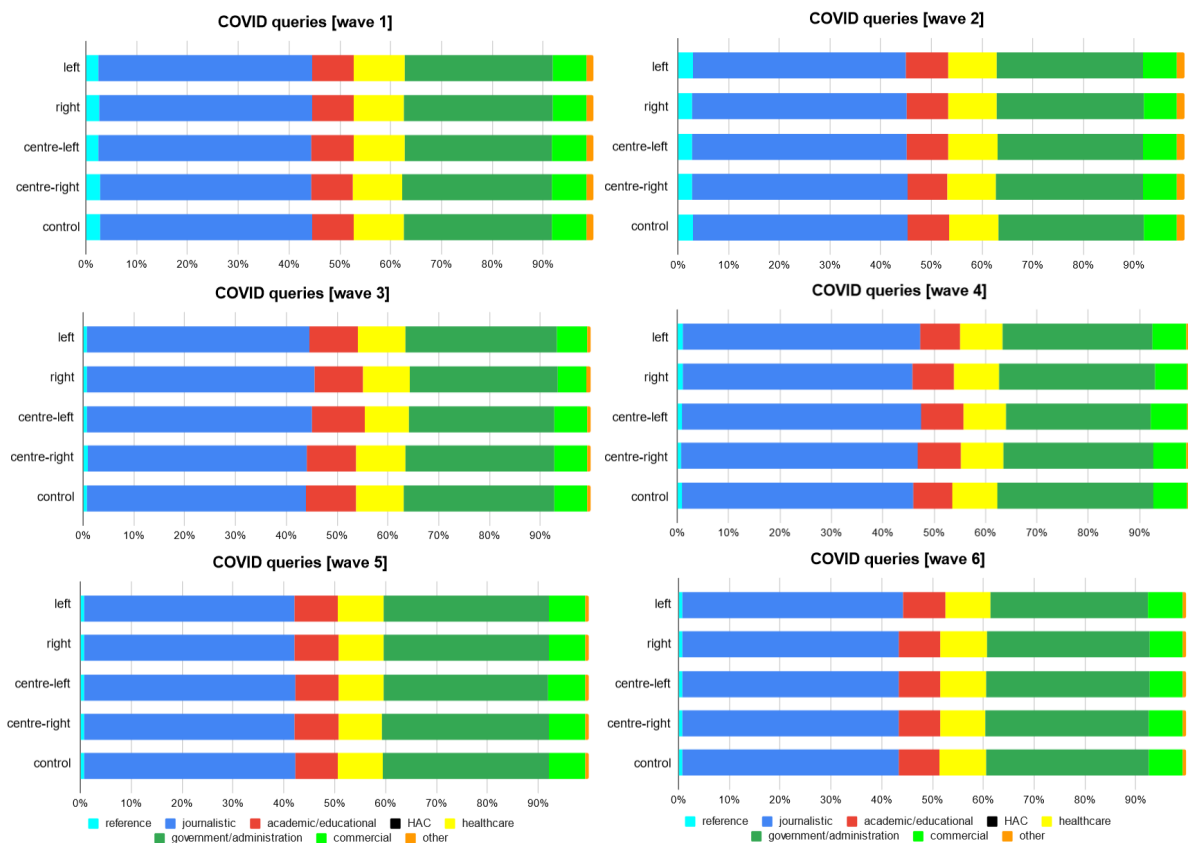


Figure 2. *Distribution of website types per algorithmic persona for Google search results [COVID queries]*

Additionally, only one persona – the centre-left one – was exposed to content from HAC media via Google Search results (albeit just for one wave for the Holocaust queries). In the majority of cases, the distribution of sources between different personas remained the same, suggesting that observed differences might be attributable not necessarily to the search algorithm recognising the implications of different web histories but instead to the randomisation of search outputs.

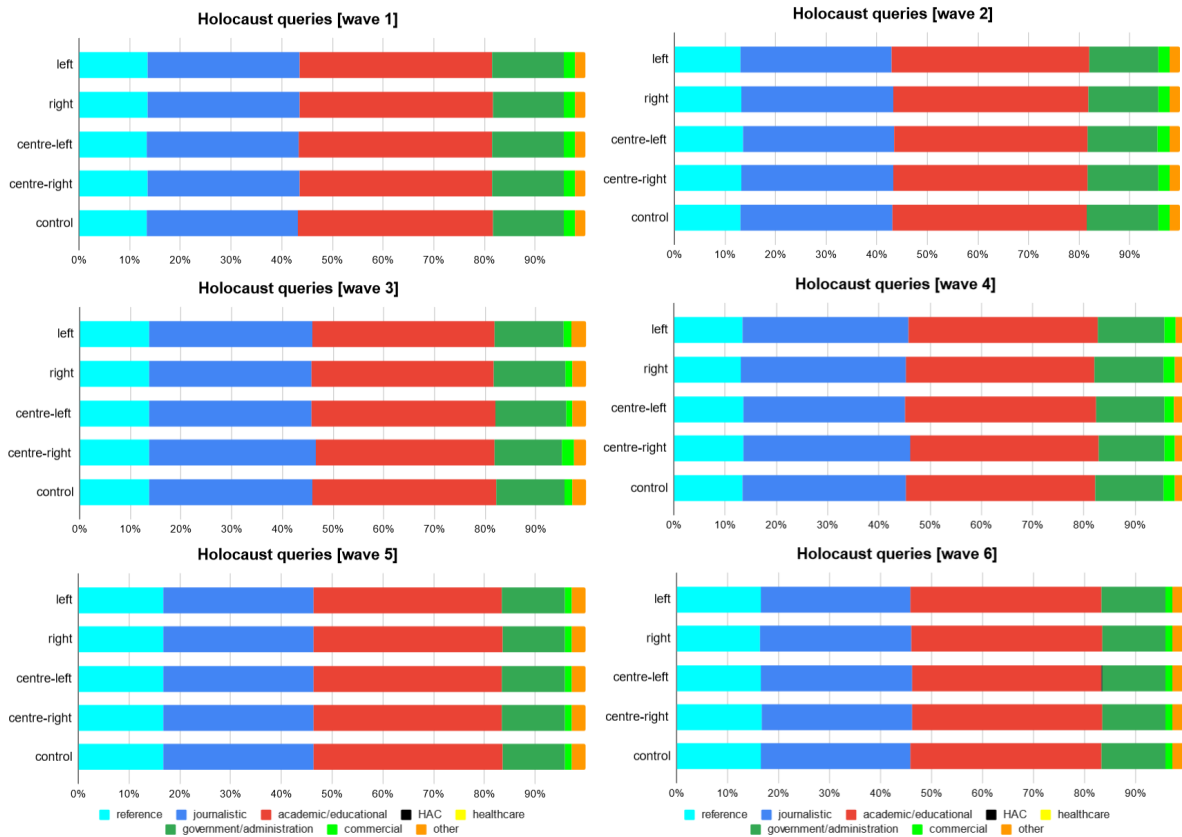


Figure 3. *Distribution of website types per algorithmic persona for Google search results*
[Holocaust queries]

Impact of Search Queries. In contrast to algorithmic personas, the examination of sources prioritised in response to individual queries (Figures 4–7) highlights substantial differences, in particular between queries focused on misinformation and those not focused on it. For COVID queries not focused on misinformation (Figure 4), the majority of content for four out of seven queries was constituted by content from government/administrative websites. For the other three queries, top 10 Google Search outputs came either from journalistic websites (“woher kommt das coronavirus” query) or from healthcare websites (“long covid” query); for the last query (“mensch tier corona”), the majority of results came from journalistic and government/administrative websites. Other types of websites remained largely under-represented, with academic sources appearing only in response to “mensch tier corona” and “corona” queries, and reference websites being prioritised for the “corona-fallzahlen” query for the first two waves.

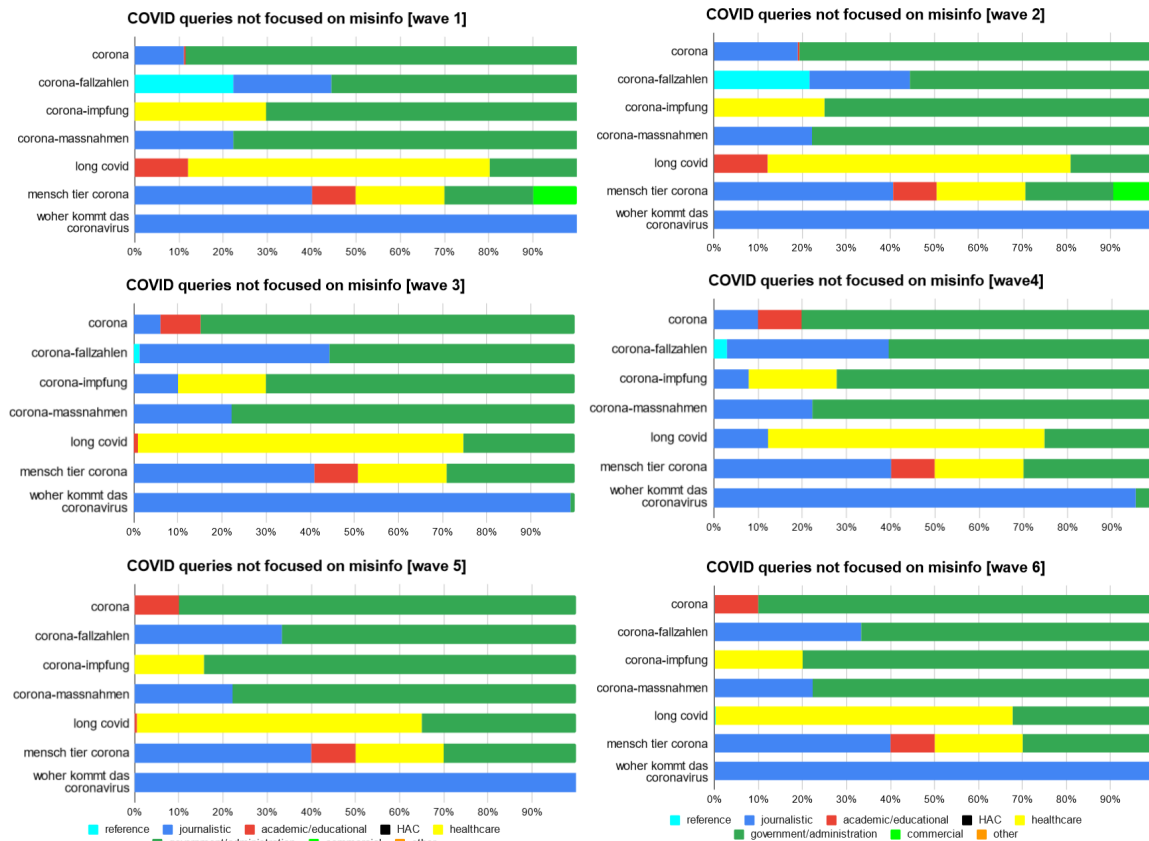


Figure 4. *Distribution of website types per search query for Google search results [COVID queries not focused on misinformation]*

For COVID queries focused on misinformation (Figure 5), we observed substantially less presence of government/administrative websites, which feature prominently only in outputs for “corona-impfung mikrochip” and “5g corona” queries. Instead, most outputs came from the journalistic websites. Another major difference concerns the large amount of outputs coming from the commercial websites, in particular for the “corona-diktatur” and “corona-lüge” queries. While healthcare websites were prioritised less often for these queries, content from academic resources appeared more often for COVID queries focused on misinformation than for ones which did not.

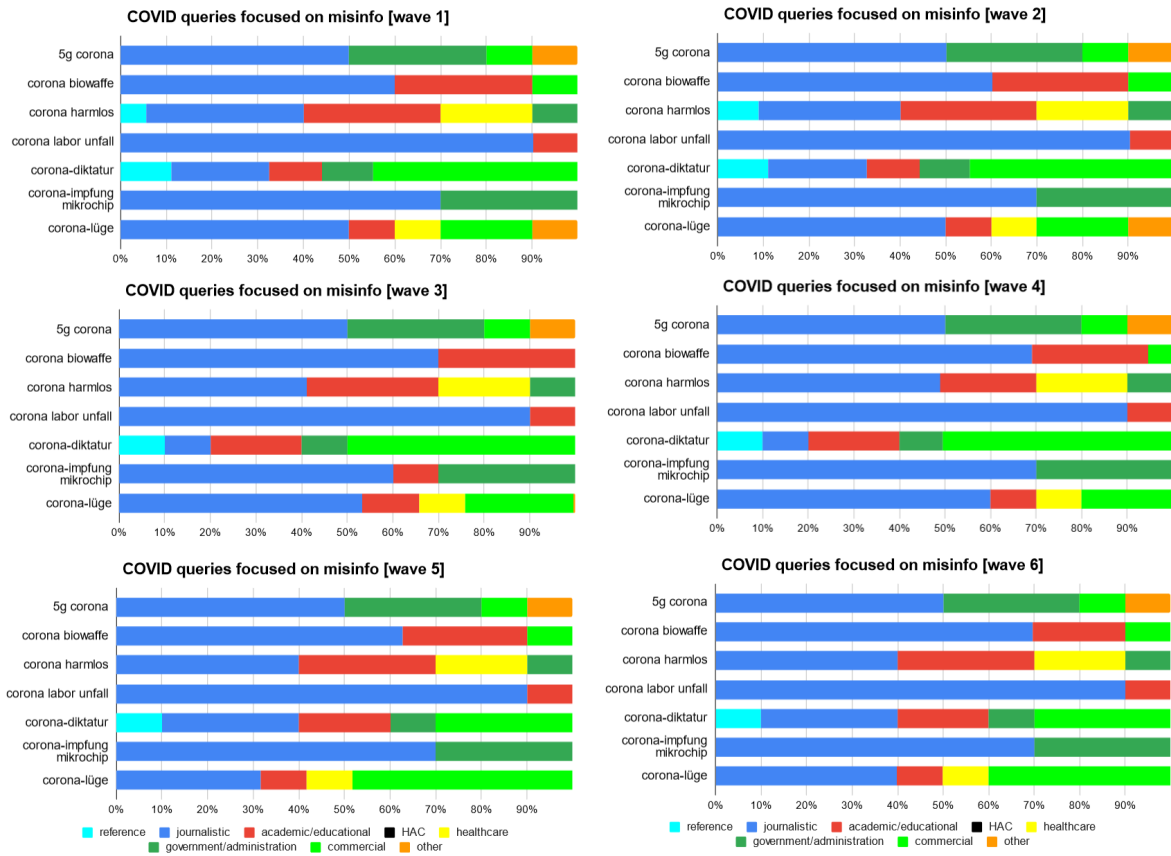


Figure 5. *Distribution of website types per search query for Google search results [COVID queries focused on misinformation]*

In the case of Holocaust queries not focused on misinformation (Figure 6), we observed a selection of sources different from those for COVID queries. Instead of the focus on government/administrative websites for COVID, the larger number of outputs came from academic/educational websites. The only Holocaust query where government/administrative websites constituted the larger portion of outputs was “holocaust erinnerung”, where outputs were often related to official statements released through websites of Bundeszentrale für politische Bildung in Germany or Bundesrat in Switzerland. The same query together with “holocaust” and “6 millionen opfer” attracted the largest number of outputs from journalistic websites (up to 40% of outputs). Finally, for some queries (e.g. “juden holocaust opfer” and “konzentrationslager”), the larger number of outputs (i.e. up to 30%) came from reference websites. However, the only query for which this specific type of source remained consistently prioritised in outputs (i.e. above 40%) was “nie

wider”. The latter query was also the only one for which the substantial number of outputs came from commercial websites.

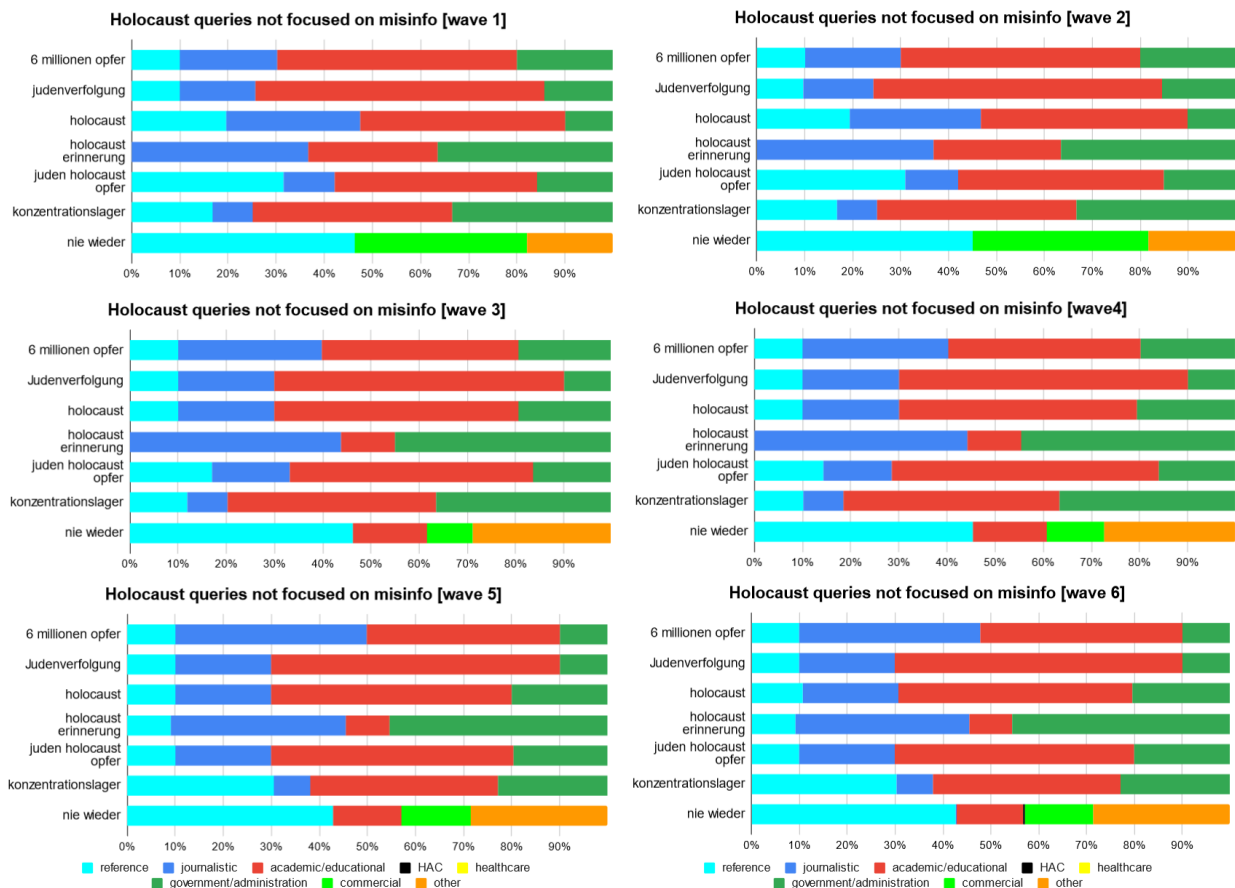


Figure 6. *Distribution of website types per search query for Google search results [Holocaust queries not focused on misinformation]*

Unlike COVID queries, where we observed substantial differences in the source composition between two groups of queries, Holocaust queries which were focused on misinformation (Figure 7) resulted in similar distributions of source categories compared with Holocaust queries not focused on misinformation. Content coming from academic/education websites was most common, followed by journalistic content. For some queries (e.g. “hitler hatte recht” and “holocaustopfer übertrieben”), reference websites constituted the large number of outputs (e.g. up to 30%). Overall, such consistency in sources prioritised for both types of Holocaust queries may have been one of the reasons for the lesser number of outputs supporting misinformation (as compared to COVID queries), as we will discuss in the next section.

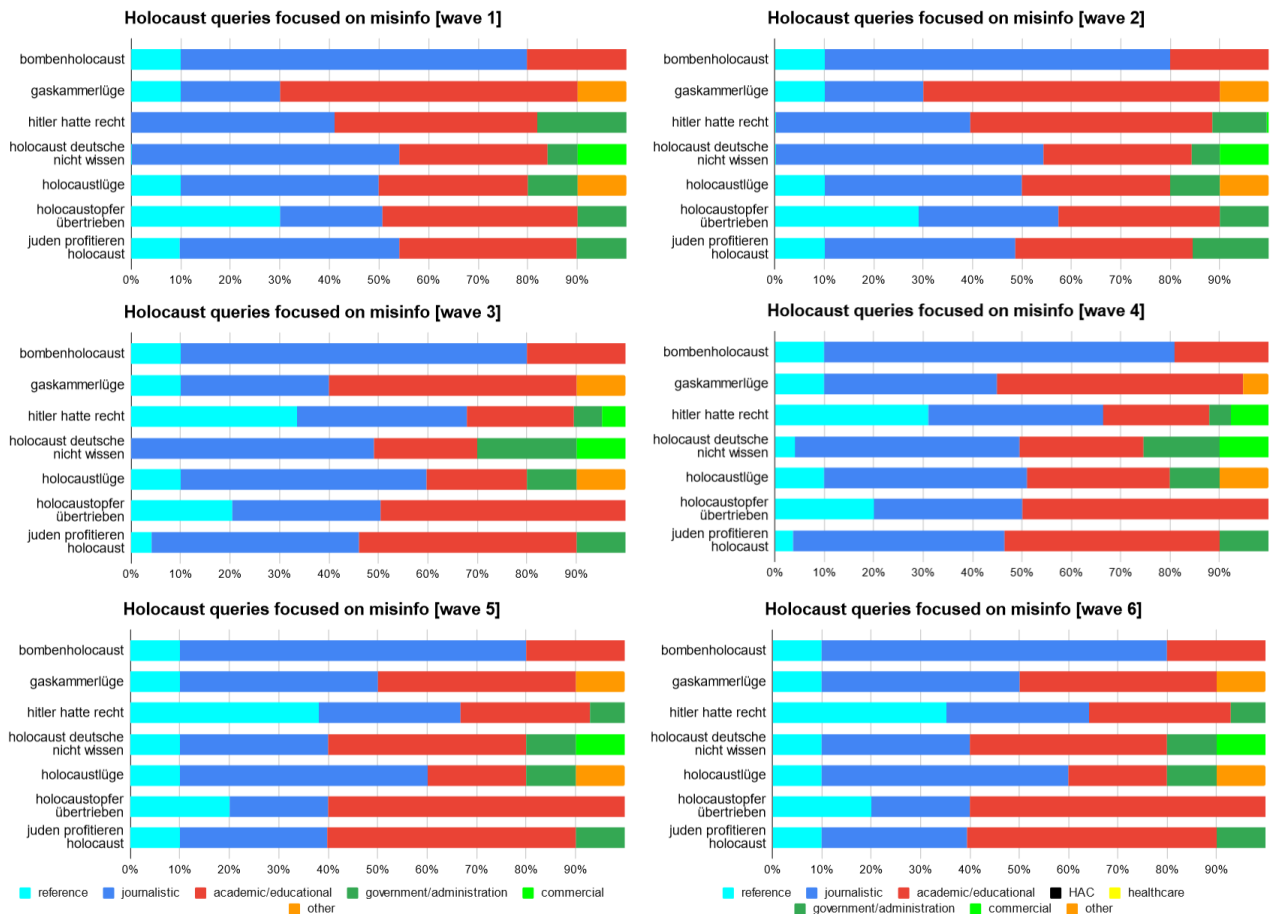


Figure 7. *Distribution of website types per search query for Google search results [Holocaust queries focused on misinformation]*

Impact of Time. Figures 2–3 show that the time factor had a limited influence on the distribution of sources to which different personas were exposed. While there was variation in the distribution of sources between waves for individual personas, such variation remained minor and was also consistent across individual personas (e.g. if one persona was exposed less to specific sources, then this also applied to other personas). In terms of the distribution of individual types of sources, we observed a minor decrease in the number of journalistic sources for all personas and an increase of government/administrative sources for waves 5–6. In terms of the variation between individual personas, we observed more consistency between personas for waves 1–2, followed by increased variation for waves 3–6.

In the case of individual queries (Figures 4–7), we observed a greater impact by the time factor. While a number of COVID and Holocaust queries remained consistent in terms of source

composition for the respective outputs, for some queries, certain types of sources appeared in some waves but not others. For instance, for COVID queries, the “corona harmlos” query resulted in outputs coming from reference websites for waves 1-2, but for the next waves these types of outputs disappeared; the same behaviour was observed for the “corona-lüge” query for the outputs coming from ‘other’ websites. In most cases, however, these time-based changes remained relatively limited (i.e. resulting in a variation of approximately 10% of outputs between waves); the few exceptions were constituted primarily by Holocaust queries, where for the “nie wider” query for waves 1–2 approximately 30% of outputs were constituted by commercial website materials, whereas for the next waves the proportion of such outputs dropped to 10%.

Prioritisation of Misinformation Stances by Google Search

General Composition of Search Outputs. Figure 8 shows that for COVID and Holocaust queries, more than half of search outputs did not mention misinformation at all. 29% outputs for COVID queries and 34% outputs for Holocaust queries debunked misinformation, whereas outputs mentioning misinformation without it being explicitly debunked constituted 11% of search results for COVID and 6% for the Holocaust. Finally, around 5% outputs for COVID and less than 1% of outputs the Holocaust promoted misinformation, thus demonstrating that Google Search algorithms show better performance in terms of filtering out misinformation-supportive content for more established forms of misinformation (e.g. concerning Holocaust denial).

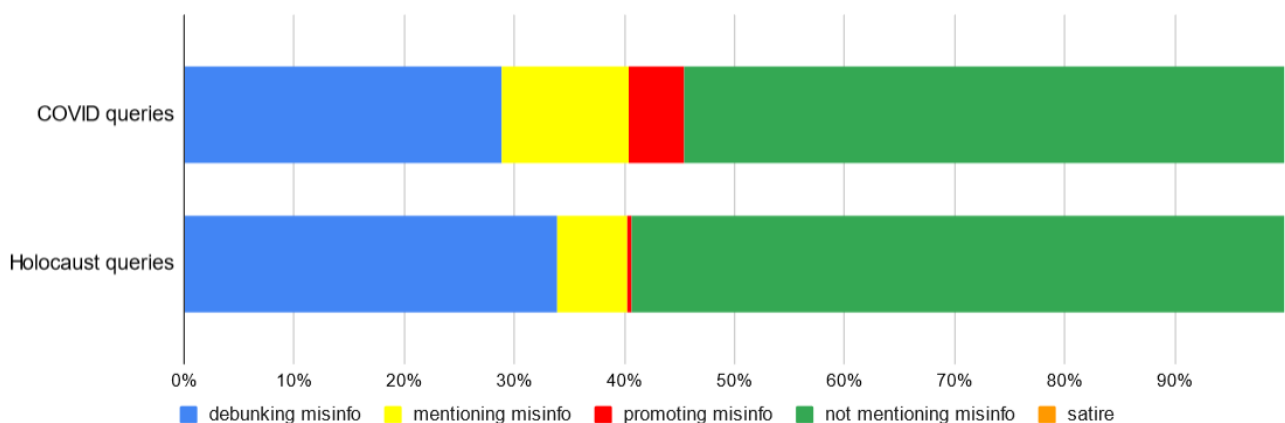


Figure 8. *Aggregated distribution of misinformation stances for all agents across all six waves for COVID and Holocaust queries for Google Search*

Impact of Algorithmic Personas. Similar to the composition of information sources discussed in the previous section, the impact of algorithmic personas on the prioritisation of different misinformation stances was minimal. Figures 9–10 demonstrate that for most cases, exposure to content with different stances was the same between individual personas. For waves 3–4 for COVID queries, there was minor variation in the prioritisation of content debunking misinformation, with it being more present for the left and right personas in wave 3 and the left and centre-right personas in wave 4.

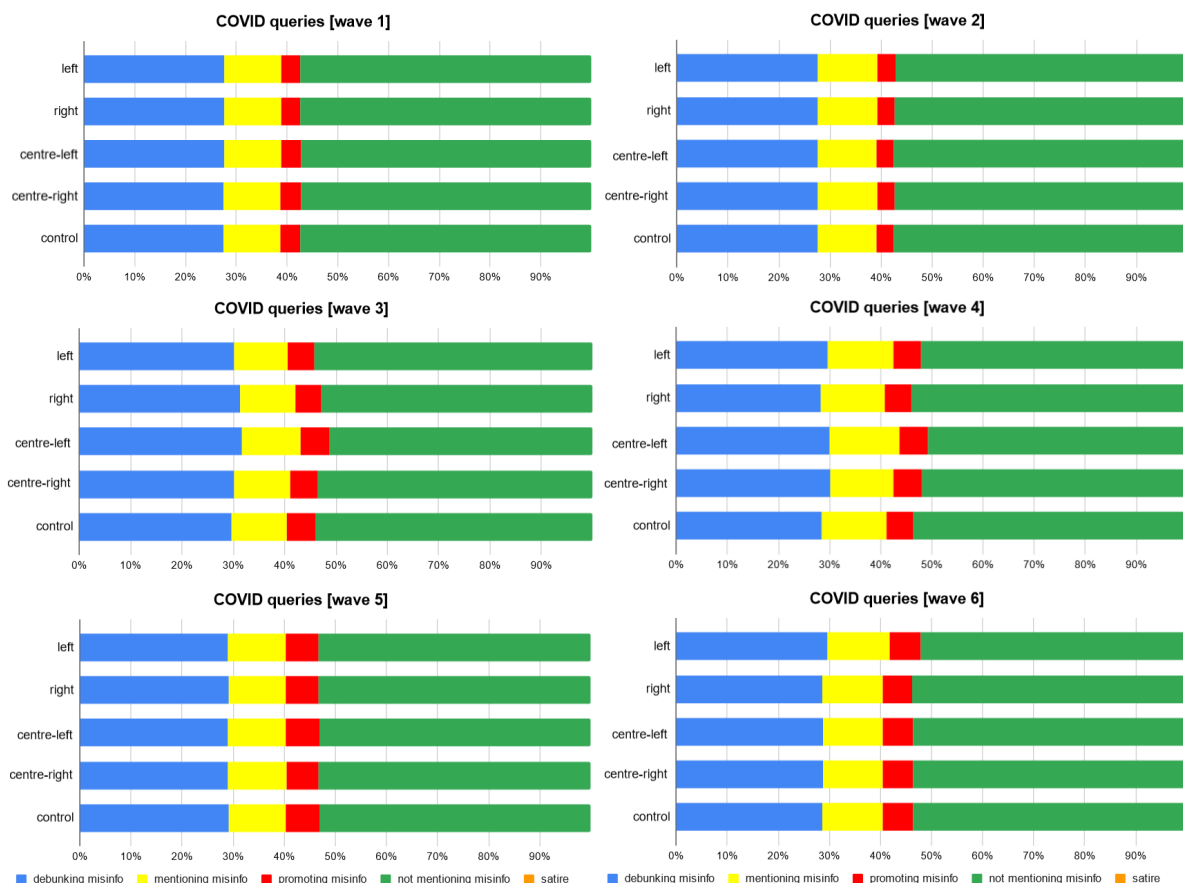


Figure 9. *Distribution of misinformation stances per algorithmic persona for Google search results [COVID queries]*

In the case of Holocaust queries, we found that for wave 4, the centre-right persona received more outputs debunking misinformation and that for wave 3 the centre-left persona received fewer outputs mentioning misinformation. In all these cases, however, variation remained very limited.

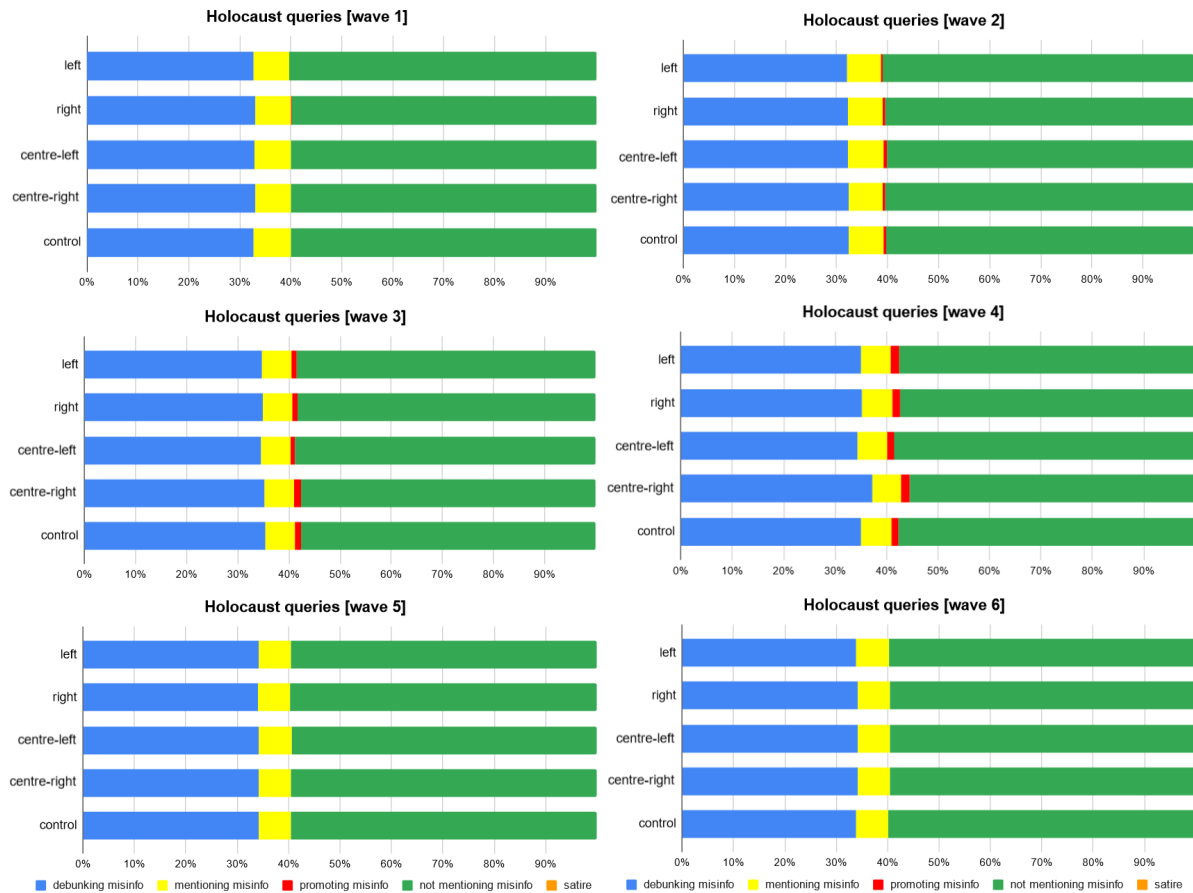


Figure 10. *Distribution of misinformation stances per algorithmic persona for Google search results [Holocaust queries]*

Impact of Search Queries. For misinformation stances, the distinction between queries focused on misinformation and those without such a focus was pronounced. Figure 11 shows that COVID queries without a focus on misinformation resulted almost exclusively in search outputs not mentioning misinformation. The only two queries that consistently deviated from this core finding were “corona-impfung” and “woher kommt das coronavirus”: for the former query, 10% of search outputs across all waves included content debunking misinformation associated with COVID vaccines, whereas outputs for “woher kommt das coronavirus” were mostly split between content

debunking and mentioning misinformation about COVID (except the first/second and the fifth/sixth waves, where 10% of outputs also promoted misinformation).

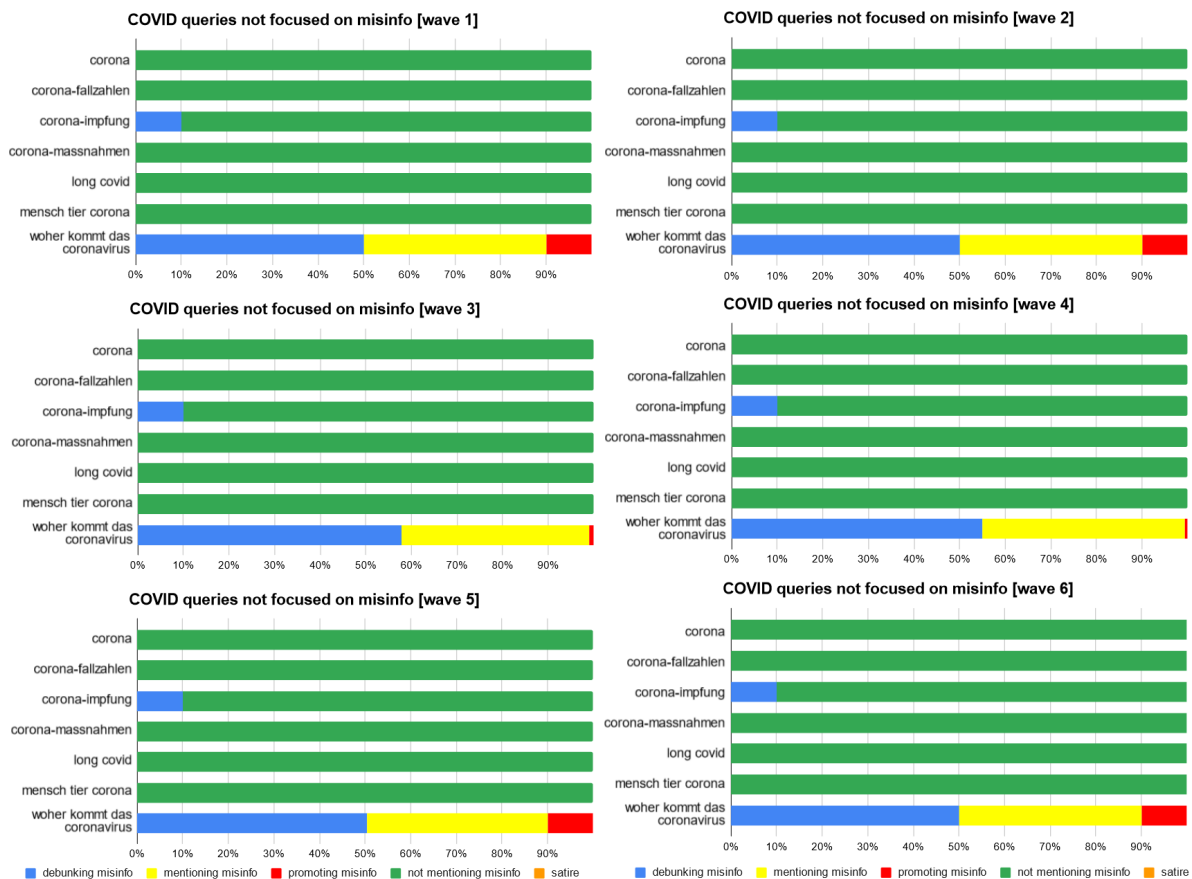


Figure 11. *Distribution of misinformation stances per search query for Google search results*

[COVID queries not focused on misinformation]

In the case of COVID queries with a focus on misinformation, a rather different situation was observed. Figure 12 demonstrates that only one query out of seven – “covid harmlos” – resulted in outputs, the majority of which did not mention misinformation. The other six queries retrieved outputs dealing with misinformation in one form or another: for three (“5g corona”, “corona biowaffe”, “corona-impfung mikrochip”), the outputs were focused on debunking misinformation, whereas another one (“corona labor unfall”) primarily offered outputs mentioning misinformation but not explicitly debunking or promoting it. The remaining two queries (i.e. “corona-diktatur” and “corona-lüge”) were distinguished by the large number of outputs (i.e. up to 40%) promoting misinformation, which might be attributable to these two topics featuring particularly prominently in Swiss political debates about the adoption of measures for dealing with the pandemic.

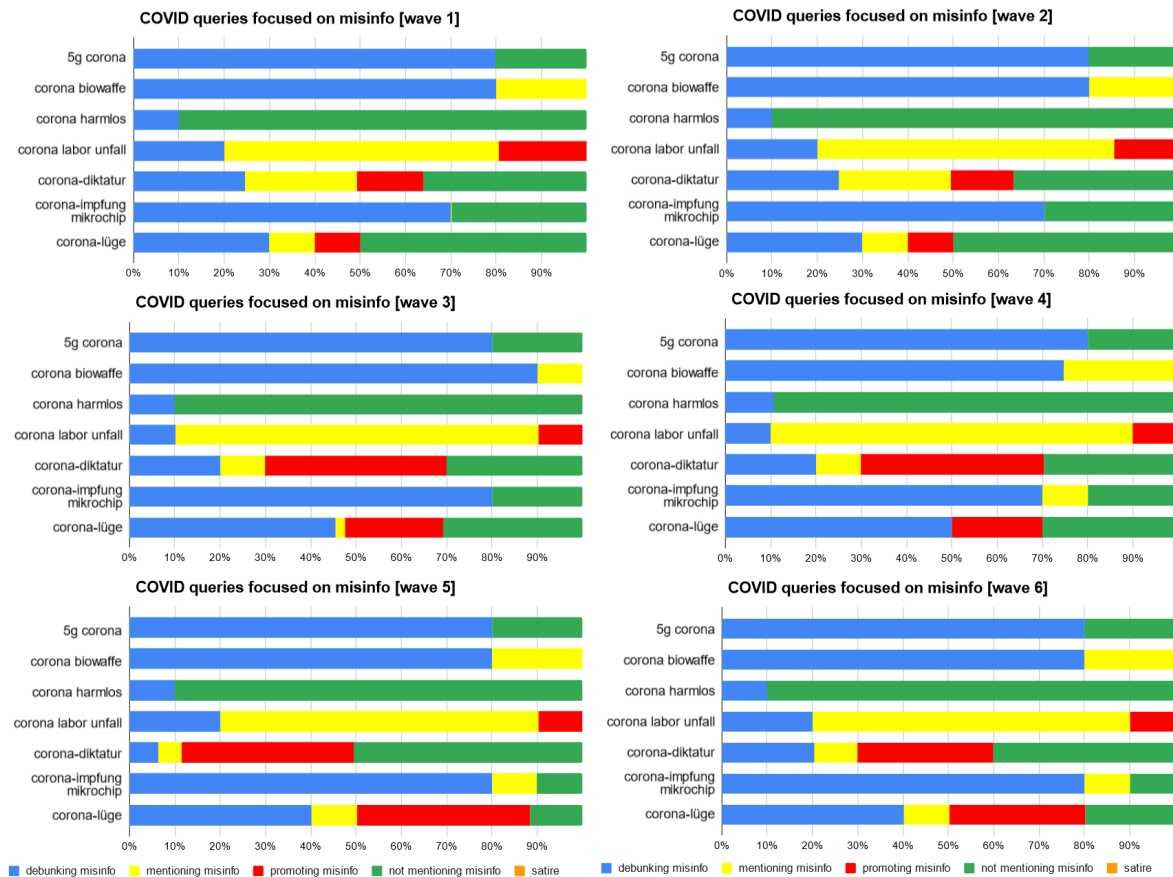


Figure 12. Distribution of misinformation stances per search query for Google search results

[COVID queries focused on misinformation]

For the Holocaust queries not focused on misinformation (Figure 13), we also observed the prevalence of search outputs not mentioning misinformation. There was, however, more debunking content compared with COVID queries, with four Holocaust queries out of seven (i.e. “6 millionen opfer”, “judenverfolgung”, “holocaust”, and “juden holocaust opfer”) consistently receiving between 10% and 30% debunking outputs. For the same queries, approximately 10% of outputs mentioned misinformation without explicitly debunking it.



Figure 13. *Distribution of misinformation stances per search query for Google search results*

[Holocaust queries not focused on misinformation]

For queries focusing on misinformation in the context of the Holocaust (Figure 14), we also observed the large number of outputs not mentioning misinformation. For three queries – “hitler hatte recht”, “holocaustopfer übertrieben” and “juden profitieren holocaust” – such outputs usually constituted the majority of results. For two other queries – “holocaust-lüge” and “gaskammer-lüge” – between 80% and 100% of outputs were constituted by items debunking misinformation. Finally, for the “holocaust deutsche nicht wissen” and “bomben-holocaust” queries, the outputs were split between content debunking and not mentioning misinformation. The latter two queries were also those for which we observed the largest number of outputs (i.e. between 15% and 20%) mentioning misinformation without explicitly debunking or supporting it. Compared with COVID queries focused on misinformation, we observed substantially less content promoting misinformation in response to

Holocaust queries, where it appeared in response only to “hitler hatte recht” and “holocaustopfer übertrieben” queries and did not exceed 10% of all outputs.

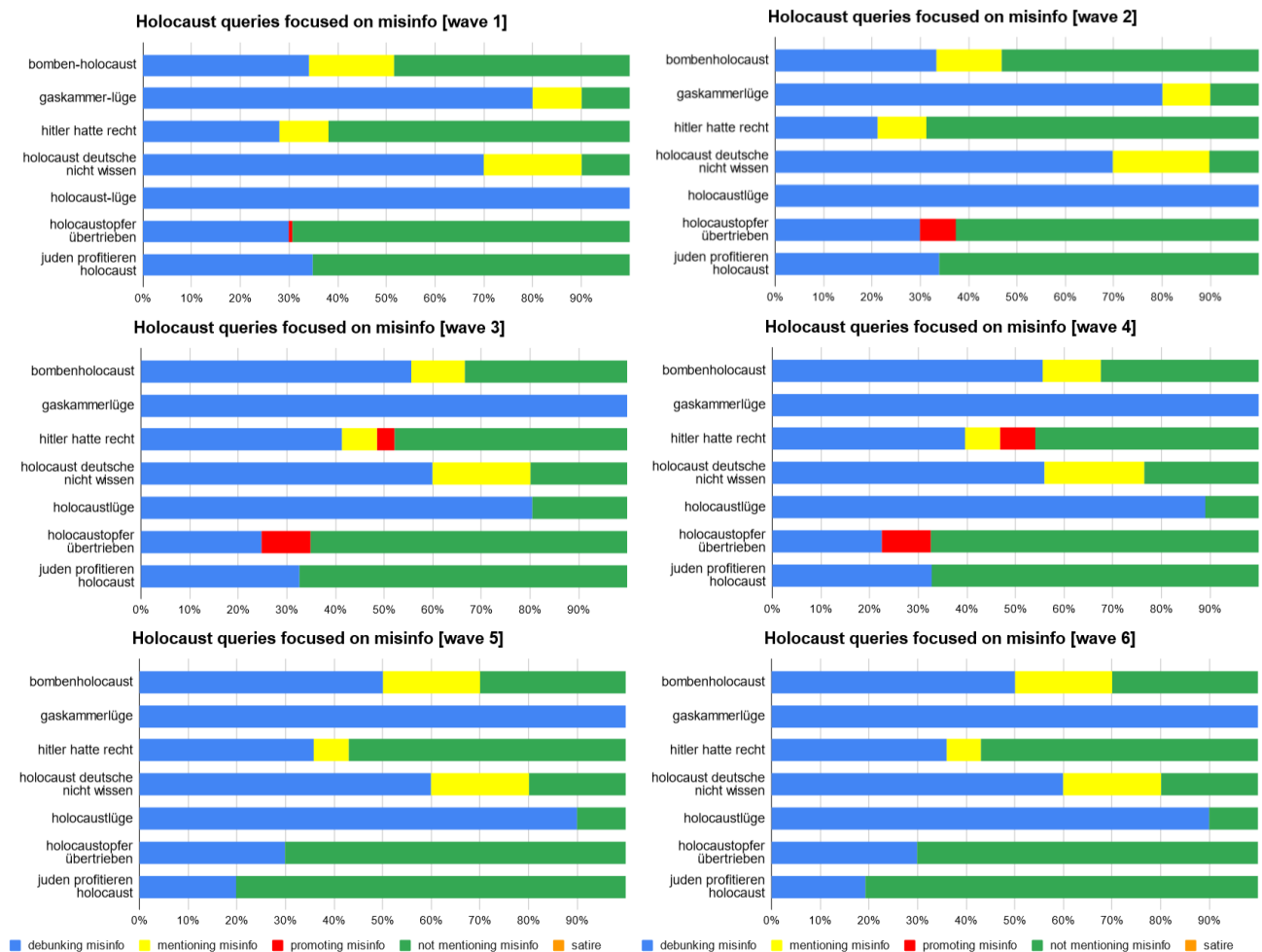


Figure 14. *Distribution of misinformation stances per search query for Google search results*

[Holocaust queries focused on misinformation]

Impact of Time. Figures 9–10 show that exposure of individual personas to misinformation via Google Search changed over time, albeit not dramatically. Specifically, for COVID and Holocaust queries, we observed an increase of search outputs promoting misinformation for waves 3–4 (3–6 for COVID queries) compared with waves 1–2. These changes tended to be consistent across personas with the exception of waves 3–4, where we observed fluctuations among personas concerning exposure to more content debunking misinformation for left, right, and centre-right personas.

For individual COVID and Holocaust queries (Figures 11-14), we observed some variation in the impact of time. The more pronounced impact was traced for queries related to misinformation both for COVID and the Holocaust, where we observed a substantial increase in the amount of content promoting misinformation for waves 3–6 (e.g. for queries “corona-diktatur”, “corona-lüge”, “hitler hatte recht” and “holocaustopfer übertrieben”). In the case of “corona-lüge”, such increase was also accompanied by a growth in the number of outputs debunking misinformation, but for “corona-diktatur”, the number of debunking results also substantially dropped. For queries not focused on misinformation, the impact of time was less pronounced, where the prevalence of specific stances remained largely consistent across all six waves.

Descriptive Analysis of YouTube Search Audit Results

Prioritisation of Information Sources by YouTube Search

General Composition of Search Outputs. Figure 15 shows that compared with Google Search, the visibility of journalistic content in YouTube search outputs was even more pronounced. If for Google Search such content constituted between 20% and 40% of outputs for Holocaust and COVID queries, then for YouTube the proportion increased to the range of 60% to 80%. Similar to Google Search, the majority of the most commonly prioritised videos originated from the channels of Germany-based media companies and outlets (e.g. DW Deutsch, ARD, and ZDF). Another category of channels (‘other’) which appeared frequently both for Holocaust and for COVID queries (e.g. between 8% and 12%) was constituted by channels owned by a broad range of non-institutional actors (e.g. individual bloggers or individual journalists). Examples of most commonly prioritised channels from the ‘other’ category included the personal channel of Martin Wehrle, a German career coach or the music channel of KDM Exclusive group. The high visibility of this specific type of sources was also common for YouTube recommendation outputs, but not for Google Search, where such content appeared in top results extremely rarely.

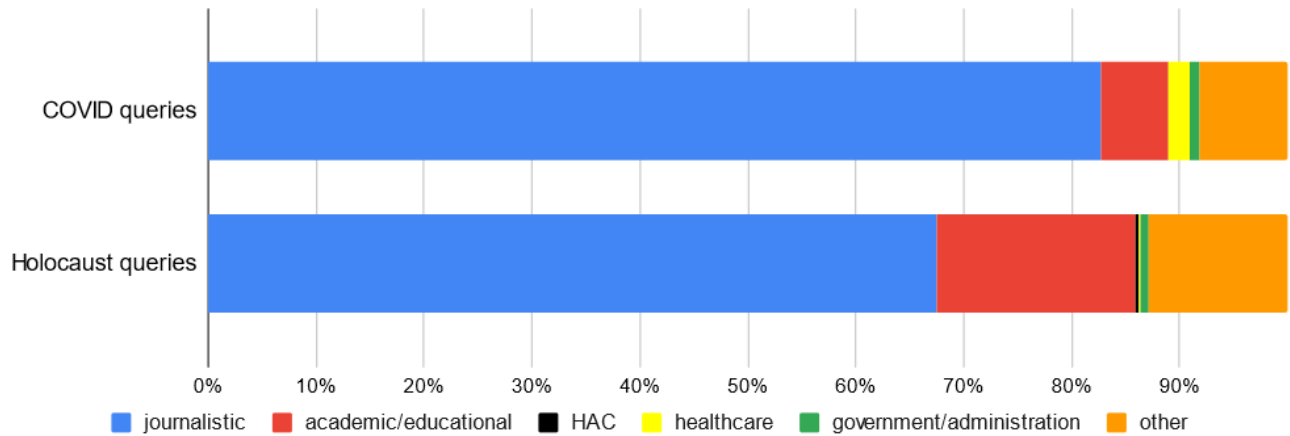


Figure 15. *Aggregated distribution of information sources for all agents across all six waves for COVID and Holocaust queries for YouTube search*

In the case of Holocaust queries, we observed a higher presence of academic/education outputs in search results, which was also common for Google Search. Examples of channels providing such outputs included Geschichte-simpleclub (i.e., a German educational channel with history-related videos) and zeitzeugen-portal (i.e., a German channel collecting videos of eye witnesses of important historical events); most of these channels were from Germany. Holocaust queries were again more likely to retrieve outputs from HAC media (e.g. MrMarxismo YouTube channel); similarly, healthcare channels (e.g. a channel of the World Health Organization) had a presence in YouTube search outputs for COVID queries, albeit substantially less prominent compared with Google. There were also a few outputs from government/administrative and commercial channels, but much less so compared with Google Search.

Impact of Algorithmic Personas. Despite the above-mentioned differences in source composition between Google and YouTube search, the impact of algorithmic personas in both cases was similarly negligible. Figures 16-17 show that like for Google, there was minor variation in the composition of outputs between different personas for YouTube search (mostly for Holocaust queries, where for some data collection waves the centre-right and left personas received more outputs from journalistic channels).

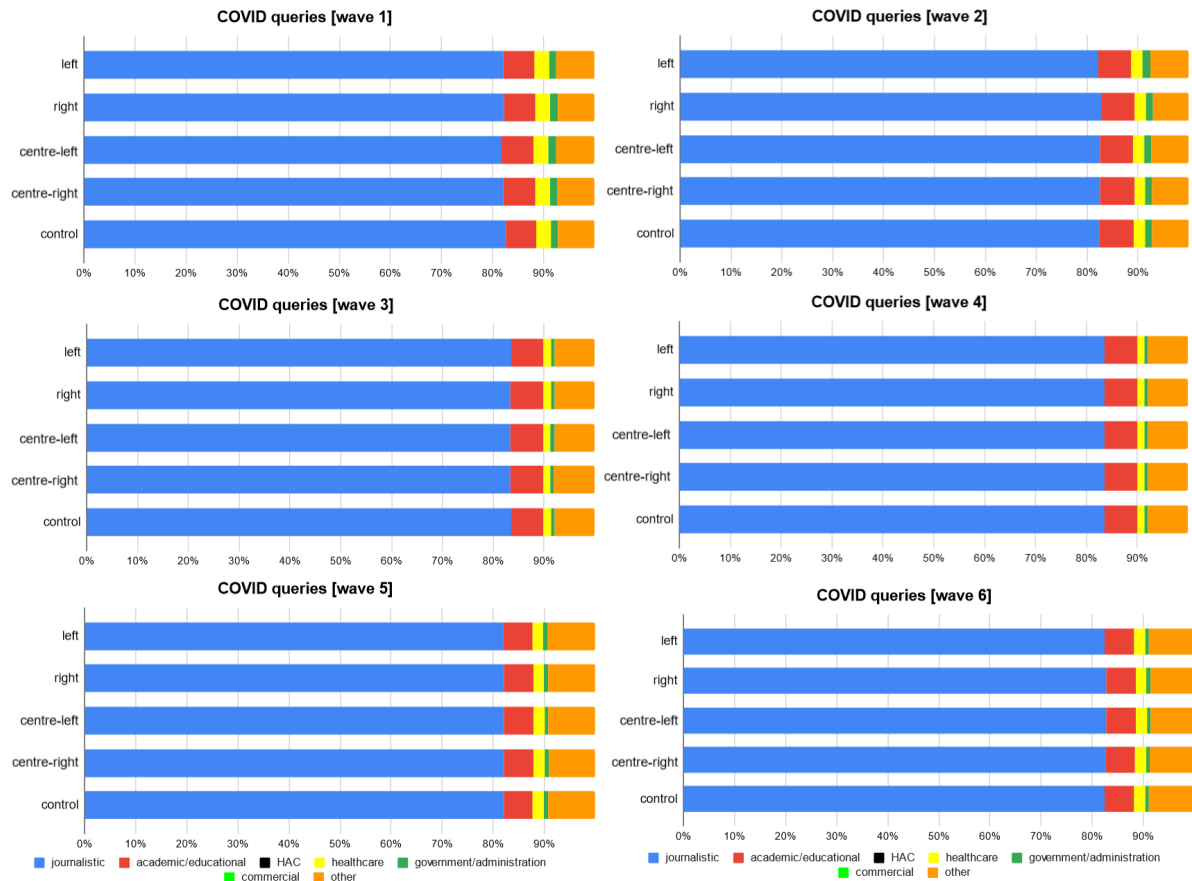


Figure 16. *Distribution of website types per algorithmic persona for YouTube search results*
[COVID queries]

Additionally, for Holocaust queries, centre-left and left personas received for some waves more outputs from HAC media and less from healthcare channels. However, just as in the case of Google, these differences were not consistent across waves and could possibly be attributable to output randomisation.

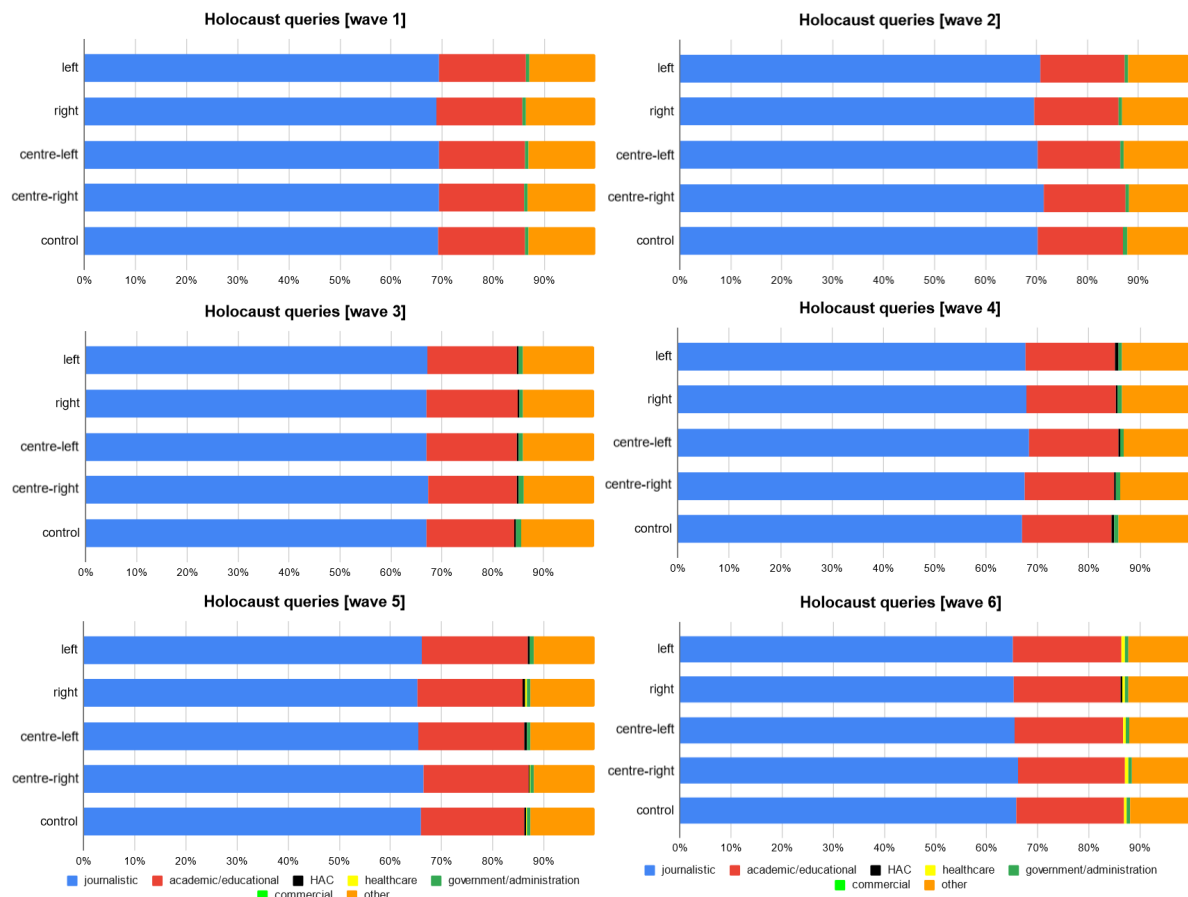


Figure 17. *Distribution of website types per algorithmic person for YouTube search results*
[Holocaust queries]

Impact of Search Queries. The analysis of sources prioritised in relation to COVID queries focused or not focused on misinformation (Figures 18–19) for YouTube search shows fewer differences compared with Google Search. In the case of queries not focused on misinformation (Figure 18), we observed that more than half of all search outputs came from journalistic channels (except the “corona” query for the fifth/sixth waves). For some queries (e.g. “corona-fallzahlen” and “corona-massnahmen”) journalistic channels were the only source of content. For the “corona” query we also found the large number of outputs coming from the ‘other’ channels; furthermore, this was one of two queries (together with “long covid”), for which we observed outputs from healthcare-related channels. The limited presence of healthcare content was another difference from Google Search, where such content was retrieved substantially more often.

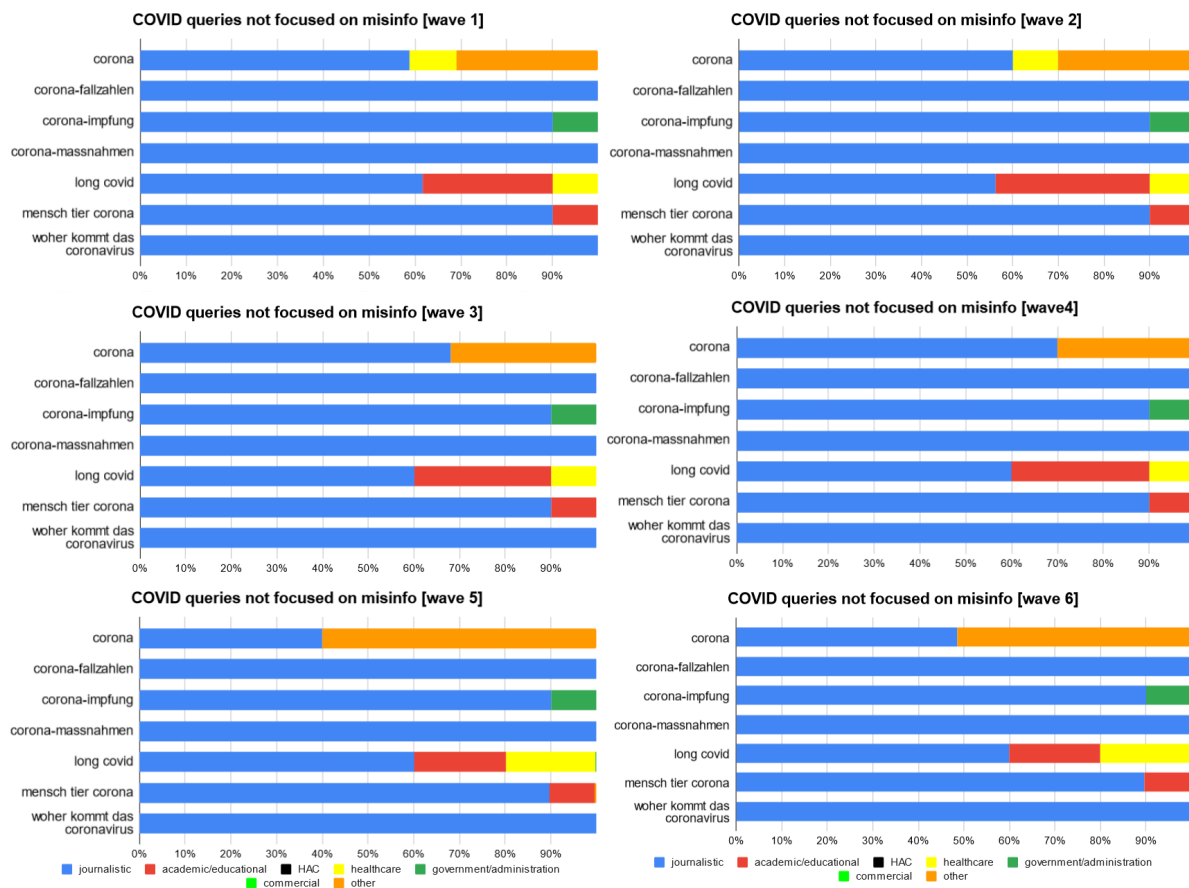


Figure 18. *Distribution of website types per search query for YouTube search results [COVID queries not focused on misinformation]*

For the COVID queries focused on misinformation (Figure 19), journalistic channels also provided the majority of content retrieved via YouTube search, albeit only one of such queries (i.e. “corona labor unfall”) resulted in retrieving outputs exclusively from the journalistic channels. For four other queries (i.e. “corona-diktatur”, “corona-lüge”, “corona-impfung mikrochip”, “5g corona”) a number of outputs (i.e. between 10% and 20%) were retrieved from the ‘other’ channels (e.g. the personal channel of Carsten Strauch, a German film producer, or the personal channel of Martin Wehrle). Similarly to the queries not focused on misinformation, only a few outputs (primarily for the “corona biowaffe” query) came from the healthcare-related channels.

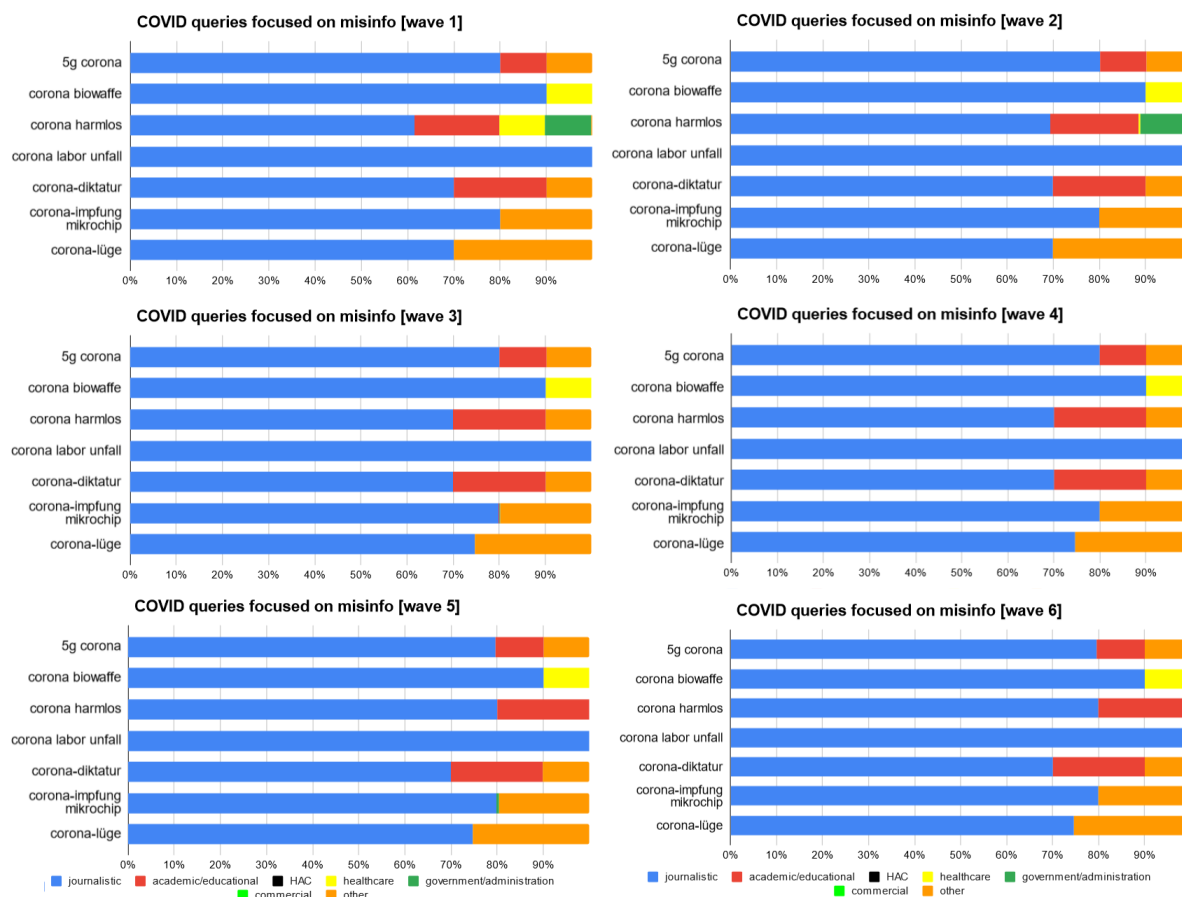


Figure 19. *Distribution of website types per search query for YouTube search results [COVID queries focused on misinformation]*

For the Holocaust queries (Figures 20–21), we also observed a high visibility of journalistic channels. For the queries not focused on misinformation (Figure 20), journalistic content constituted the consistent majority for four queries out of seven (i.e. “holocaust erinnerung”, “6 millionen opfer”, “konzentrationslager”, “juden holocaust opfer”). The other three queries resulted either in the same or a larger number of outputs coming from the academic/educational channels or, in the case of the “nie wieder” query, ‘other’ channels. The latter case can be attributed to YouTube search prioritising non-Holocaust-related content in response to the query, such as popular culture items (e.g. Germanophone pop songs entitled “Nie Wieder” from Nimo or Bushido). Also, for one of queries - i.e. “konzentrationslager” - we found content retrieved from a HAC channel (i.e. MrMarxismo).

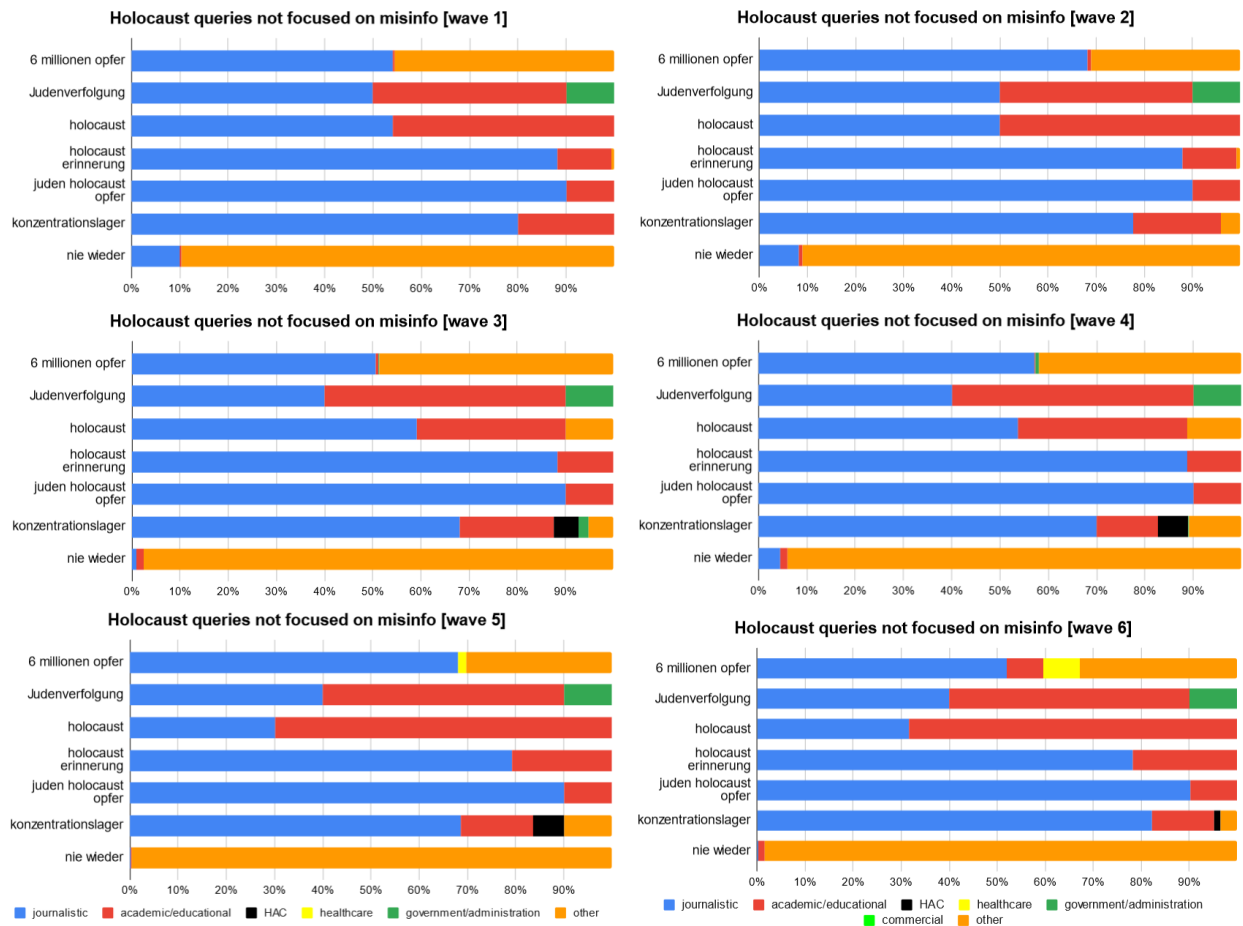


Figure 20. *Distribution of website types per search query for YouTube search results [Holocaust queries not focused on misinformation]*

In the case of Holocaust queries focused on misinformation (Figure 21), journalistic content constituted more than 50% of outputs for six out of seven misinformation-focused queries (with the exception of the “bomben-holocaust” query, where roughly the same number of outputs came from academic/educational channels, such as University of California Television or Dinge erklärt channels). Besides journalistic and academic/educational channels, we also observed some outputs coming from ‘other’ channels, in particular for the “6 millionen opfer” query. Unlike Google Search outputs for misinformation-focused Holocaust queries, we did not observe any outputs from government/administration channels or commercial channels.



Figure 21. *Distribution of website types per search query for YouTube search results [Holocaust queries focused on misinformation]*

Impact of Time. The factor of time had a limited impact on variation in the distribution of sources between algorithmic personas (Figure 16-17). Similar to Google Search, we observed minor fluctuations across the waves, such as a small drop in the number of journalistic outputs for Holocaust queries and an increase of outputs from ‘other’ as well as academic/educational channels. These fluctuations, however, usually affected all personas equally, so they can likely be attributed not to the impact of specific web histories but instead to overall changes in source relevance as perceived by the YouTube search algorithm. The only case when the influence of the factor of time on the sources to which personas were exposed was more impactful concerned the growing presence of HAC media content for waves 3–6 for the Holocaust queries, albeit such increase remained low.

In terms of the impact of time on the selection of sources for individual queries (Figures 17-21), we again observed major similarities between YouTube search and Google Search. For most COVID and Holocaust queries, we observed little change in terms of source composition between waves. However, for some queries, different waves resulted in the presence of different types of sources: for instance, for the “konzentrationslager” query, outputs for waves 3–6 included HAC content constituting between 2% and 7% of outputs (depending on the wave), whereas for waves 1–2 such content was absent. In the case of the “nie wider” query, we observed a reverse situation where journalistic content present in the outputs for waves 1–4 disappeared for waves 5–6. In a few cases, the change in the distribution of sources was rather substantial as, for instance, for the “corona” query, where the proportion of content from ‘other’ channels jumped from 30% for waves 1–4 to 60% for waves 5–6, but such major fluctuations remained rare.

Prioritisation of Misinformation Stances by YouTube Search

General Composition of Search Outputs. Figure 22 shows that similar to our observations on source composition, we found that YouTube search followed the same pattern in terms of prioritisation of misinformation stances as Google Search (with the only exceptions being a lesser proportion of content debunking misinformation and a higher number of outputs falling into the category of satire). For the COVID queries, 60% of YouTube search outputs did not mention misinformation. The second-most common stance was misinformation debunking, which was present in 22% of outputs, followed by 14% of outputs mentioning misinformation. The remaining 4% of outputs were divided between content items promoting misinformation or containing satirical claims.

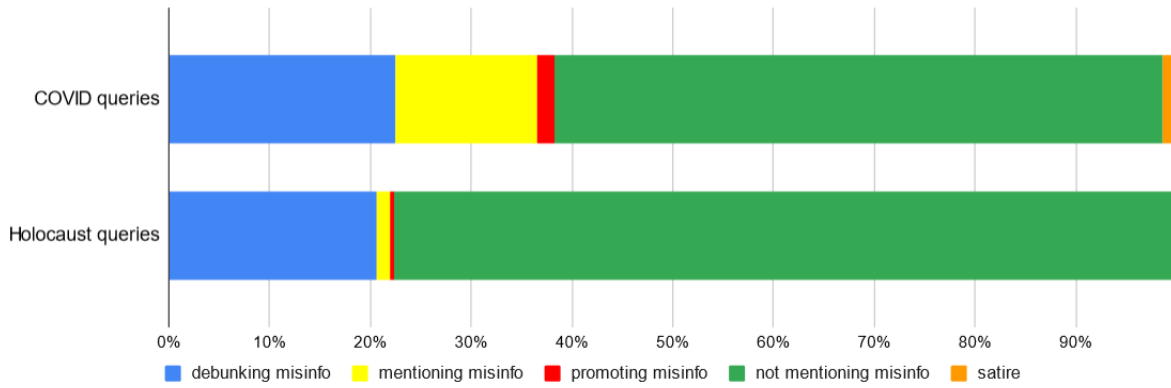


Figure 22. *Aggregated distribution of misinformation stances for all agents across all six waves for COVID and Holocaust queries for YouTube search*

In the case of Holocaust search outputs (Figure 16), an even larger number of outputs (77%) did not mention misinformation, whereas misinformation-debunking items constituted 21% (i.e. almost the same as in the case of COVID queries). However, the number of items promoting or simply mentioning misinformation for the Holocaust was substantially smaller and together with satirical items constituted approximately 2%.

Impact of Algorithmic Personas. Figures 23–24 show that similar to Google Search, algorithmic personas had little influence on the prioritisation of specific stances on Holocaust and COVID misinformation on YouTube search. In the case of COVID queries (Figure 23), we observed minor variation in the exposure of personas to videos debunking misinformation; however, such variation occurred only during the first two waves and was rather inconsistent between them. If during wave 1, the control persona was more exposed to debunking content, then during wave 2 it received less outputs with debunking stance compared to other personas. Instead, during wave 2 left and right personas received more outputs which debunked misinformation.

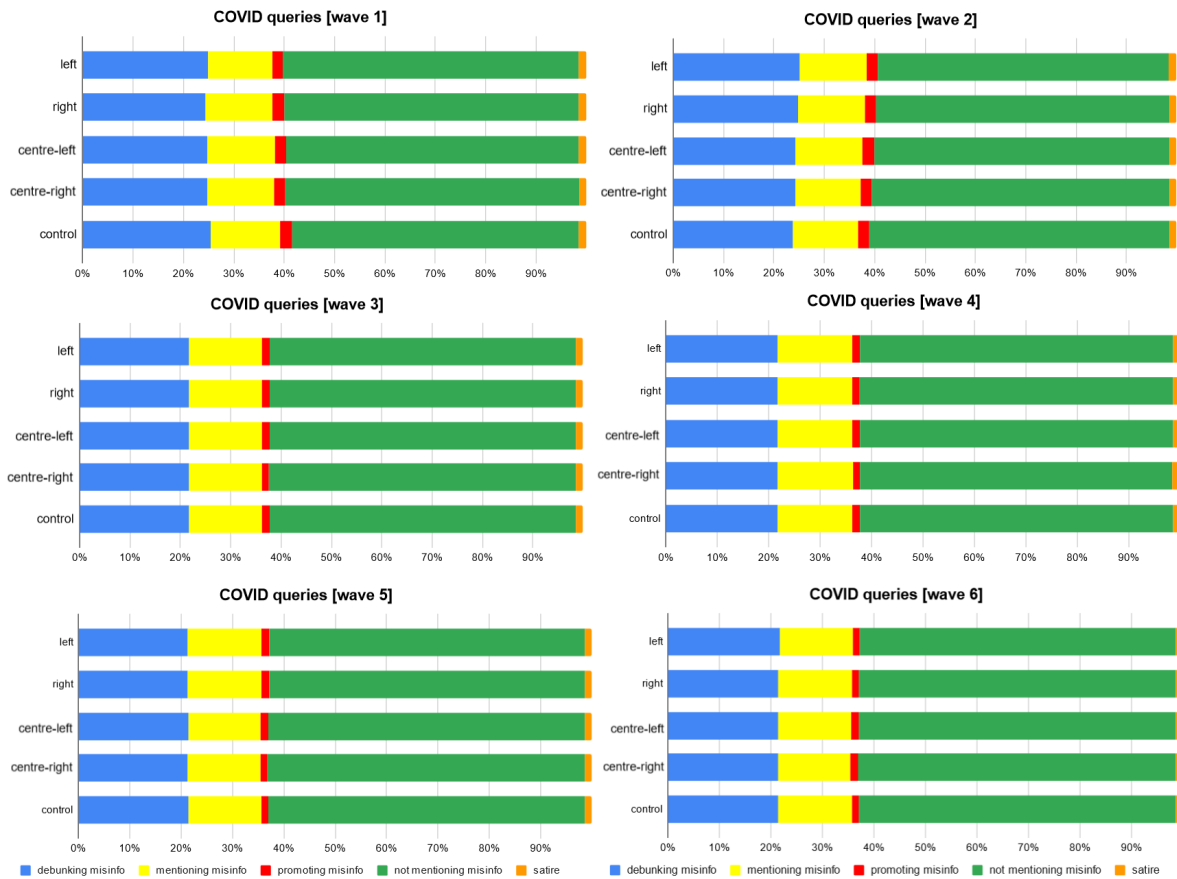


Figure 23. *Distribution of misinformation stances per algorithmic persona for YouTube search results [COVID queries]*

In the case of Holocaust queries (Figure 24), we observed minor variations between personas in terms of exposure to content which was not only debunking, but also promoting misinformation. Unlike COVID queries, these variations occurred not only during the first two waves. In several cases, left and centre-right personas received more exposure to content debunking misinformation and the left persona also received more exposure to content promoting misinformation for one of waves. However, similar to Google Search, the consistency and the score of the differences between personas was limited.

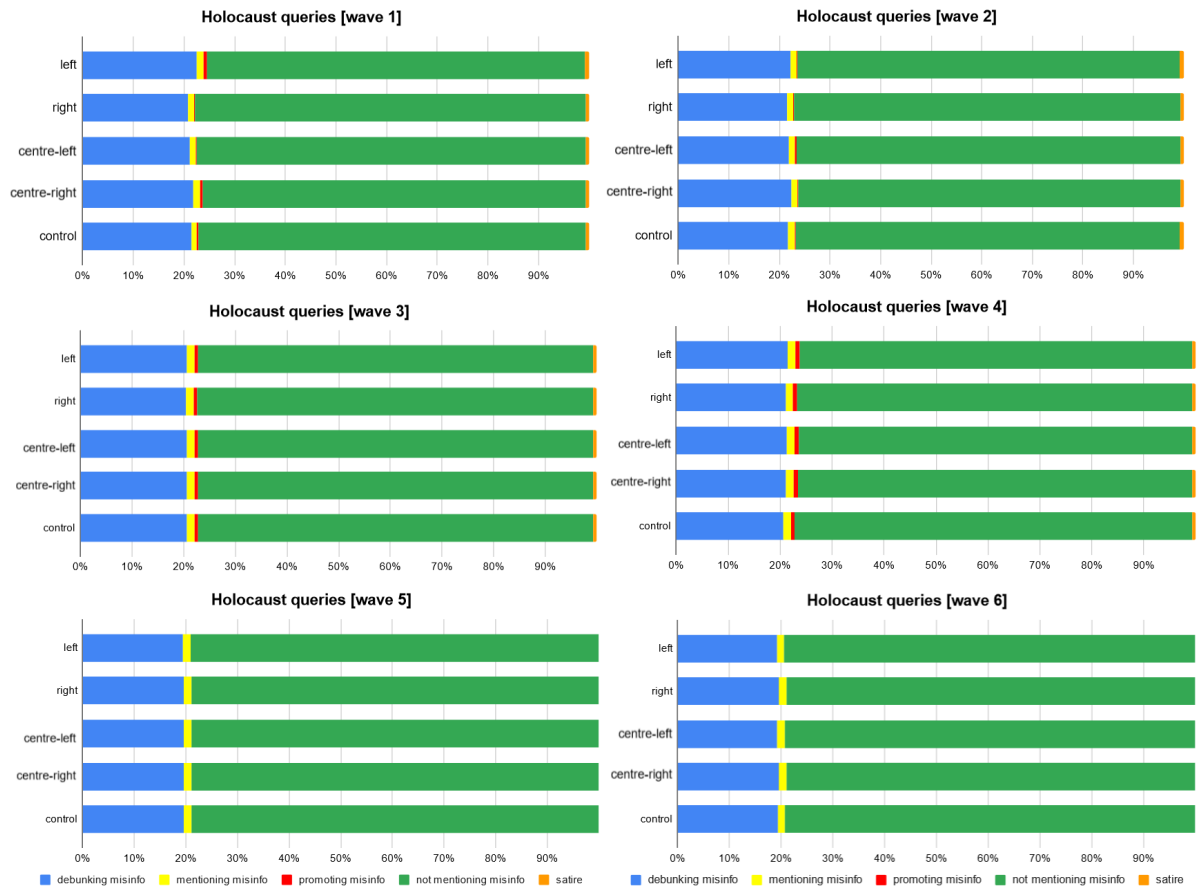


Figure 24. *Distribution of misinformation stances per algorithmic person for YouTube search results [Holocaust queries]*

Impact of Search Queries. Similar to the Google Search outputs, we observed substantial differences in stances on misinformation between individual queries for YouTube search (Figures 25–28). For COVID queries not focused on misinformation (Figure 25), most outputs did not mention misinformation at all. Just as in the case of Google Search, the only major exception was related to the “woher kommt das coronavirus”, outputs for which contained a number of videos debunking, but also mentioning misinformation. It was also the only query, for which videos promoting misinformation were retrieved: their number was the same as for Google Search (i.e. around 10% of outputs). At the same time, for YouTube search more queries not focused on misinformation (e.g. “corona-impfung”, “corona-massnahmen”) resulted in outputs mentioning misinformation without its explicit debunking.

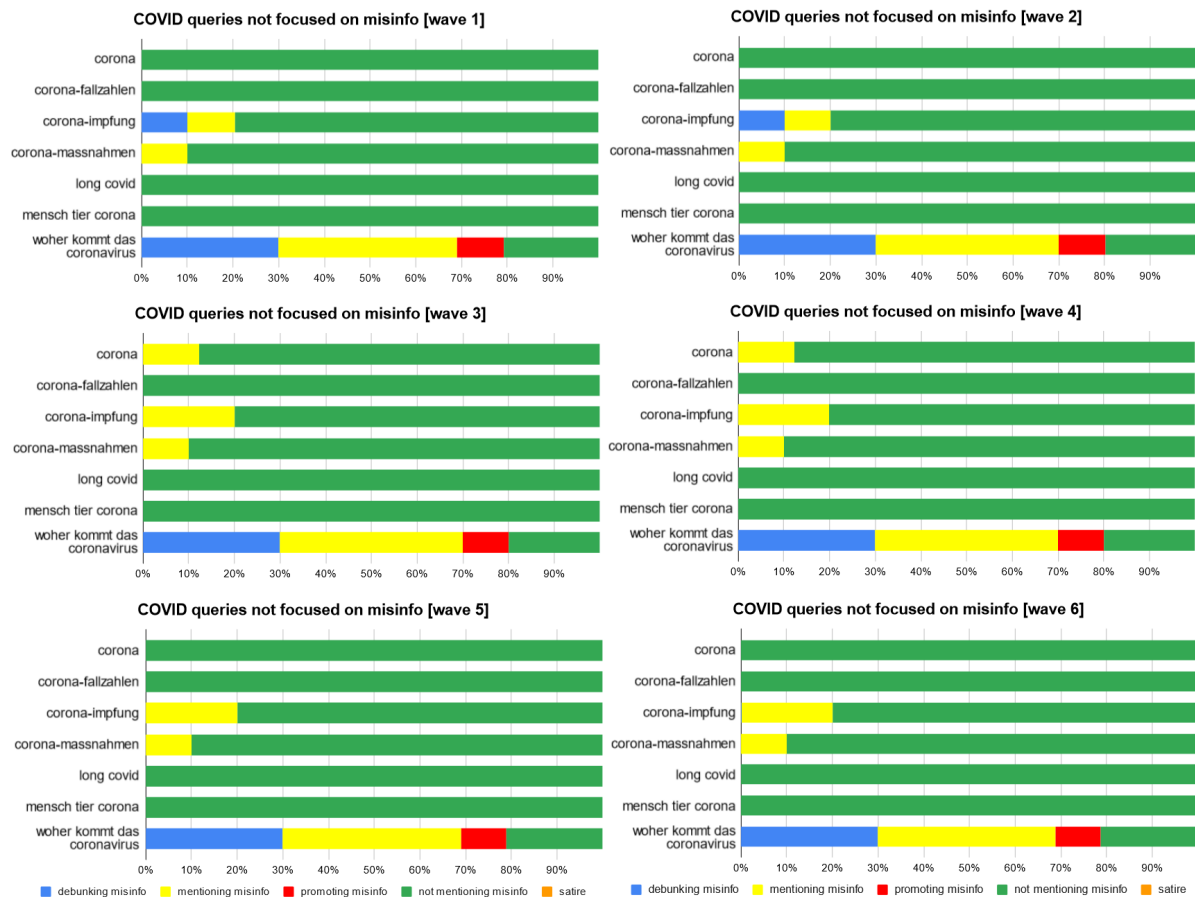


Figure 25. *Distribution of misinformation stances per search query for Google search results*
[COVID queries not focused on misinformation]

For COVID queries focused on misinformation (Figure 26), the number of outputs not mentioning misinformation was also rather high. Such outputs were prevalent for three queries out of seven (i.e. “corona harmlos”, “corona-lüge”, “corona-impfung mikrochip”), thus making this stance more prevalent than in the case of Google Search. One of the remaining four queries (i.e. “corona 5g”) was dominated by debunking outputs (as in the case of Google Search), whereas for the remaining queries most outputs were constituted by a combination of content debunking and mentioning misinformation (or only mentioning in the case of “corona labor unfall”). The “corona labor unfall” query was also the only one attracting outputs promoting misinformation, albeit such outputs remained relatively few (i.e. around 10%). Overall, the number of outputs promoting misinformation was lesser for YouTube search than for Google Search, but it is important to keep in mind that for YouTube we examined only first three minutes of each video.

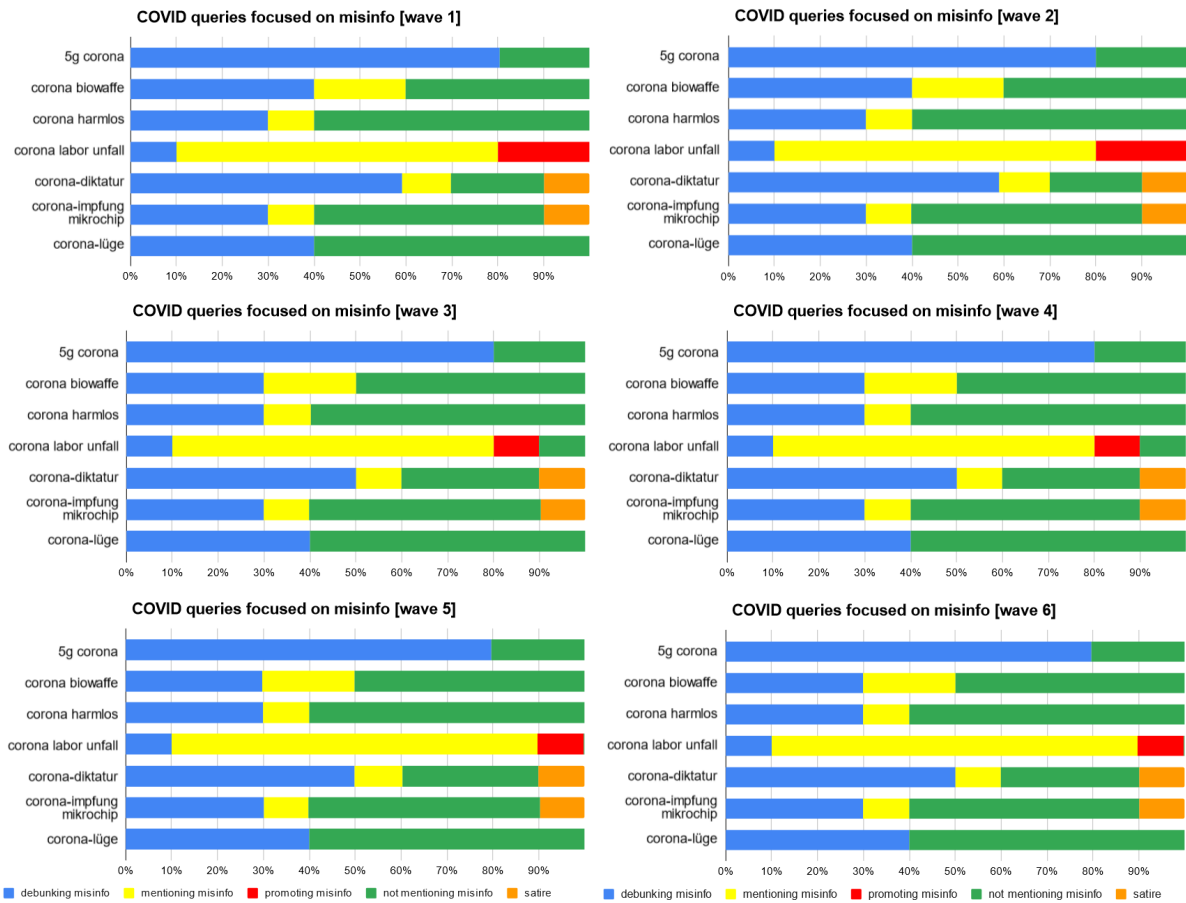


Figure 26. *Distribution of misinformation stances per search query for YouTube search results*
[COVID queries focused on misinformation]

In the case of Holocaust queries not focused on misinformation (Figure 27), we observed an even larger presence of outputs not mentioning misinformation than in the case of COVID queries. One of the queries - i.e. “nie wieder” - resulted in all outputs not mentioning misinformation. Unlike Google Search outputs, we did not find any outputs mentioning misinformation. Instead, three queries (i.e. “holocaust erinnerung”, “juden holocaust opfer”, “holocaust”) resulted in a substantial number of outputs (i.e. around 20%) debunking misinformation. We also observed a few outputs promoting misinformation, but it happened only for one wave and for a single query (i.e. “holocaust”).

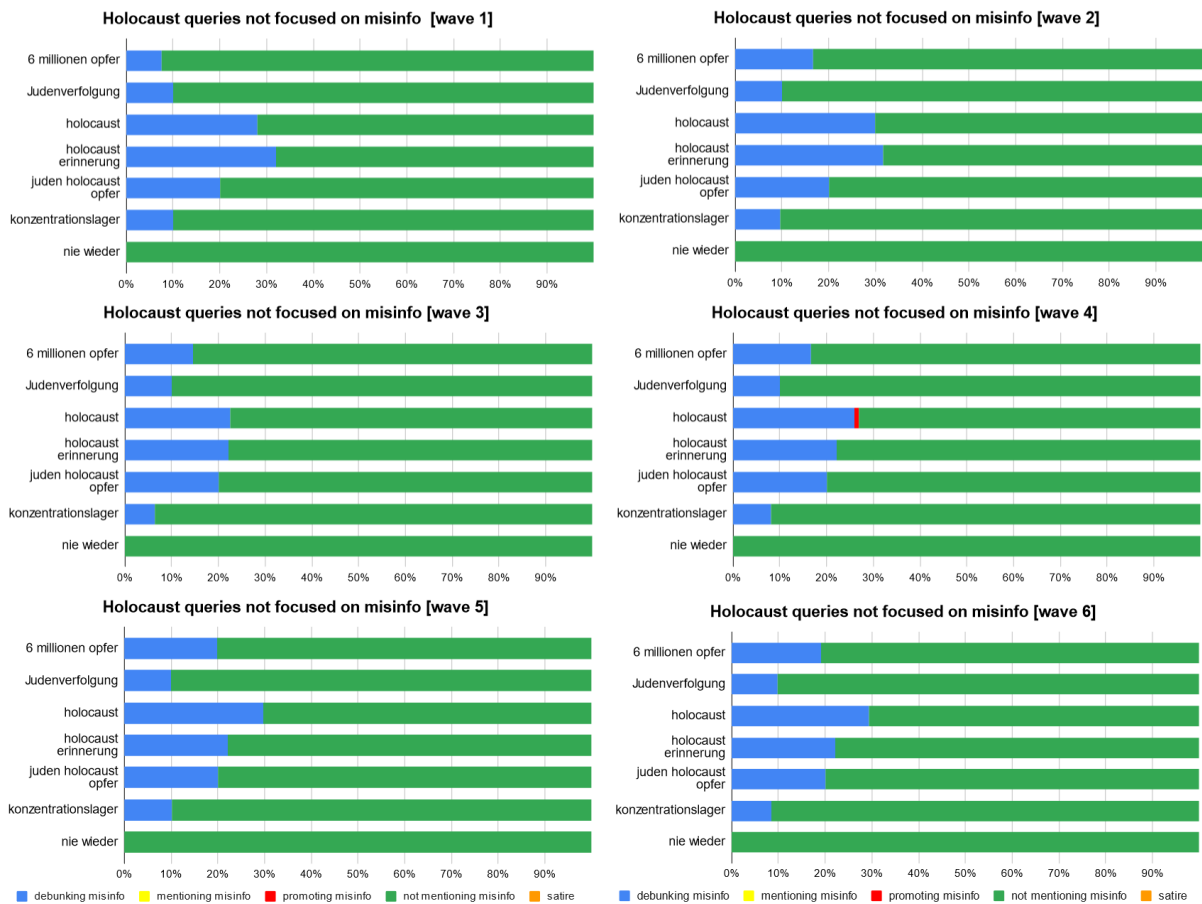


Figure 27. *Distribution of misinformation stances per search query for YouTube search results*

[Holocaust queries not focused on misinformation]

In the case of Holocaust queries focused on misinformation (Figure 28), we also observed that the majority of outputs consistently did not mention misinformation for all queries except “gaskammer-lüge” and “holocaust-lüge”. For the latter two queries, there was a large number of outputs debunking misinformation (up to 60%). Additionally, together with “bomben-holocaust”, “gaskammer-lüge” and “holocaust-lüge” were the only queries which resulted in outputs mentioning disinformation, but not explicitly debunking or promoting it; in all cases, the misinformation-mentioning outputs constituted around 10%. Compared with Google Search, YouTube search outputs contained less results promoting misinformation: these appeared only for the “holocaustopfer übertrieben” and “bomben-holocaust” queries (the former query was also the one responsible for such outputs for Google search). However, Google Search also was characterised by the larger number of outputs debunking misinformation for the most of queries.



Figure 28. *Distribution of misinformation stances per search query for YouTube search results*
[Holocaust queries focused on misinformation]

Impact of Time. Similar to Google Search, we observed a limited impact of time on the prioritisation of specific stances for YouTube search between algorithmic personas (Figures 23-24). There was variation between waves 1–2 and waves 3–6 for COVID queries, with the videos offering a debunking stance gaining less visibility for the four latter waves. In the case of Holocaust queries, there was an increase in videos promoting conspiracies for waves 3–4, but this increase was rather minor.

In the prioritisation of sources for individual queries (Figures 25-28), we also found similarities to Google Search. In the majority of cases, the distribution of stances for individual cases was even more consistent over time than in the case of Google Search, but for some queries we observed changes. For instance, for the “bomben-holocaust” query, we observed an increase of videos mentioning misinformation in waves 3–4, whereas for the “corona-impfung” query outputs debunking disinformation disappeared for waves 3-6. In the majority of these cases, the time-based changes

affected around 10% of the system outputs. In terms of cross-query changes, we also observed the tendency for the decreasing number of outputs promoting misinformation for waves 3-6 for both Holocaust and COVID queries, whereas for Google Search the opposite trend was identified.

Descriptive Analysis of YouTube Recommendation Audit Results

Prioritisation of Information Sources by YouTube Recommendations

General Composition of Recommendation Outputs. Similar to Google Search and YouTube search, content from journalistic channels constituted a large proportion of recommended items (i.e. around 56% for COVID seeds and 52% for Holocaust seeds). Many of the most commonly recommended items again came from the YouTube channels of Germany-based outlets (in particular, Spiegel and ZDF), but compared with Google Search and YouTube search, we observed substantially more content from Swiss channels (e.g. SRF Wissen and SRF Doc). This observation may indicate more localised outputs of the YouTube recommender system compared with the other two, in particular concerning the journalistic content.

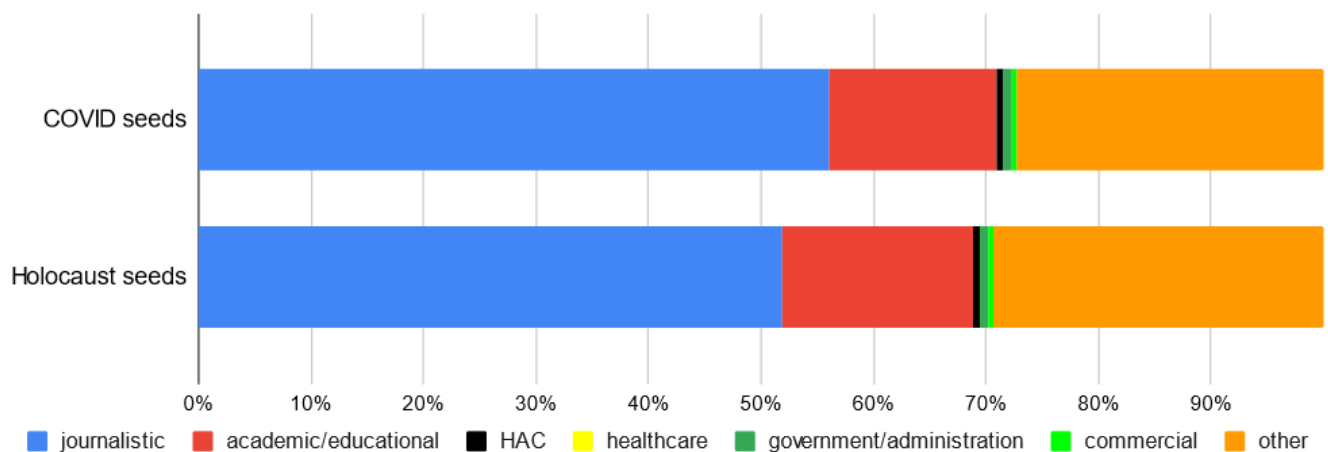


Figure 29. *Aggregated distribution of information sources for all agents across three waves for COVID and Holocaust seeds for YouTube recommendations*

The second-most common source of content (i.e. 27% of first 20 recommender outputs for COVID seeds and 29% for Holocaust seeds) was constituted by channels from the 'other' category, which were less represented in the outputs of search systems. The composition of recommendations

from this category was broad and included content about project management (e.g. from Agile Leipzig channel), IT consultancy (e.g. Computacenter channel), and entertainment (e.g. Best of Bud Spencer & Terence Hill channel). While a number of these channels were from Germany, there were also a number of Anglophone international channels. Finally, academic/educational channels were the third-most common source of recommendations (i.e. 15% of outputs for COVID seeds and 17% of outputs for Holocaust seeds). Unlike Google Search and YouTube search, in the case of YouTube recommendations academic/educational content was primarily coming from German university channels (e.g. Universität Stuttgart and Universität Konstanz channels).

Impact of Algorithmic Personas. Compared with Google and YouTube searches, we observed more differences between sources prioritised by the YouTube recommendation algorithms in relation to algorithmic personas (Figure 30). While for all five personas, journalistic channels tended to be the prevalent sources of recommendations (except wave 2), the distribution of other types of channels varied. For instance, for the COVID seeds, agents associated with the centre-right persona experienced more exposure to commercial channels through YouTube recommendations; the same person was among several personas which were more exposed for commercial recommendations for the Holocaust seeds. Additionally, agents with the centre-right persona were the ones most exposed to HAC channels for the Holocaust seeds.

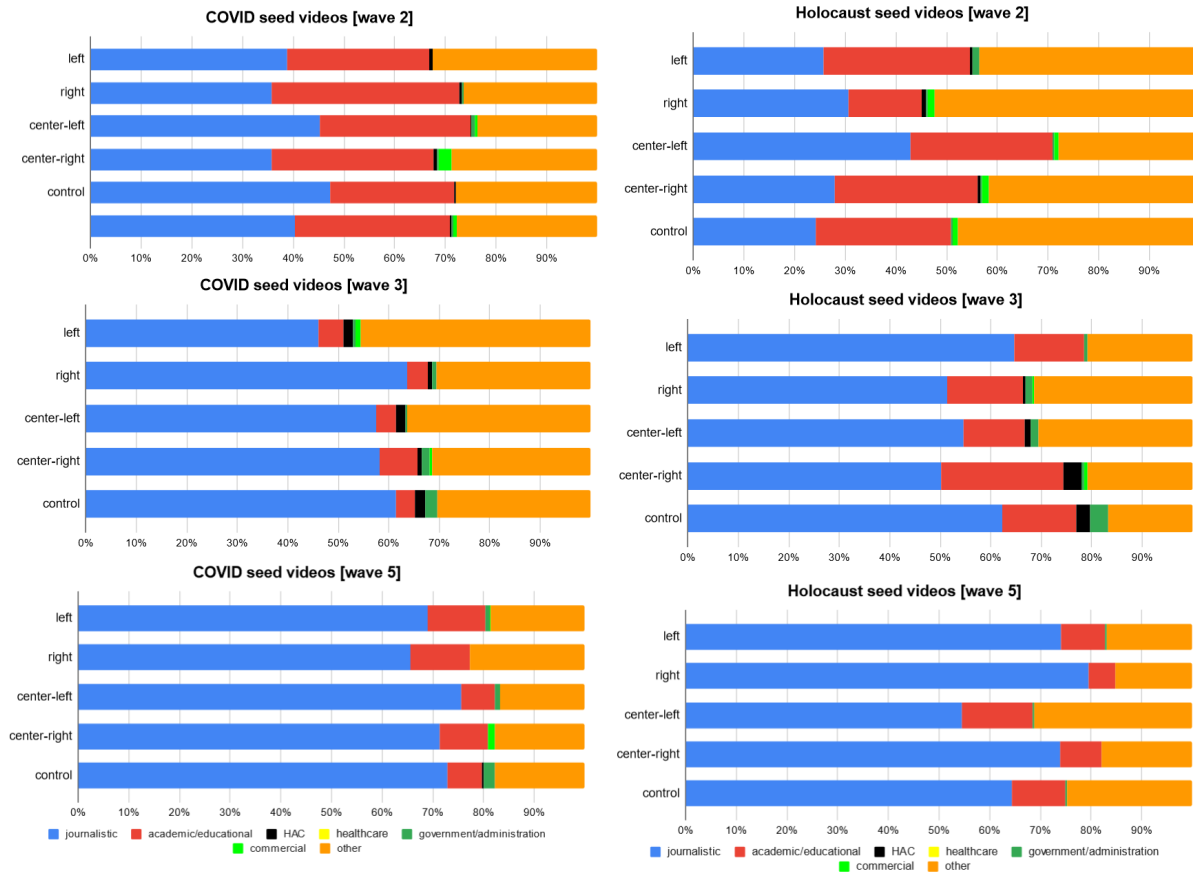


Figure 30. *Distribution of channel types per algorithmic person for YouTube recommendation chains [steps 2-20]*

Besides the content for centre-right persona, we also observed some consistent distinctions for the control persona, where there was the tendency to receive more recommendations from government/administration YouTube channels (e.g. in the case of COVID seed videos, but also occasionally for the Holocaust seeds). However, in most cases, the differences seemed more dependent on the data collection wave: for wave 2, agents received fewer recommendations for journalistic channels and more for academic ones for both the COVID and Holocaust seeds, whereas for the next two waves the distribution changed. Similarly, the number of HAC-related recommendations increased for wave 3 for both seeds, but then these recommendations almost disappeared for wave 5.

Impact of Video Seeds. Figure 31 highlights a number of differences between YouTube search and YouTube recommendation outputs in relation to COVID and Holocaust misinformation.

While in both cases journalistic and academic/educational channels constituted the larger number of system outputs, for YouTube recommendations we observed a substantially higher number of outputs from ‘other’ channels, in particular those related to entertainment (e.g. music) and also consultancy services, in particular the field of IT. In the case of YouTube search, these ‘other’ channels were prominently present only for a few queries, but, in the case of recommendations, they constituted between 20% and 40% of outputs with more content from these channels appearing for the Holocaust seeds.

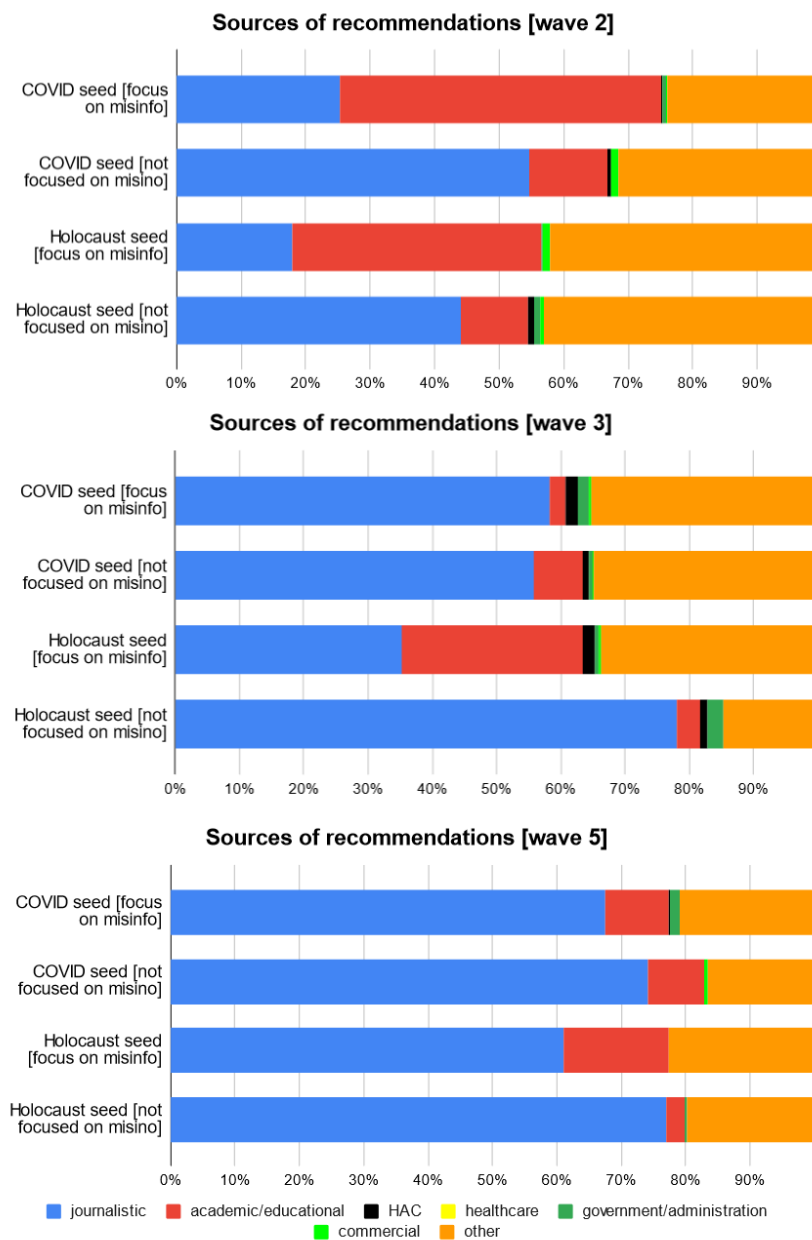


Figure 31. *Distribution of channel types per recommendation seed for YouTube recommendation chains [steps 2-20]*

In terms of the differences between specific seeds, we found that seeds focused on misinformation attracted more recommendations from academic/educational channels that potentially can be caused by the algorithm aiming to prevent users from being exposed to more misinformation-focused content. However, these specific seeds were also more likely to receive recommendations from HAC channels than seeds which did not focus on misinformation, thus highlighting that, to a certain degree, the argument (e.g. Chen et al., 2021) that YouTube recommendation algorithms expose users to more extreme (even if not explicitly containing false claims) content if they start watching it might be still valid.

Impact of Time. Compared with YouTube and Google searches, the impact of the time factor on source composition for YouTube recommendations was substantially more pronounced (Figures 30-31). Both for Holocaust and for COVID seeds in response to specific personas, we observed profound changes in the visibility of journalistic channels, which constituted between 20% and 40% of recommendations for wave 2 and between 53% and 70% for waves 3 and 5. Similarly, there were substantial fluctuations in the number of recommendations from 'other' channels, the amount of which for certain personas (e.g. right and control) for the Holocaust seeds changed from 50% of recommendations for wave 2 to approximately 20% for wave 5.

We likewise observed substantial impact of time for the recommendations associated with individual video seeds. These changes, however, were rather consistent independently of the seed, thus suggesting that changes might have been attributed to the large-scale changes in recommendation relevance (e.g. due to different types of content being more intensively engaged during specific waves). Under these circumstances, during wave 2, we observed more recommendations coming from 'other' and academic/educational channels, whereas for waves 3 and 5 more recommendations originated from journalistic channels. Such volatility, however, sharply contrasted with Google Search and YouTube search, where large-scale changes in output compositions for individual queries occurred rather rarely.

Prioritisation of Misinformation Stances by YouTube Recommendations

General Composition of Recommendation Outputs. A major distinction of YouTube recommendation outputs compared with Google and YouTube search results concerns the extremely low number of recommended items with any stance on misinformation. Figure 32 shows that both for COVID and Holocaust seeds around 99% of recommendations did not mention misinformation (at least within 3 minutes of each YouTube recommended video). The number of outputs which promoted misinformation was negligible for both types of seeds; however, while this amount was substantially lesser than for Google Search and YouTube search, we also observed a much lower amount of content debunking misinformation. While for Google Search we found around 30% of debunking outputs (10% less for YouTube search), for YouTube recommendations we observed less than 1% of content debunking misinformation for Holocaust seeds and almost none for COVID seeds.

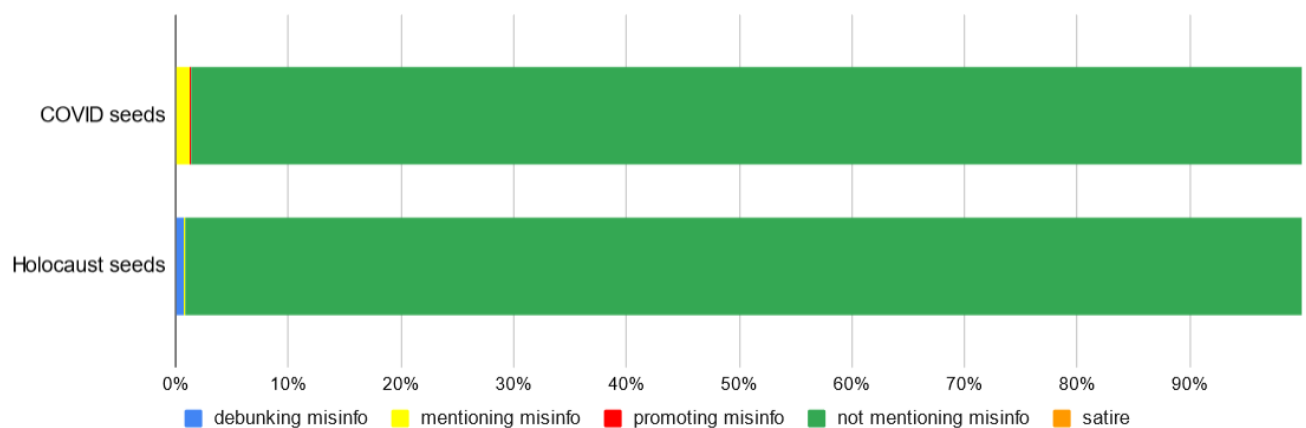


Figure 32. *Aggregated distribution of misinformation stances for all agents across three waves for COVID and Holocaust seeds for YouTube recommendations*

Impact of Algorithmic Personas. Similar to Google and YouTube searches, we observed a limited impact of algorithmic persona for YouTube recommendations. In the case of recommendations (Figure 33), the process of tracing the impacts of personas was complicated by the above-mentioned prevalence of recommendations not mentioning misinformation. However, despite the low number of misinformation-related recommendations, we found that their distribution was rather unequal between personas. Out of five personas, only the centre-left persona (for

Holocaust seeds) received any recommendation promoting or mentioning misinformation. The same persona was receiving more recommendations debunking misinformation for Holocaust seeds. At the same time, the low number of persona-specific recommendations related to misinformation stances as well as the lack of consistency across waves makes it difficult to decisively attribute these variations to the impact of web history on the YouTube recommendation algorithm.

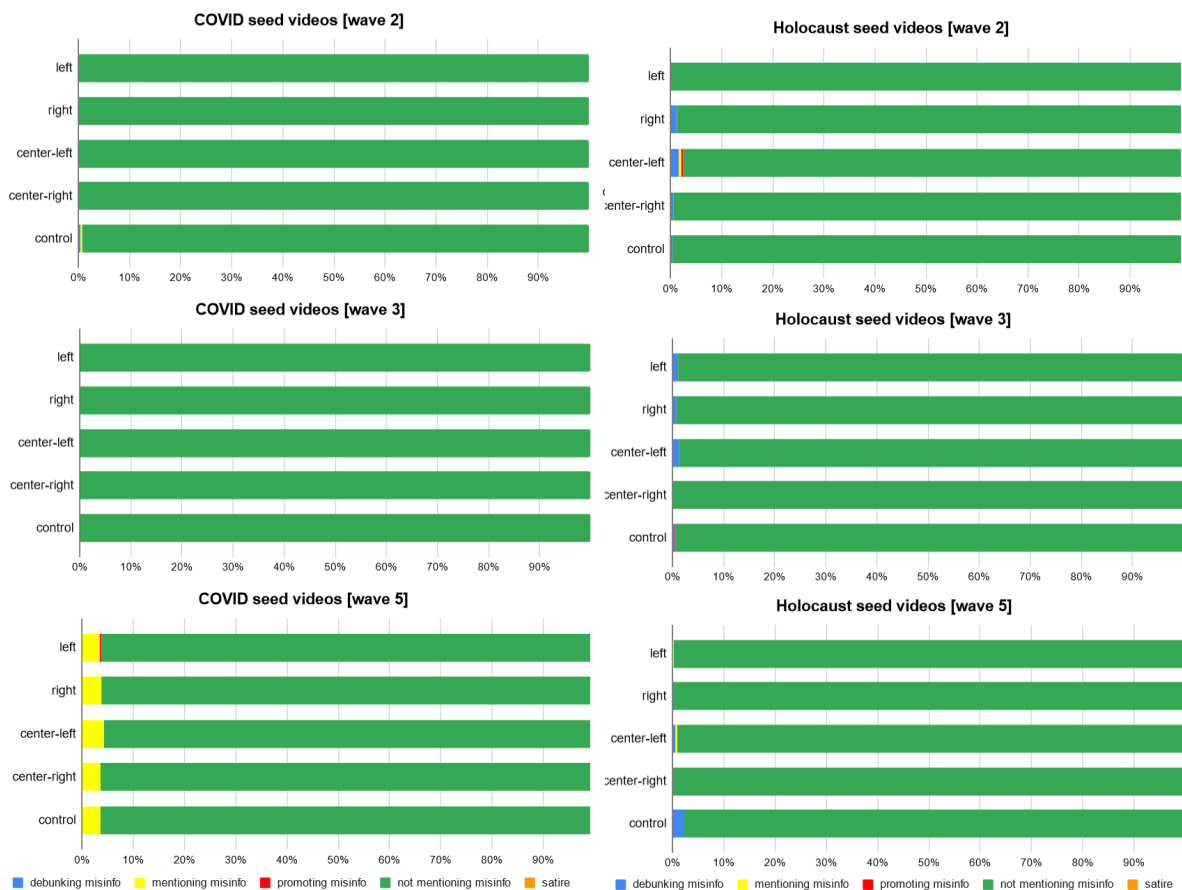


Figure 33. *Distribution of misinformation stances per algorithmic person for YouTube recommendation chains [steps 2-20]*

Impact of Video Seeds. Figure 34 shows little difference in the stance of recommended videos between seed videos focused and not focused on misinformation. For all seeds, the overwhelming majority of recommendations did not mention misinformation. The most pronounced exception was COVID seeds not focused on misinformation, for which around 6% of items mentioned misinformation, albeit for a single wave. In the case of Holocaust seeds, the number of misinformation-related recommendations was slightly higher for the seeds not focused on

misinformation. These observations support findings from other recent audits of YouTube recommendations (e.g. Juneja et al., 2023), which suggest that changes in the YouTube recommendation algorithm decreased the likelihood of it promoting misinformation content and keeping users in “rabbit holes” of misinformation (Brown et al., 2022).

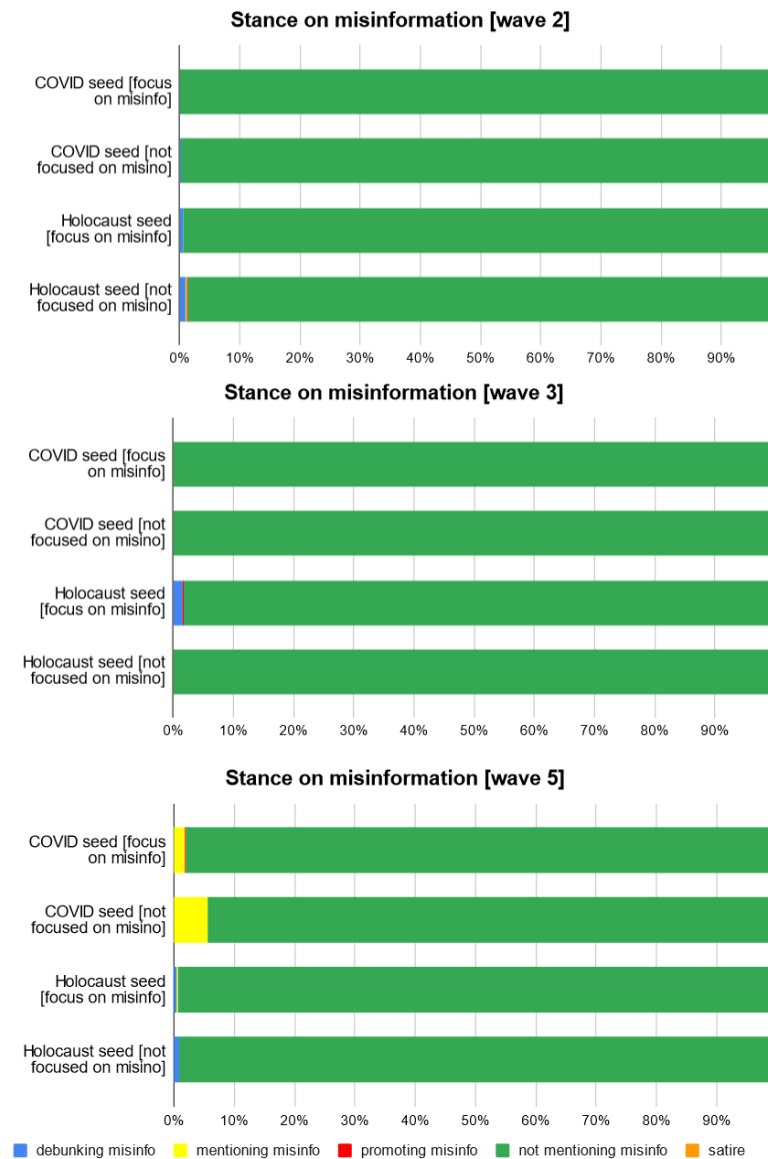


Figure 34. *Distribution of misinformation stances per recommendation seed for YouTube recommendation chains [steps 2-20]*

Impact of Time. While the overall impact of the time factor seems to be low due to the prevalence of content not mentioning misinformation, Figures 33-34 highlights that it had implications for the presence of misinformation-relevant content in YouTube recommendations. In the case of

algorithmic personas, we observed more debunking content for recommendations related to the Holocaust seeds for centre-left and right personas wave 2, but for the later waves the presence of such content for these personas decreased. Similarly, for the centre-left persona for the Holocaust seeds there were more recommendations promoting misinformation, but for the later waves such recommendations disappeared.

In the case of individual seeds, we also observed the impact of time on the exposure to specific misinformation stances. While for waves 2 and 3 (i.e. late May and early June) there was minimal presence of recommendations dealing with misinformation and such presence related primarily to Holocaust-related seeds, for wave 5 we observed an increase in recommendations mentioning misinformation in relation to COVID seeds not focused on misinformation. At the same time, in the case of COVID seeds not focused on misinformation, we observed a rather consistent prioritisation of recommendations not mentioning misinformation across all three waves.

Regression Analysis – Bringing it All Together

Regression Analysis Results for Google Search

Several observations stand out based on the regression analysis for Google Search (Table 4): first, the log-odds of encountering outputs promoting misinformation rather than debunking it was higher in the later waves; however, encountering outputs that did not mention misinformation had lower odds than results debunking misinformation in the later waves.

Table 4. *Output of Multinomial Regression for Google Search*

	Misinformation simply mentioned	Misinformation promoted	No misinformation mentioned
(Intercept)	0.37 ^{***}	0.13 ^{***}	0.64 ^{***}
	(0.05)	(0.08)	(0.03)
Query type: COVID [not focused on misinfo]	1.86 ^{***}	0.61 ^{***}	14.53 ^{***}
	(0.04)	(0.08)	(0.03)
Query type: Holocaust [focused on misinfo]	0.35 ^{***}	0.09 ^{***}	1.17 ^{***}
	(0.04)	(0.08)	(0.02)
Query type: Holocaust [not focused on	1.41 ^{***}	0.00 ^{***}	11.61 ^{***}

	Misinformation simply mentioned	Misinformation promoted	No misinformation mentioned
misinfo]			
	(0.04)	(0.00)	(0.03)
Browser: Firefox	1.01	1.00	1.00
	(0.03)	(0.05)	(0.02)
Persona: right	0.99	0.98	0.99
	(0.04)	(0.07)	(0.03)
Persona: left	1.01	0.99	0.99
	(0.04)	(0.07)	(0.03)
Persona: center-right	0.99	1.01	0.99
	(0.04)	(0.07)	(0.03)
Persona: control	1.00	1.02	1.00
	(0.05)	(0.08)	(0.03)
Wave: 2	1.02	1.03	1.02
	(0.05)	(0.09)	(0.03)
Wave: 3	0.85***	1.56***	0.90***
	(0.05)	(0.09)	(0.03)
Wave: 4	0.99	1.69***	0.92*
	(0.05)	(0.09)	(0.03)
Wave: 5	0.92	1.62***	0.90**
	(0.05)	(0.08)	(0.03)
Wave: 6	0.96	1.48***	0.92**
	(0.05)	(0.09)	(0.03)
AIC	127207.69		
BIC	127595.46		
Log Likelihood	-63561.84		
Deviance	127123.69		
Num. obs.	75564		
K	4		
***p < 0.001; **p < 0.01; *p < 0.05			

Taken together, these observations suggest that the differences in the results encountered in later waves became more substantial in terms of stance on misinformation. Persona and browser type did not have a significant relationship with the likelihood of encountering misinformation. Finally, as COVID queries focused on misinformation were the reference category, the regression output demonstrated that these were the most likely to elicit misinformation-promoting results across all query types; conversely, queries not focused on misinformation for both Holocaust and COVID were the most likely to elicit the results that do not mention misinformation or only mentioned misinformation (as compared to debunking it). Furthermore, Holocaust queries (whether focused on misinformation or not) were less likely to elicit misinformation-promoting results compared to COVID

queries. This observation supports our suggestion that content moderation for Google Search performed better in the case of the Holocaust than for COVID.

Regression Analysis Results for YouTube Search

Table 5 presents the results of a multinomial logistic regression for YouTube search. The reference categories and predictors are identical to Google Search (see above), and the same applies to interpretations of the coefficients.

Several observations stand out based on the analysis: first, the log-odds of encountering results promoting rather than debunking misinformation was lower in the later waves of observation – a finding that is directly opposite to Google Search. Furthermore, encountering results that did not mention misinformation had higher odds than the results debunking misinformation in the later waves – again the direct opposite to Google Search. Neither persona nor browser type had a significant relationship with the likelihood of encountering misinformation, similarly to the results for Google Search.

Table 5. *Output of Multinomial Regression for YouTube Search*

	Misinformation simply mentioned	Misinformation promoted	No misinformation mentioned
(Intercept)	0.39***	0.06***	0.88***
	(0.05)	(0.12)	(0.04)
Query type: COVID [not focused on misinfo]	4.69***	6.35***	17.96***
	(0.05)	(0.09)	(0.04)
Query type: Holocaust [focused on misinfo]	0.24***	0.55***	2.78***
	(0.05)	(0.10)	(0.02)
Query type: Holocaust [not focused on misinfo]	0.00***	0.03***	5.77***
	(0.00)	(0.50)	(0.03)
Browser: Firefox	1.01	0.99	1.01
	(0.03)	(0.07)	(0.02)
Persona: right	1.01	0.99	1.01
	(0.05)	(0.11)	(0.03)
Persona: left	1.01	1.00	1.00
	(0.05)	(0.11)	(0.03)
Persona: center-right	1.00	0.98	1.00
	(0.05)	(0.11)	(0.03)

	Misinformation simply mentioned	Misinformation promoted	No misinformation mentioned
Persona: control	1.02	0.99	1.01
	(0.05)	(0.12)	(0.03)
Wave: 2	0.99	0.96	0.99
	(0.06)	(0.12)	(0.03)
Wave: 3	1.20**	0.91	1.11**
	(0.05)	(0.12)	(0.03)
Wave: 4	1.18**	0.92	1.08*
	(0.05)	(0.12)	(0.03)
Wave: 5	1.21***	0.62***	1.17***
	(0.05)	(0.13)	(0.03)
Wave: 6	1.21***	0.62***	1.17***
	(0.05)	(0.13)	(0.03)
AIC	107272.97 107659.95 -53594.49 107188.97 74151 4		
BIC			
Log Likelihood			
Deviance			
Num. obs.			
K			
*** p < 0.001; ** p < 0.01; * p < 0.05			

Regarding query type, several findings emerge: first, in the case of COVID queries not focused on misinformation, there was a higher likelihood of encountering all types of results as compared to those debunking misinformation. For Holocaust queries, that was not the case: results mentioning misinformation or promoting it were less likely to be encountered than those debunking it; however, results not mentioning misinformation at all were even more likely to appear than those debunking it. This observation suggests that content moderation in relation to the Holocaust worked better than in relation to COVID – a finding similar to the corresponding observation regarding Google Search.

Conclusions

In this section, we provide a short summary of our findings in response to the four RQs. Specifically, we sum up our observations on the composition of system outputs in relation to prioritised sources and misinformation and its variation between individual algorithmic content selection systems as well as the impacts of algorithmic personas, user queries, and time factors on such composition. We then provide an overview of the study implications and possible policy recommendations. related to

increasing awareness of algorithmic misinformation, establishing monitoring infrastructure for tracking changes in algorithmic content selection systems over time, and fostering normative discussion on what the role of these systems should be.

Summary of Findings

Journalistic Content Usually Receives Most Attention From Algorithmic Content Selection Systems

Our findings indicate that for all three algorithmic content selection systems (i.e. Google Search, YouTube search, and YouTube recommender), the largest number of system outputs was from journalistic websites and YouTube channels associated with journalistic media. While there was variation in the composition of sources for individual queries (e.g. more academic/educational sources for Holocaust queries) and across platforms (e.g. more content coming from individuals, such as bloggers, for YouTube search and recommendations), journalistic content tended to be prioritised by both YouTube and Google algorithms, thus making journalistic outlets the primary source of information about both true and false claims in relation to the Holocaust and to COVID. The prioritisation of journalistic content also remained mostly consistent over time, even while for some queries there were certain fluctuations in its visibility.

Google Search and YouTube Search Tend to Prioritise Foreign Content (Particularly for Journalistic and Academic/Educational Sources)

The examination of outputs of algorithmic content selection systems also highlights that for many of the most common categories (e.g. journalistic, academic/educational, and other sources), these systems tend to prioritise foreign content, in particular the one coming from Germany. Such a pattern is particularly pronounced for Google Search and YouTube search, where we observed little content coming from Swiss journalistic media (with the exception of NZZ and Swissinfo) and close to none from Swiss academic/educational institutions. For YouTube recommendations, we observed more content coming from Switzerland-based outlets, thus suggesting more output localisation from this specific system / more successful algorithmic optimisation efforts from Swiss outlets for YouTube

recommendations (as contrasted by, potentially, more successful Google search optimisation strategies from Germany-based journalistic outlets). At the same time, for government/administrative sources, we found a strong prevalence of outputs related to Switzerland.

Low Presence of Content Supporting Misinformation (Except for Some Queries)

For the three systems we audited, the number of outputs explicitly supporting misinformation was low, in particular in response to queries related to the Holocaust. However, the presence of outputs explicitly supporting misinformation – especially in the top search results for YouTube and Google – remains concerning, in particular considering that for some queries such outputs constituted a relatively large portion of results (e.g. up to 40% of outputs for the queries “corona-diktatur” and “corona-lüge ” for Google Search). Additionally, for a number of queries, a substantial number of outputs mentioned misinformation without explicitly debunking it. While per se such outputs are not as problematic as ones explicitly supporting misinformation, exposure to these outputs might contribute to individuals engaging with misinformation less critically than if it were directly debunked.

YouTube Might Be a Less Likely Place for Misinformation Exposure Than Google

Overall, our analysis shows that users might be more likely to be exposed to misinformation via Google Search than YouTube search and, especially, YouTube recommendations. It is important to be careful with interpreting this finding considering the limitations of our analysis (and the general difficulty of reliably detecting misinformation on YouTube due to the often substantial length of videos together with the challenges of audiovisual format). However, it also aligns with recent studies (e.g. Juneja et al., 2023; Ledwich & Zaitsev, 2020) suggesting that YouTube puts substantial effort into culling the spread of misinformation via the platform’s algorithmic content selection system. It also stresses the importance of algorithm audits for identifying problematic behaviour / problematic content and putting pressure on platforms to make changes countering these.

Algorithmic Personas Have Hardly Any Impact on Misinformation Exposure

Our findings from both descriptive and regression analysis show little effect of algorithmic persona on exposure to misinformation-related content via Google and YouTube. While we observed minor variation in outputs for specific personas, the degree of these variations was minimal and could be attributed to the randomisation of system outputs, especially when considering the absence of a consistent pattern across different waves. This observation indicates that algorithmic content selection systems do not necessarily use information about individuals' web history related to browsing activities outside the platforms where these systems are deployed. However, more research is required for a more thorough examination of this assumption, ideally relying on more complex algorithmic personas simulating both within-platform and external browsing activities.

Choice of Query (But Not of Video Seed) Determines Exposure to Misinformation

The key factor which influenced exposure to misinformation-related content for Google and YouTube searches turned out to be the choice of query. We observed that queries focused on misinformation resulted in more content promoting misinformation than those without such a focus, thus indicating that Google and YouTube searches often put more emphasis on satisfying user information needs than on nudging users towards content which might be more likely to debunk misinformation. We also found that COVID-related queries resulted in more content supporting misinformation than Holocaust-related ones, which suggests that it is more difficult for algorithmic systems to deal with recent/emerging misinformation stories. By contrast, for YouTube recommendations, we did not find much evidence for video seeds focused on misinformation resulting in recommendations promoting misinformation, but it is also important to take into consideration that we used a sample of only four seeds (as contrasted by 28 search queries), which might have influenced our observations.

Factor of Time Affects Exposure to Misinformation

Another important factor affecting exposure to misinformation is the time point the algorithmic content selection system was used. While for a number of queries the degree of change over time remained minimal, for some we observed substantial changes between waves which, in a number of cases, resulted in increased exposure to content promoting misinformation.

Implications and Policy Recommendations

Increase Awareness of Algorithmic Misinformation in the Population

The first recommendation that can be provided based on our observations is to put more effort into increasing awareness about potential exposure to misinformation via algorithmic systems, including systems which are often viewed as authoritative (e.g. web search engines). Even while companies, such as Google or Microsoft, are known to put effort in countering the spread of misinformation, our study demonstrates that content supporting or mentioning (but not explicitly debunking) misinformation still appears to be present, in particular for more recent and emerging forms of misinformation. Considering that in the future we expect new misinformation narratives to emerge, it will likely be difficult for companies powering algorithmic content selection systems to immediately accommodate these systems to counter these upcoming narratives. Under these circumstances, it is essential to increase awareness of the general Swiss population about the possibility of being exposed to misinformation via content selection systems and enabling more possibilities for individuals to obtain critical digital literacy skills.

Establish Monitoring Infrastructure to Track How Algorithms Deal With Misinformation Over Time

The second recommendation concerns conducting more longitudinal research on algorithmic systems to acquire better understanding of the role of the time factor in the context of exposure to misinformation via algorithmic content selection systems. The differences in the prevalence of misinformation which we observed over a relatively short period pose a major challenge for auditing the performance of algorithmic systems. First, it makes it harder to detect whether individuals using these systems at a certain time point were more likely to be exposed to misinformation, thus making it more complicated to trace the effects of such exposure. Second, it makes it more difficult to identify whether algorithmic systems might be subjected to malperformance. One possibility to address this challenge is establishing a monitoring infrastructure tracking changes in outputs by algorithmic

systems dealing with selected forms of misinformation. Such monitoring infrastructure can then be used not only by scholars but also the general public in Switzerland.

Foster Normative Discussion on What the Role of Algorithmic Content Selection Systems Should Be

Our final recommendation relates to the necessity of fostering normative discussion on the societal roles of algorithmic content selection systems, in particular in the context of Switzerland. Our observations about the importance of user input for exposure to misinformation highlight that algorithmic systems often retrieve content most fitting to the user's inquiry, even if such content promotes misinformation. While such a "liberal" (Helberger, 2019) model of algorithmic logic can be viewed as the one maximising user agency, it also might have a number of negative societal implications, for example, creating more risk of societal polarisation and the spread of misinformation. It might be important to consider other possible models of algorithmic content selection considering the importance of citizens being properly informed about different societal matters for direct democracies such as Switzerland.

References

- Allington, W. (2017). Holocaust denial online. *Journal of Contemporary Antisemitism*, 1(1), 33-54.
- Bandy, J. (2021). Problematic machine behavior: A systematic literature review of algorithm audits. *Proceedings of the ACM on Human-Computer Interaction*, 5, 1-34.
- Bandy, J., & Diakopoulos, N. (2021). Curating quality? How Twitter's timeline algorithm treats different types of news. *Social Media+ Society*, 7(3), 1-17.
- Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, 35(2), 180-191.
- Bradshaw, S. (2019). Disinformation optimised: gaming search engine algorithms to amplify junk news. *Internet Policy Review*, 8(4), 1-24.

Brennen, J. S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). *Types, sources, and claims of COVID-19 misinformation*. Oxford Internet Institute.

Brown, M. A., Bisbee, J., Lai, A., Bonneau, R., Nagler, J., & Tucker, J. A. (2022). Echo chambers, rabbit holes, and algorithmic bias: How YouTube recommends content to real users. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4114905

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, 8(4), 1-14

Bryant, L. V. (2020). The YouTube algorithm and the alt-right filter bubble. *Open Information Science*, 4(1), 85-90.

Chen, L., Mislove, A., & Wilson, C. (2016). An empirical analysis of algorithmic pricing on amazon marketplace. In *Proceedings of the 25th international conference on World Wide Web* (pp. 1339-1349). ACM.

Chen, A., Nyhan, B., Reifler, J., Robertson, R., & Wilson, C. (2021). Exposure to alternative & extremist content on YouTube. *Anti-Defamation League*. <https://www.adl.org/resources/report/exposure-alternative-extremist-content-youtube>

Chen, Y., & Wang, L. (2022). Misleading political advertising fuels incivility online: A social network analysis of 2020 US presidential election campaign video comments on YouTube. *Computers in Human Behavior*, 131. doi: <https://doi.org/10.1016/j.chb.2022.107202>

Cornia, A. C., Sehl, A., Levy, D. A., & Nielsen, R. K. (2018). *Private sector news, social media distribution, and algorithm change*. Oxford Internet Institute.

Courtois, C., Slechten, L., & Coenen, L. (2018). Challenging Google Search filter bubbles in social and political information: Disconforming evidence from a digital methods case study. *Telematics and Informatics*, 35(7), 2006-2015.

Cuan-Baltazar, J. Y., Muñoz-Perez, M. J., Robledo-Vega, C., Pérez-Zepeda, M. F., & Soto-Vega, E. (2020). Misinformation of COVID-19 on the internet: infodemiology study. *JMIR Public Health and Surveillance*, 6(2). doi: <https://doi.org/10.2196/18444>

Datareportal. (2022). Digital 2022: Switzerland. *Datareportal*. <https://datareportal.com/reports/digital-2022-switzerland>

Donzelli, G., Palomba, G., Federigi, I., Aquino, F., Cioni, L., Verani, M., ... & Lopalco, P. (2018). Misinformation on vaccination: A quantitative analysis of YouTube videos. *Human Vaccines & Immunotherapeutics*, 14(7), 1654-1659.

Elkin, L. E., Pullon, S. R., & Stubbe, M. H. (2020). 'Should I vaccinate my child?' Comparing the displayed stances of vaccine information retrieved from Google, Facebook and YouTube. *Vaccine*, 38(13), 2771-2778.

Epstein, R., & Robertson, R. E. (2015). The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. *Proceedings of the National Academy of Sciences*, 112(33), E4512-E4521.

Gabarron, E., Oyeyemi, S. O., & Wynn, R. (2021). COVID-19-related misinformation on social media: A systematic review. *Bulletin of the World Health Organization*, 99(6), 455-463.

Galindo, F., & Garcia-Marco, J. (2017). Freedom and the Internet: Empowering citizens and addressing the transparency gap in search engines. *European Journal of Law and Technology*, 8(2), 1-18.

Gaufman, E. (2015). World War II 2.0: Digital memory of fascism in Russia in the aftermath of Euromaidan in Ukraine. *Journal of Regional Security*, 10(1), 17-35.

Ghezzi, P., Bannister, P. G., Casino, G., Catalani, A., Goldman, M., Morley, J., ... & Floridi, L. (2020). Online information of vaccines: Information quality, not only privacy, is an ethical responsibility of search engines. *Frontiers in Medicine*, 1-7.

González-Aguilar, J. M., & Makhortykh, M. (2022). Laughing to forget or to remember? Anne Frank memes and mediatization of Holocaust memory. *Media, Culture & Society*, 44(7), 1307-1329.

Goodrow, C. (2021, September 15). On YouTube's recommendation system. *YouTube Official Blog*.

<https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/>

Google Search. (n.d.). How results are automatically generated. *Google*.
<https://www.google.com/search/howsearchworks/how-search-works/ranking-results/>

Google Search Central. (n.d.a). In-depth guide to how Google Search works. *Google*.
<https://developers.google.com/search/docs/fundamentals/how-search-works>

Google Search Central. (n.d.b). A guide to Google Search ranking systems. Google.
<https://developers.google.com/search/docs/appearance/ranking-systems-guide>

Google Search Central. (n.d.c). Introducing a new way for sites to highlight COVID-19 announcements on Google Search. Google.
<https://developers.google.com/search/blog/2020/04/highlight-covid-19-announcements-search>

Guhl, J., & Davey, J. (2020). *Hosting the 'holohoax': A snapshot of holocaust denial across social media*. The Institute for Strategic Dialogue.

Featherstone, J. D., & Zhang, J. (2020). Feeling angry: the effects of vaccine misinformation and refutational messages on negative emotions and vaccination attitude. *Journal of Health Communication*, 25(9), 692-702.

Haim, M., Arendt, F., & Scherr, S. (2017). Abyss or shelter? On the relevance of web search engines' search results when people google for suicide. *Health Communication*, 32(2), 253-258.

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, 6(3), 330-343.

Hamidy, E. (2022). "Krymnash" on YouTube: Formation of Singularities. In *Proceedings of the 9th International Conference on Behavioural and Social Computing (BESC)* (pp. 1-5). IEEE.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). Measuring personalization of web search. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 527-538). ACM.

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 7(8), 993-1012.

Hermann, M. & Krähenbühl, D. (2020, November 27). Parlamentarier-Rating: Im Nationalrat rückt die Mitte nach links – der Ständerat hingegen könnte zum «Bremsklub» werden. NZZ.
<https://www.nzz.ch/schweiz/parlamentarierrating-wohin-sich-die-raete-bewegt-haben-Id.1588933?reduced=true>

Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M., & Watts, D. J. (2021). Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences*, 118(32), 1-8.

Hussein, E., Juneja, P., & Mitra, T. (2020). Measuring misinformation in video search platforms: An audit study on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4, 1-27.

IHRA. (2021). *Understanding Holocaust distortion: Contexts, influences and examples*. IHRA.

Juneja, P., & Mitra, T. (2021). Auditing e-commerce platforms for algorithmically curated vaccine misinformation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-27). ACM.

Juneja, P., Bhuiyan, M. M., & Mitra, T. (2023). Assessing enactment of content regulation policies: A post hoc crowd-sourced audit of election misinformation on YouTube. *arXiv*. <https://doi.org/10.48550/arXiv.2302.07836>

Kaiser, J., Rauchfleisch, A., & Cordova, Y. (2021). Fighting Zika with honey: An analysis of YouTube's video recommendations on Brazilian YouTube. *International Journal of Communication*, 15, 1244-1262.

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd CHI Conference on Human Factors in Computing Systems* (pp. 3819-3828). ACM.

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C., & Mislove, A. (2015). Location, location, location: The impact of geolocation on web search personalization. In *Proceedings of the 2015 Internet Measurement Conference* (pp. 121-127). ACM.

Kulshrestha, J., Eslami, M., Messias, J., Zafar, M. B., Ghosh, S., Gummadi, K. P., & Karahalios, K. (2019). Search bias quantification: investigating political bias in social media and web search. *Information Retrieval Journal*, 22, 188-227.

Kuznetsova, E., & Makhortykh, M. (2023). Blame it on the algorithm? Russian government-sponsored media and algorithmic curation of political information on Facebook. *International Journal of Communication*, 17, 971–992.

Ledwich, M., & Zaitsev, A. (2020). Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. *First Monday*, 25(3). <https://doi.org/10.5210/fm.v25i3.10419>

Leerssen, P. (2020). The Soap Box as a Black Box: Regulating transparency in social media recommender systems. *European Journal of Law and Technology*, 11(2), 1-51.

Lev-On, A. (2008). The democratizing effects of search engine use: On chance exposures and organizational hubs. In A. Spink and M. Zimmer (Eds.), *Web Search: Multidisciplinary Perspectives* (pp. 135-149). Springer.

Leuener, R. (2017, August 16). NZZ Companion: How we successfully developed a personalised news application. *Medium*.

<https://medium.com/@rouven.leuener/nzz-companion-how-we-successfully-developed-a-personalised-news-app-d3c382767025>

Li, H. O. Y., Pastukhova, E., Brandts-Longtin, O., Tan, M. G., & Kirchhof, M. G. (2022). YouTube as a source of misinformation on COVID-19 vaccination: A systematic analysis. *BMJ Global Health*, 7(3), 1-6.

Lynas, M. (2020, April 20). COVID: Top 10 current conspiracy theories. *Alliance for Science*. <https://allianceforscience.org/blog/2020/04/covid-top-10-current-conspiracy-theories/>

Makhortykh, M. (2018). #NoKievNazi: Social media, historical memory and securitization in the Ukraine crisis. In V. Strukov and V. Apyrshchenko (Eds.), *Memory and Securitization in Contemporary Europe* (pp. 219-249). Palgrave Macmillan.

Makhortykh, M., Urman, A., & Ulloa, R. (2020). How search engines disseminate information about COVID-19 and why they should do better. *Harvard Kennedy School Misinformation Review*, 1(1), 1-12.

Makhortykh, M., Urman, A., & Ulloa, R. (2021a). Hey, Google, is it what the Holocaust looked like? Auditing algorithmic curation of visual historical content on Web search engines. *First Monday*, 26(10). doi: <https://doi.org/10.5210/fm.v26i10.11562>

Makhortykh, M., Urman, A., & Ulloa, R. (2021b). Detecting race and gender bias in visual representation of AI on web search engines. In *Advances in Bias and Fairness in Information Retrieval* (pp. 36-50). Springer.

Makhortykh, M., Urman, A., & Ulloa, R. (2022a). Memory, counter-memory and denialism: How search engines circulate information about the Holodomor-related memory wars. *Memory Studies*, 15(6), 1330-1345.

Makhortykh, M., Urman, A., & Wijermars, M. (2022b). A story of (non) compliance, bias, and conspiracies: How Google and Yandex represented Smart Voting during the 2021 parliamentary elections in Russia. *Harvard Kennedy School Misinformation Review*, 3(2), 1-16.

Markmann, S., & Grimme, C. (2021). Is YouTube still a radicalizer? An exploratory study on autoplay and recommendation. In *Disinformation in Open Online Media* (pp. 50-65). Springer.

Mikians, J., Gyarmati, L., Erramilli, V., & Laoutaris, N. (2012). Detecting price and search discrimination on the internet. In *Proceedings of the 11th ACM workshop on Hot Topics in Networks* (pp. 79-84). ACM.

Mittelstadt, B. (2016). Automation, algorithms, and politics| auditing for transparency in content personalization systems. *International Journal of Communication*, 10, 4991–5002.

Möller, J., Helberger, N., & Makhortykh, M. (2019). *Filter bubbles in the Netherlands?* <https://dare.uva.nl/search?identifier=2d8db249-cb3a-4eae-b514-56897c08a2d6>

Noble, S. U. (2018). *Algorithms of oppression*. New York University Press.

Norocel, O. C., & Lewandowski, D. (2023). Google, data voids, and the dynamics of the politics of exclusion. *Big Data & Society*, 10(1), 1-14.

Pamment, J. (2020). *The EU's role in fighting disinformation: Crafting a disinformation framework*. Carnegie Endowment for International Peace.

Paramita, M. L., Orphanou, K., Christoforou, E., Otterbacher, J., & Hopfgartner, F. (2021). Do you see what I see? Images of the COVID-19 pandemic through the lens of Google. *Information Processing & Management*, 58(5), 1-16.

Pradel, F. (2021). Biased representation of politicians in Google and Wikipedia search? The joint effect of party identity, gender identity and elections. *Political Communication*, 38(4), 447-478.

Puschmann, C. (2019). Beyond the bubble: Assessing the diversity of political search results. *Digital Journalism*, 7(6), 824-843.

Ognyanova, K., Lazer, D., Robertson, R. E., & Wilson, C. (2020). Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School Misinformation Review*, 1(4), 1-19.

Otterbacher, J., Bates, J., & Clough, P. (2017). Competent men and warm women: Gender stereotypes and backlash in image search results. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6620-6631). ACM/

Ramponi, M. (2022, December 23). How ChatGPT actually works. *AssemblyAI*.
<https://www.assemblyai.com/blog/how-chatgpt-actually-works/>

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira Jr, W. (2020). Auditing radicalization pathways on YouTube. In *Proceedings of the 2020 Conference on fairness, Accountability, and Transparency* (pp. 131-141). ACM.

Riley, J. K. (2022). Angry enough to riot: An analysis of in-group membership, misinformation, and violent rhetoric on TheDonald.win between election day and inauguration. *Social Media+ Society*, 8(2), 1-12.

Statcounter. (2023). Search Engine Market Share Switzerland. *Statcounter*.
<https://gs.statcounter.com/search-engine-market-share/all/switzerland>

Steiner, M., Magin, M., Stark, B., & Geiß, S. (2022). Seek and you shall find? A content analysis on the diversity of five search engines' results on political queries. *Information, Communication & Society*, 25(2), 217-241.

Srba, I., Moro, R., Tomlein, M., Pecher, B., Simko, J., Stefancova, E., ... & Bielikova, M. (2023). Auditing YouTube's recommendation algorithm for misinformation filter bubbles. *ACM Transactions on Recommender Systems*, 1(1), 1-33.

Trielli, D., & Diakopoulos, N. (2022). Partisan search behavior and Google results in the 2018 US midterm elections. *Information, Communication & Society*, 25(1), 145-161.

Tripodi, F. B. (2022). *The propagandists' playbook: How conservative elites manipulate search and threaten democracy*. Yale University Press.

Tufekci, Z. (2018). YouTube, the great radicalizer. *The New York Times*, 10(3), 2018.

Vincent, N., Johnson, I., Sheehan, P., & Hecht, B. (2019). Measuring the importance of user-generated content to search engines. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 505-516). AAAI.

Vlasceanu, M., & Amodio, D. M. (2022). Propagation of societal gender inequality by internet search algorithms. *Proceedings of the National Academy of Sciences*, 119(29), 1-8.

Unkel, J., & Haim, M. (2021). Googling politics: Parties, sources, and issue ownerships on Google in the 2017 German federal election campaign. *Social Science Computer Review*, 39(5), 844-861.

Ulloa, R., Makhortykh, M., & Urman, A. (2022). Scaling up search engine audits: Practical insights for algorithm auditing. *Journal of Information Science* (online first). doi: <https://doi.org/10.1177/01655515221093029>

Ulloa, R., Richter, A. C., Makhortykh, M., Urman, A., & Kacperski, C. S. (2022b). Representativeness and face-ism: Gender bias in image search. *New Media & Society* (online first). doi: <https://doi.org/10.1177/14614448221100699>

Urman, A., & Makhortykh, M. (2022). "Foreign beauties want to meet you": The sexualization of women in Google's organic and sponsored text search results. *New Media & Society* (online first). doi: <https://doi.org/10.1177/14614448221099536>

Urman, A., Makhortykh, M., & Ulloa, R. (2021). Auditing source diversity bias in video search results using virtual agents. In *Companion Proceedings of the Web Conference 2021* (pp. 232-236). ACM.

Urman, A., Makhortykh, M., & Ulloa, R. (2022a). The matter of chance: auditing web search results related to the 2020 US presidential primary elections across six search engines. *Social Science Computer Review*, 40(5), 1323-1339.

Urman, A., Makhortykh, M., Ulloa, R., & Kulshrestha, J. (2022b). Where the earth is flat and 9/11 is an inside job: A comparative algorithm audit of conspiratorial information in web search results. *Telematics and Informatics*, 72, 1-15.

Urman, A., Makhortykh, M., & Ulloa, R. (2022c). Auditing the representation of migrants in image web search results. *Humanities and Social Sciences Communications*, 9(1), 1-16.

Weber, M. S., & Kosterich, A. (2018). Coding the news: The role of computer code in filtering and distributing news. *Digital Journalism*, 6(3), 310-329.

Williams, E. M., & Carley, K. M. (2023). Search engine manipulation to spread pro-Kremlin propaganda. *Harvard Kennedy School Misinformation Review*, 4(1): 1-13.

Yesilada, M., & Lewandowsky, S. (2022). Systematic review: YouTube recommendations and problematic content. *Internet Policy Review*, 11(1), 1-22.

YouTube. (n.d.a). YouTube Search. *YouTube*.

<https://www.youtube.com/howyoutubeworks/product-features/search/>

YouTube. (n.d.b). Recommended videos . *YouTube*.

<https://www.youtube.com/howyoutubeworks/product-features/recommendations/>

Zavadski, A., & Toepfl, F. (2019). Querying the Internet as a mnemonic practice: how search engines mediate four types of past events in Russia. *Media, Culture & Society*, 41(1), 21-37.

Zuiderveen Borgesius, F., Trilling, D., Möller, J., Bodó, B., De Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Review*, 5(1). 1-16.

Zweig, K. (2017, April 7). Watching the watchers: Epstein and Robertson's "Search Engine Manipulation Effect". *Algorithm Watch*. <https://algorithmwatch.org/en/watching-the-watchers-epstein-and-robertsons-search-engine-manipulation-effect/>

Appendix A1: Codebook for Manual Content Labelling

Descriptions of coding variables

- **ACCESS: Is the item available (i.e. can the link be opened)?**
 - **Y==Yes**
 - **N==No.** Do only use this option if neither the WayBackMachine nor Google cache or Swissdrex succeed at retrieving the content.

- **RELATED: Is the item related to the COVID/Holocaust?**
 - **C==Covid**
 - **H==Holocaust** ("Holocaust" here also includes items related to Hitler, Nazi Germany, and its modern day expression)
 - **N==neither.** If you chose N, you still have to do the coding for the SOURCE. COVID_MISINFO and HOLOCAUST_MISINFO will be 0, STANCE will be 9.
 - **99==Unclear** (i.e. information is not in German or English). If you chose 99, leave COVID_MISINFO, HOLOCAUST_MISINFO, and STANCE empty. If you can code SOURCE, please do so.

- **SOURCE: What is the item's source type [text pages]?**
 - **Type of the domain for the Google search:**
 - **1==reference website** (e.g. Wikipedia or IMDB). This category includes websites publishing statistics or datasets that are not directly controlled by governments (code 7), media outlets (code 2) or universities (code 3).
 - **2==journalistic media** (e.g. NZZ or a news aggregator such as Google News);
 - **3==academic/educational site** (e.g. UniBern or Yad Vashem -> institution attached to it). This includes educational campaigns by NGOs. If the organisation has a strong political leaning, you should consider labelling it as code 5.

- **4**==social media (e.g. Facebook or YouTube). If YouTube, then mark the type of the channel (see below);
- **5**==HAC media. This category includes right- and left-wing hyperpartisan websites promoting explicitly ideologically biased agenda / alternative websites that claim to be a replacement to mainstream news / conspiracy websites that promote common conspiracy theories (e.g. qanon.pub);
- **6**==healthcare related website. This includes healthcare institutions like doctors / hospitals, health advice provided by pharmaceutical companies or healthcare networks and magazines associated with health networks. It does not include health-related general interest magazines (e.g. Psychologie Heute, GEO Gesundheit): these outlets belong to code 2. If the website or video simply explains scientific research without providing general health information, then it likely belongs to code 3.
- **7**==government and administration (e.g. ministry of healthcare), including Robert Koch Institute (RKI). URL suffixes which indicate a government website (e.g. .gov) are a strong indicator for code 7.
- **8**==commercial and (non-media) company websites (e.g. amazon.com)
- **9**==other (non-conspiracy) websites (e.g. individual blog, comedy/satire, lifestyle). Please specify in the type in comments.
- **Type of the YouTube channel for YouTube (for hints, see also equivalent above)**
– make sure to click on the link to the channel website and check “Kanalinfo” there:
 - **2**==journalistic channel (e.g. of NZZ). If the channel belongs to an individual journalist, please code as 9. If it is a journalistic collective, code as 2.
 - **3**==academic/educational site (e.g. of Yad Vashem). This includes educational campaigns by NGOs. If the organisation has a strong political leaning, you should consider labelling it as code 5 instead.

- **5==HAC media.** This category includes right- and left-wing hyperpartisan channels promoting explicitly ideologically biased agenda / alternative channels that claim to be a replacement to mainstream news / conspiracy channels that promote common conspiracy theories (e.g. pro-QAnon channel);
 - **6==healthcare related channels** (healthcare institutions like doctors/hospitals) and channels associated with health networks. If the channel simply explains scientific research without providing general health information, then it probably shall be coded as 3;
 - **7==government and administration** (e.g. ministry of the healthcare), including Robert Koch Institute (RKI) and Landes- und Bundeszentralen;
 - **8==commercial and (non-media) channels;**
 - **9==other:** channels of individual vloggers, individual journalists, comedy/satire, lifestyle, etc. Please specify in the comments.
- **COVID_MISINFO: To what COVID misinformation does the item refer to?** Type of misinformation [if multiple types are present, then list them separated by commas]. Do only code explicitly mentioned or debunked misinformation.
 - **0==no misinformation present;**
 - **1==5G spreads COVID or is in fact responsible for COVID deaths;**
 - **2==COVID is a form of bioweapon;**
 - **3==COVID was made in the lab;**
 - **4==Big Pharma, corporations or other economic elites conspire to use COVID for their economic gain;**
 - **5==the severity of COVID is much lower than it is claimed.** Drawing the line here is not always that easy, but analyses that point out that more recent variants cause less severe symptoms or are less likely to overwhelm the health system because of high vaccination rates (e.g. <https://www.mdr.de/wissen/covid-bald-harmloser-als-grippe->

[100.html](#)) should not be coded as 5. Broader claims, such as Covid being “eine für die Masse der Bevölkerung ungefährliche Krankheit”, qualify for code 5;

- **6==**COVID is used by the governments/secret powers/powerful individuals to install dictatorship/control over people;
- **7==**the COVID is a hoax/does not actually exist;
- **8==**(long-term) side effects of COVID-19 vaccines, in particular claims that these vaccines can cause infertility, change DNA, do not protect from severe illness, spread the virus itself, or are used to insert microchips. Just mentioning or worrying about the possibility of (long-term) side effects without specifying them does not justify code 8, because the vaccines do in fact have side effects and we can not say that there is a scientific consensus that COVID vaccines would not have any side effects in the long term.
- **9==**other misinformation [add a comment describing the misinformation in 3-5 words]

● **HOLOCAUST_MISINFO: To what Holocaust misinformation does the item refer to?**

Type of misinformation [if multiple types are present, then list them in ascending order separated by commas (without space)]. Do only code explicitly mentioned or debunked misinformation:

- **0==**no misinformation present;
- **1==**the number of Holocaust victims (i.e. 5-6 millions) is inflated;
- **2==**the majority of Germans, Nazis or Hitler himself were not aware of the Holocaust;
- **3==**present-day Jews profit from the Holocaust;
- **4==**overall, Hitler/Nazis did more good than bad things: this category includes articles or videos that focus on the positive aspects of Nazi Germany, praise Hitler and downplay or ignore any atrocities committed during the National Socialism era.
- **5==**gas chambers did not exist/were fake/were just used for delousing;
- **6==**the Holocaust is a lie or never happened;

- **7**==misleading comparisons: other atrocities are worse than the Holocaust (e.g., bombing of Dresden or the climate change/abortions); the British concentration camps were equivalent to the Nazi concentration camps.
 - **8**==justifications for the Holocaust, such as painting the Jews as an inherently evil race
 - **9**==other misinformation [add a comment describing the misinformation in 3-5 words]
- **STANCE: What is the stance of the item on the COVID-/Holocaust-related misinformation?**
 - **1**==promotes misinformation (e.g., gas chambers never existed or the Swiss government uses COVID to create a dictatorship);
 - **0**==misinformation is mentioned but simply described - i.e. neither debunked, nor promoted (e.g. some people claim that gas chambers never existed or the Swiss government uses COVID to create a dictatorship). Also includes cases, where both sides of the argument are presented equally.
 - **-1**==explicitly debunks misinformation (e.g. statements that claims about the gas chambers not existing are fake or that the Swiss government does not use COVID to create a dictatorship despite some people arguing for that). This implies a clear signal that the misinformation statement is false or at least that the reader should doubt the validity of the claim. It can take the form of the explicit counterclaim (e.g. this claim is false/this claim is incorrect/this claim is a lie) or a more implicit statement expressing doubt, such as by adding “angeblich” or by denoting a claim as “Verschwörungstheorie”. Note that by this standard, articles or videos taking a relatively “neutral” stance may only barely qualify for a -1. An example is this Wikipedia entry: https://de.wikipedia.org/wiki/Politisches_Testament_Adolf_Hitlers. Only the words in italic prevent this article from being classified as an article simply describing misinformation: “Darin gab er einem *angeblichen* Weltjudentum die Schuld am Zweiten Weltkrieg, um die von ihm betriebene Vernichtung der europäischen

Juden zu rechtfertigen. [...] Stattdessen wies er dem *fiktiven* "Weltjudentum" die Kriegsschuld zu".

- **9**==information on the topic of COVID/the Holocaust is presented, but misinformation is not explicitly mentioned or debunked. If there is no misinformation present (HOLOCAUST_MISINFO==0, resp. COVID_MISINFO==0) then this should be coded as 9 as well.
 - **99**==satire or not applicable in other ways
- **COMMENT:** use this column to indicate any doubts you have about your own coding (in particular about the STANCE). This could be due to you watching only part of a video but suspecting it might contain something dubious later on, or because the codebook instructions do not fit the particular case very well or just generally flag an item for our review because it strikes you as suspicious.