

Data Donation as a Method for Investigating Trends and Challenges in Digital Media Landscapes at National Scale

The Danish Population's
Use of YouTube as an
Illustrative Case

ANJA BECHMANN
MIRIAM KROMAN BREMS
MARCUS K. OLESEN
JESSICA G. WALTER
DAVID WEGMANN

NORDICOM

D T L A B
CENTER FOR DIGITAL SOCIAL RESEARCH



DANMARKS FRIE
FORSKNINGSFOND
INDEPENDENT RESEARCH
FUND DENMARK



Nordicom is funded with
support from the
Nordic Council of Ministers

This publication was produced by DATALAB – Center for Digital Social Research at Aarhus University, and published in collaboration with Nordicom at the University of Gothenburg.

Data Donation as a Method for Investigating Trends and Challenges in Digital Media Landscapes at National Scale

The Danish Population's
Use of YouTube as an
Illustrative Case

Anja Bechmann, Miriam Kroman Brems,
Marcus K. Olesen, Jessica G. Walter,
& David Wegmann

© The authors

ISBN: 978-91-89864-18-4

DOI: 10.48335/9789189864184

Nordicom

University of Gothenburg

PO Box 713 SE-405 30

Gothenburg

Cover: Karin Andén

Cover Photo: Scandinav bildbyrå

Cite this work:

Bechmann, A., Brems, M. K., Olesen, M. K., Walter, J. G., & Wegmann, D. (2025) *Data donation as a method for investigating trends and challenges in digital media landscapes at national scale: The Danish population's use of YouTube as an illustrative case*. Nordicom, University of Gothenburg.
<https://doi.org/10.48335/9789189864184>

Acknowledgements

The data collection was commissioned to and executed by the Danish survey institute Epinion. We thank, in particular, Thorkil Klint and Trine Møller from Epinion for co-creating the data collection setup, for adjustments to our tests throughout a year before launch, and for executing this novel approach to nation-wide data donation even though it was untested ground. We also thank doctoral student Anton Elias Holt for helping us with the ethical approval of the study and student assistant Julie Kærgaard Poulsen for providing us with screen dumps for the Appendix of this publication.

Funding

The study is part of the project SOCIAL MEDIA INFLUENCE (2023–2028), funded by the Independent Research Fund Denmark (grant number 10.46540/3097-00188B). Furthermore, the study has received SHAPE strategic funding from Aarhus University.

Table of contents

0. Foreword.....	5
1. Executive summary	6
2. Introduction	9
2.1 Changing digital media landscapes and data access	9
2.2 Objectives and contribution to the field	12
2.3 Structure of the publication.....	13
3. Overview of the data collection steps	15
4. Recruitment.....	17
4.1 Panel description.....	17
4.2 Incentives	18
4.3 Invitation material.....	18
4.4 Reminders	18
4.5 Completions over time.....	19
4.6 Target population	20
5. Survey 1: Screening participants into the study.....	23
5.1 Screening based on YouTube use	24
5.2 Screening based on sociodemographic profile	25
5.3 Screening based on consent to donating YouTube data	26
5.4 Screening based on availability of stored YouTube data	28
5.5 Drop-off during completion of the survey.....	31
6. Survey 2: Guiding participants to request their data	32
6.1 Step-by-step guide to request data	32
6.2 Drop-off in Survey 2	36
7. Survey 3: Guiding participants to download and securely donate their data.....	39
7.1 Applied data validation criteria	40
7.2 Drop-off in Survey 3	43
8. Survey 4: Demographic distribution and duration.....	45
8.1 Completion times.....	46
9. The donated data files.....	49
9.1 Sociodemographic profiles and the donated data	52
9.2 Viewing data.....	55
9.3 Search data.....	61
9.4 Subscription data	63
9.5 Comment data	64
10. GDPR registration and ethical reflections	66

11. Discussion and conclusion	67
12. References	72
13. Appendix: Screenshots of Surveys 1–4.....	78
13.1 Survey 1	81
13.2 Survey 2.....	93
13.3 Survey 3.....	99
13.4 Survey 4.....	108

0. Foreword

As social media continues to play an increasingly central role in everyday life and public discourse, so too does the need to understand the mechanisms through which platforms influence citizens' attitudes, values, and engagement in democratic processes. This report presents a novel methodological approach to studying these dynamics, with a particular focus on the Nordic context.

The purpose of this publication is to provide a detailed documentation of a case study that explores the potential of data donation as a relatively new method for collecting digital trace data. The study investigates to what extent this method can be scaled to national populations, and how it may help generate deeper insights into digital media usage and behaviour. Using Denmark as an illustrative case, and YouTube as the platform of interest, the study offers valuable reflections on the challenges and opportunities of this emerging methodological field.

The study has been designed and executed by members of DATALAB – Center for Digital Social Research, which is an interdisciplinary research centre at Aarhus University in Denmark, led by Professor Anja Bechmann. Through careful documentation and critical discussion of the data collection process, methodological design, and the quality of the donated data, this publication seeks to inspire other researchers in the broader field of digital communication.

The overall aim of Nordicom is to stimulate and promote research-based knowledge about media and communication in the Nordic region and to strengthen cooperation between Nordic media researchers. The opportunity to publish and disseminate this work in collaboration with DATALAB fits very well within that aim. By sharing insights from this case, we aim to support future methodological innovation and comparative research efforts across the Nordic region.

Ultimately, my colleagues and I hope that this collective endeavour will contribute to a more nuanced and empirically grounded understanding of digital media usage, behaviours, data infrastructures, and platform influence on democracy. We extend our sincere thanks to DATALAB, whose work has made this publication possible, and we look forward to the interest and dialogue it may generate among scholars across the Nordic countries.

Gothenburg, April 2025
Jonas Ohlsson
Director, Nordicom

1. Executive summary

This publication demonstrates how data donation can be scaled to national population level to understand trends in digital media usage and potential challenges for digital democracies. We describe and discuss the results of our novel methodological setup combining data donation and surveys about the Danish population's use of YouTube as an illustrative case.

Through three consecutive survey steps, the study collected trace data from participants using quotas for gender, age, and education to approximate a sample representative of the adult Danish population of YouTube users, resulting in a final donated data sample of $n = 1,063$. This sample size is much larger than other similar data donation studies conducted. In a fourth step, these data were linked with a comprehensive post-survey including questions about YouTube use, social media use, news use, mental health, personality traits, conspiracy mentality, political views, political interest, political efficacy, and more.

The publication documents the data collection procedure and analyses response rates, screening out and drop-off at different steps of the data collection, and how well the final sample matches the adult Danish population of YouTube users (as estimated in a prior representative survey study). Furthermore, we provide a descriptive statistical overview of the donated data and discuss ethical considerations and clearance of the study. The methodological design was developed by the research team. The Danish survey institute Epinion was responsible for conducting the data collection and technical implementation. The key takeaways from the data collection process in the different steps are as follows:

Feasibility: The study demonstrated that data donation is scalable, as it succeeded in collecting 1,063 valid data donation files from participants using quotas to approximate a sample mirroring the adult Danish YouTube population (as estimated by a prior representative survey study). However, achieving representativeness proved challenging, as summarised below.

Study context: The study was conducted in Denmark, which might impact the success of scaling data. Like the rest of the Nordic countries, the Danish media context is characterised by high social media penetration and high trust in societal institutions. This means that a large share of the general population belongs to the target group, and there might be a higher willingness to donate.

Study design: That the study succeeded in scaling data may also be impacted by the study design. Participants were asked to donate data first and were only allowed to proceed to the post-survey if they donated valid data. This ensured that participants unwilling to donate data were screened out early and that participants unable to successfully donate were screened out after the donation steps.

Fieldwork: The fieldwork period was slightly longer than three months, including a pilot phase of approximately one month. Following the pilot phase, criteria for how much YouTube data respondents had to donate to be eligible to participate were eased, which led to increased completion rates. This highlights that data density must be weighed against project feasibility.

Participation duration: For respondents who completed all steps of the study, the median completion time was 2.3 days. Studying the quickest 95 per cent only (not accounting for outliers) reveals that there are no gender, age, or educational differences in duration for participating in the study (starting Survey 1 and completing Survey 4).

Sample representativeness: Regarding the representativeness of the final sample, comparison to a study of self-reported YouTube use among the adult Danish population conducted by Epinion shows that the oldest age group (55+) is underrepresented by 11 percentage points in our final sample. This is the biggest deviance when looking at sociodemographic indicators (age, gender, education).

Willingness to donate data: Close to half (49.2%) of respondents who initially expressed interest in participating and were eligible to do so declined to participate when informed that they must donate their data. It is especially the oldest age group (55+) who declined, indicating that lacking confidence in their own technical skills may have played a central role. However, there was also a sizeable drop-off among the younger age groups following information about the required data donation. This may install skewness in the final sample, as respondents willing to donate might be qualitatively different from those unwilling to donate.

Availability of data: To participate, respondents must have had a Google account and have allowed their YouTube data to be stored on their account (otherwise, they had no data to donate). A sizeable share of respondents otherwise eligible and willing to participate were screened out due to these requirements. A central limitation to the method is thus that it cannot capture YouTube usage patterns from users who do not have a Google account or do not allow data storage.

Drop-off in the technical steps: The oldest age group (55+) was more likely to drop off during the technical steps of the data donation process, where respondents had to request, download, and donate their data. This means that older participants must be oversampled to achieve a final sample representative of the user group and, even then, the older participants who complete the full survey might better represent the more technically skilled older population. While differences between education groups were observed, there was no systematic pattern where lower or higher education is consistently associated with drop-off in the technical steps.

Data validation: Donated data was validated by Epinion based on predefined criteria. Of those participants that uploaded data but failed data validation ($n = 496$), most respondents failed because they had not watched enough videos during the last year (23%) or had no data entry reaching back to before 1 January 2021(36%). These requirements were enforced to ensure that the donated files contained enough data and to allow cross-event analyses over time. The third most prevalent reason for failing validation was that respondents had not uploaded a watch-history file or the wrong file format for the watch history (21%). The remainder of participants who failed validation failed for other reasons.

(In)completeness of the donated data: Analysis of the data available in the donated files emphasises that data donation – even when applied in the context of famously data-hungry platforms such as YouTube – necessarily provides incomplete data. First, only data that YouTube records, stores, and makes accessible to the user via the logfiles can be obtained: As these logfiles only contain timestamps for the beginning of

video-playback, but not when or if playback was terminated, the data does not yield which videos participants have seen, but rather for which ones playback was initiated. Furthermore, participants have some agency over the data that is being stored about them by changing the data collection settings or retroactively deleting previously collected data from their account. Although data donations might currently be the best way of obtaining insights into user behaviour on platforms, researchers must be aware of these caveats when making conclusions based on logfile data.

Recency bias in the donated data: Analysis of the donated data shows that the further back in time we go, the fewer respondents have available data, and the available data itself becomes increasingly sparse. This must be considered if making analysis of YouTube usage back in time, as analyses of older data would, consequently, draw on less data from a thinner pool of participants. Conversely, analyses of more recent usage patterns can be based on much more data from a larger sample of participants.

Types of data: Of the participants who completed the study, all donated viewing data (as this was a requirement), 96.3 per cent donated search data, 90.4 per cent donated subscription data, and 65.0 per cent donated comment data. The latter suggests that more than half of the participants have, at some point, engaged in discussions on the platform. This is interesting since YouTube is often characterised as a pure streaming service rather than a social media platform.

Niche consumption patterns: The combined donated watch-history files contained 15 million videos. Filtering this sample for duplicates yielded nearly 7 million unique videos. The most viral videos in the sample were watched by less than a third of the participants. Both patterns in the sample point to niche consumption with a long tail of videos watched by few people.

International and national content: Although data was collected in Denmark, English-language content dominates among the watched videos in the dataset: Computational audio analysis of a random sample of 10,956 videos across participants showed that 79.1 per cent of the videos are in English and only 16.3 per cent are in Danish. At the same time, however, analysis of the channels that most donating participants subscribed to suggests that there is also an interest in nationally produced content, for example, from Danish influencers.

Dominance of platform-native infotainment content: Analysis of the list of channels most participants subscribe to show a dominance of international and national influencers and infotainment YouTube channels. Only one of the twenty most subscribed channels is run by a traditional Danish media organisation. These findings suggest that influencers play an important role on YouTube in terms of what users actively choose to follow and that platform-native infotainment content is dominant, whereas traditional Danish media have a comparatively small footprint.

2. Introduction

The purpose of this publication is to provide a detailed documentation of a case study employing a novel methodological setup to explore to what extent data donation as a relatively new method for collecting digital trace data can be scaled to national populations and thus provide increased knowledge of digital media usage and behaviour. We use Denmark as an illustrative case and YouTube as an object of interest due to its position as one of the largest social media platforms and streaming services in the Nordic countries and worldwide, especially for news (Newman et al., 2024). Through careful documentation and discussion of the data collection, methodological setup, and quality of the donated data, we hope to inspire other researchers within the broader field of digital communication to build on our experiences in the pursuit of future methodological developments. This collective endeavour will over time hopefully increase our knowledge of digital media, usage, behaviours, data infrastructures, and platform influence on democracy.

2.1 Changing digital media landscapes and data access

In an increasingly digital media landscape, the Nordic countries are among the populations in the world with the highest social media platform reach (Danmarks Statistik, 2023; Eurostat, 2024; NordMedia Network, 2024; Statista, 2024). The portion of people not using social media has decreased to a minimum and platforms are increasingly used for almost all aspects of life – socialising, work, entertainment, information, and news-seeking (Kulturministeriet, 2021). Globally, video content on platforms such as YouTube, TikTok, and Instagram Reels is on the rise, and YouTube is used for news by 31 per cent across the populations covered in the *Reuters Institute Digital News Report*. Influencers on the platforms thus become potentially more central for these parts of the populations' information diet (Newman et al., 2024).

Social media platforms therefore constitute an obvious setting for studying digital media usage and behaviour as it unfolds in different degrees of harmony or disorder – from mobilising people for good causes and against regimes to inauthentic and malicious coordinated influence operations (Bennett & Livingston, 2018; Bradshaw & Howard, 2018; González-Bailón, 2017; Howard, 2020; Mitchell et al., 2019; Buning, et al., 2018; Quandt, 2018, 2021; Wardle & Derakhshan, 2017).

In line with ideologies from the early Internet history (Abbate, 2000; Turner, 2008), the idea behind platforms such as YouTube and Facebook was to connect people and promote free speech and access to information for all people worldwide (see, e.g., Kaplan, 2025), albeit platforms have later constructed moderation regimes to meet demands from actors like users, regulators, and advertisers (Gillespie, 2018). Platforms have played a significant role in mobilising people, for instance, against regimes (Bennett et al., 2014) or against people in power, in the case of the Occupy Wall Street and Black Lives Matters movements (Freelon et al., 2018), but they have also acted as tools for regimes to conduct genocide in Myanmar (Warofka, 2018; Zaleznik, 2021).

At no time in communication history have people left so many recorded traces behind of communicative behaviours that in principle allows for the decoding of not only

communication logics and dynamics but also the social fabric in general (González-Bailón, 2017; Lazer et al., 2021). At the same time, however, most of the platforms are not based within the legal jurisdiction of the Nordic countries, and the European Union has fought for almost 20 years to increase the regulation of primarily American- (but also Chinese-) owned platforms, with a special emphasis on privacy, increased transparency and data access, systemic challenges to democracy, and well-being (see e.g., Buning et al., 2018). Algorithmic and data transparency become a geopolitical battlefield as these data traces on Nordic and European Union citizens – with a potential to understand fundamental social questions for the twenty-first century and to disclose challenges to democracies and well-being – are held by private companies in the US without incentives to share. The platform companies can tune their algorithms for visibility of some content and actors over others that potentially directly, or more likely indirectly, affect opinion formation and online and offline behaviour during crises and elections. Moreover, the platforms can change their policies on a whim because they are private companies. Increased use of social media platforms comes at a societal price (Bechmann, 2020) if influential actors are too opaque or democratic societies have too little knowledge of what constitutes influence, leaving them unable to mitigate harmful patterns to citizens' trust in societal institutions and well-being (Brennen et al., 2020; Charquero-Ballester et al., 2021).

The digital communication and Internet research community has heavily criticised the platforms for the lack of data-sharing with academic researchers affiliated with research organisations that would allow for general knowledge of platform usage in national and global populations and especially independent scrutiny of systemic risks and potential influences of platforms on democracy and well-being (Bruns et al., 2018; de Vreese & Tromble, 2023; Freelon, 2018). Since 2014, the platforms have increasingly restricted access to data for research purposes to the extent that it is almost impossible to answer fundamental questions with dedicated access solutions, which are primarily constructed for American researchers at wealthy and reputable research institutions where acquisition of data or extended collaborations have been conducted with platforms (Moeller & Bechmann, 2019). The European Union has increased the pressure on data access for researchers affiliated with a research organisation (as defined in Article 2 of Directive 2019/790, and referred to in especially the Digital Services Act (2022/2065) Article 40 (4)). This creates hope for future solutions where in principle (after a vetting procedure), all independent research that addresses systemic risks must be granted access to adequate data from the platforms (Horowitz et al., 2024; Walter & Bechmann, 2023). However, exactly this regulation is now a battlefield where platforms have withdrawn from the voluntary code of practice already signed (EFCSN, 2025; The Brussels Times, 2023) and the Digital Services Act is now being debated as part of a larger geopolitical question of accountability and regulation of Big Tech in the European Union.

This uproar and unstable situation with regards to platform data has forced independent researchers to be creative in how to approach platform studies and the collection of data for large-scale trace data studies. Some are increasingly scraping data from platforms with the fear of being sued for violating terms of services (Thomas, 2021). Others have constructed dedicated software, among other reasons to supplement available mandatory repositories, only to find errors and being shut down by platforms (Hatmaker, 2021), creating a chilling effect on the international research

community. Others yet have designed digital computational tools to access data through previous application programming interfaces and tried to keep up with new solutions in an endless and downward spiral towards ever more limited access (Bechmann & Vahlstrup, 2015; Rieder, 2013; Rieder et al., 2018).

The General Data Protection Regulation (2016/679) implementation in May 2018, stipulates that users must be able to transport their personal data out of any platform (Article 20). Researchers have also tried to utilise this regulation to study digital platform trace data by asking users to share their data portability logfiles or other types of user-created data such as screenshots with them (Razi et al., 2022; van Driel et al., 2022; Wedel et al., 2024; Wei et al., 2020). In contrast to application programming interface trace data, the method is thus user-centric as it follows the activity of the donating user and the activity traces that the person leaves behind, and the method relies on collaborating with the user (Breuer et al., 2023; Ohme et al., 2021, 2024; van Driel et al., 2022).

Data donation as a method has been applied to relatively small samples of users so far, with samples and the number of donating participants varying from $n = 102$ (van Driel et al., 2022), $n = 150$ (Breuer et al., 2023), to $n = 435$ (Welbers et al., 2024) (see also Ohme et al., 2021; Pfiffner, 2024). While application programming interfaces as a tool for data collection has proved its potential to provide knowledge about media usage at scale (e.g., Hussain et al., 2018), data donation has not been applied at scale so far for studying platform usage of a national population where the sample size of the donating participants potentially allows for inferences to the population of interest. Even at a smaller scale, data donation studies are scarce. van Driel and colleagues (2022) obtained Instagram data download packages from participants in their study for analysing how the platform is used by their participants, with one result showing that “liking” was the most common activity. Breuer and colleagues (2023) collected Facebook data and compared the obtained information with a study that used a browser plugin to collect similar user data. Ohme and colleagues (2021) collected screenshots about log data for smartphone usage with the result that the donated data showed a higher usage than self-reports about smartphone usage of the same respondents. These and similar studies show the potential of donated data for providing different insights into media usage than alternative methods (e.g., tracking or self-reports), while at the same time discussing potential disadvantages, such as biased participation (e.g., van Driel et al., 2022). In addition, Boeschoten and colleagues (2022) have developed a framework to assess errors occurring in data donation studies. So far, however, the framework has not been discussed in the context of a study conducted at scale (which we do by looking at drop-outs at different stages of our study). Furthermore, the study reported here is internationally one of the first studies to carry out a data donation-based data collection at national scale. We therefore find it important to assess whether methodological findings from existing small-scale studies can be confirmed in this study and can thus fertilise the discussion of implications of an at-scale approach for donated data. We therefore outline our procedure step-by-step and discuss these implications.

2.2 Objectives and contribution to the field

The objective and contribution of this publication is to describe the novel study design constructed to scale data donations to national populations and discuss its potentials and limitations as a way of providing increased knowledge of digital media usage and behaviour. The YouTube data donation as an illustrative case on national population scalable data donation is part of a larger study where the objective is to understand social media influence in national populations, with one of the cases of the research project being YouTube as an infrastructure for influence. The project, “Social Media Influence” (funded by Independent Research Fund Denmark), focuses on what and who influences behaviour and opinion formation on social media in relation to especially systemic risks to democracy and well-being.

Apart from developing the data donation method for collection at national scale, the larger project aims to contribute to the field of digital communication by presenting new empirical findings on potential influence patterns and dynamics by linking donated data with a post-survey answered by all participants. How actors influence citizens via media is one of the oldest questions in media and communication studies, especially during elections and crises, when people need and seek out media for information. Such influence has been difficult to measure due to numerous contextual factors, and new theories point to greater agency of citizens (Allcott et al., 2019; Allcott & Gentzkow, 2017; Benkler et al., 2018). The Covid-19 pandemic, Russia’s war in Ukraine, the latest violent developments in the Israel–Palestine conflict, and the American and European Parliament elections have reignited interest in media influence, but now in a much more complex global media environment with personalised and algorithmic curation concealing who is viewing and potentially influenced by which type of content shared on social media. Where influence is conceptualised as the power to affect, effect is the result (Katz & Lazarsfeld, 1955; McQuail & Deuze, 2022). Coining an infrastructure theoretical understanding of influence makes the study sensitive to hidden dynamics of influence in social media as a socio-algorithmic infrastructure both scripted by and scripting behaviour. The study moves beyond state of the art with an empirically grounded (Gruzd & Wellman, 2014) and non-predefined actor perspective on influence inspired by Science and Technology Studies (Bechmann & Bowker, 2019; Bowker et al., 2010; Edwards et al., 2009; Plantin et al., 2018) and focuses both on what and who influences in relation to platform communication.

Effect studies (Valkenburg & Peter, 2013) show that influence may take place at cognitive, emotional, or excitative (action) levels by, for example, igniting, confirming, or changing knowledge, values, opinions, attitudes, emotions, and actions of people. This study reconfigures such conceptions into how the infrastructures of social media potentially enable these through analyses of,

1. content viewing as a result of influence in a dual way (what people view may be influenced by algorithmic recommendations, and user-decisions on what to view and search for also impact future algorithmic curation; thus potentially influencing what is made visible to themselves and other users in the future; Bucher, 2012);
2. other engagement than viewing as traces of citizens’ digital actions and content having influenced them (Rogers, 2018) and in turn feeding into potential further amplification of the content entity of engagement;

3. self-reported perceived influence by citizens, either of the platform as an infrastructure in itself, associated actors and channels such as influencers and microcelebrities (Abidin, 2016; Marwick, 2013), or the content, structures, and formats of the platform.

By theoretically configuring influence to encompass different modes of influence, the larger project assumes, despite limitations, that distinct socio-algorithmic influence patterns can be identified in relation to different political, psychological, and sociodemographic user-profiles as well as in different topical, event, or situational contexts.

For this publication, the data collection process description and discussion illustrates the ability to collect donated logfiles for a sample collected to be representative of adult Danish YouTube users as an illustrative case on data donation as a method for understanding digital media usage and behaviour at national population scale. However, for the larger project, the data collection also served the purpose of understanding how the behaviour reflected in the trace data is connected to certain offline events (sampling across events more than four years back in time), to certain general media usage profiles, psychological profiles (e.g., based on trust, happiness, neuroticism, conspiracy mentality), political profiles, sociodemographic profiles, and to self-reported influence, measured influence effects (experiment), and context of the donated data, for example, whether other people are using the donated account.

Even though answering these larger questions of influence is not our objective in this publication, it is still the main reason for including the additional post-questionnaire of the study as a supplement to the individual-level donated logfiles.

Here we provide a methodological design for what steps can potentially be taken to craft a national population-scale data donation study (including a post-questionnaire to investigate correlations between the behaviour reflected in the data and other variables of interest). Compared with earlier data donation studies, the study presented here has succeeded in collecting logfiles from a larger sample and has used quotas to approximate a sample matching the adult Danish population of YouTube users. We discuss this in the context of specificities of the Nordic region and media system in the discussion and conclusion section.

2.3 Structure of the publication

This publication is structured to mirror the data collection process, and the benefits and challenges of each step are addressed. After this, we present general reflections on ethics and discuss, in relation to existing studies, to what extent the setup was successful in fulfilling the aim of the study, namely, to collect national population-scale data to understand YouTube usage.

The data collection for this study was done in four steps, administered as four interlinked online surveys. Donation data and responses to the post-survey were linked at the level of individual respondents:

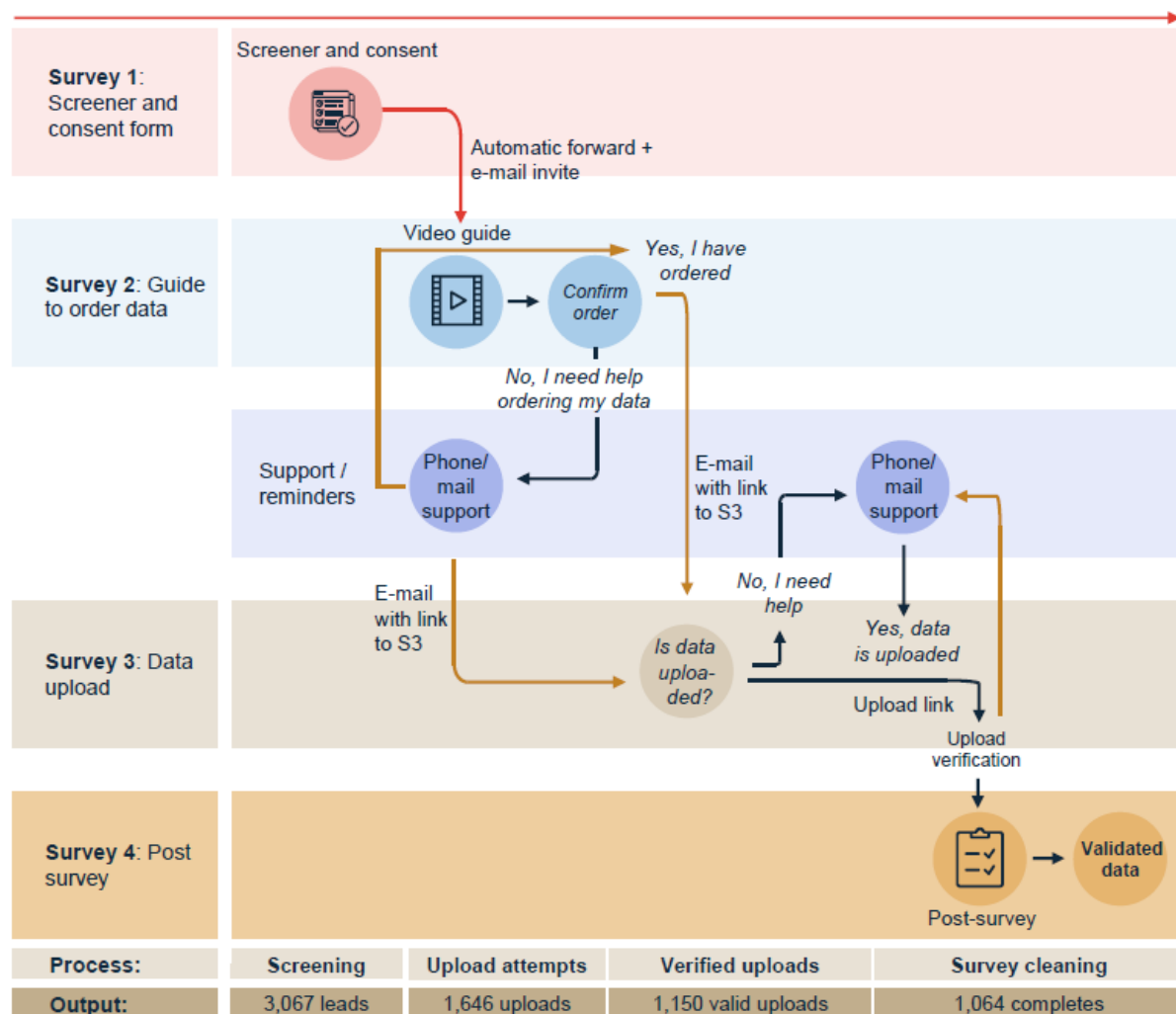
- Survey 1: Respondents were screened for their eligibility to participate and complied to participate following informed consent.
- Survey 2: Respondents received instructions on how to request their YouTube data.
- Survey 3: Respondents received instructions on how to download and share their data with the research project using a secure upload solution.
- Survey 4: A comprehensive post-survey was administered.

We first provide a brief overview of the four steps of the data collection process. Next, we account for the target population, recruitment of respondents, and how participation in the study developed over the field period. After this follows a detailed account for the screening process and analysis of which demographic groups were more likely to be screened out or drop off at different stages. Next, we account for how respondents were instructed to request, download, and securely donate their data and analyse what sociodemographic groups were more likely to drop off during the technically demanding part of the data donation process. After this, the final sample is compared to the target population and a description of participants' study completion time is provided. Upon this, we review the content and quality of the donated data and discuss the ethical considerations regarding the study. Finally, we provide a concluding discussion on the potentials and limitations of data donation as a method to understand usage trends and challenges on a national population scale. The Appendix contains screenshots of Surveys 1–4 and thus documents the implementation of the screener and data donation process as well as the post-survey questionnaire.

3. Overview of the data collection steps

The data collection process was divided into four consecutive steps, set up as interlinked surveys (see Figure 1).

Figure 1 Flow of data collection steps



Source: Epinion (exception: valid leads computed by authors)

In Survey 1, respondents were screened for their eligibility to participate in the study based on questions about their gender, age, level of education, YouTube usage, and willingness to donate their YouTube data (see specification of screening criteria below). In total, 48,405 panellists were invited to participate by e-mail, out of whom 22,733 completed at least parts of Survey 1. Out of these respondents, 9 turned out to be duplicates (reported by Epinion) and 4 had bypassed the technical setup that should ensure that no respondents who completed the screener could re-enter it (indicated by a later Survey 1 start time than end time). These 13 observations were excluded from the analysis, leaving 22,720. Of these respondents, 3,067 (13.5%) completed the screener, matched the screening criteria, gave informed consent to share their data, and provided their contact information, making them valid leads.

These respondents received an e-mail with a link to Survey 2. Of them, 2,726 respondents (88.9% of the valid leads) started the survey, where they were given instructions on how to order their YouTube activity data through Google's Takeout solution. In addition to the guide, Survey 2 contained a set of control questions to ensure that respondents had ordered their data correctly. At the end of the survey, respondents who confirmed that they had ordered data correctly were informed that they would receive their data within five days, but typically within a few hours, and were provided a link to Survey 3.

Survey 3 guided respondents to download and name the zip file containing the data they had ordered through Google Takeout and provided them with a link to securely upload their file to Epinion. 2,007 respondents started Survey 3 (65.4% of the valid leads). According to Epinion, who managed the data upload, 1,646 respondents (53.7% of the valid leads) successfully uploaded their data. When a respondent's data file was uploaded, the data was validated against a set of parameters (see below) to confirm that the data had been downloaded and uploaded correctly and that the file contained the necessary data. According to Epinion, 1,150 data uploads were verified as valid (corresponding to 69.9% of the total uploads or to 37.5% of the valid leads). Respondents with valid data uploads who completed Survey 3 were provided a link to Survey 4.

Survey 4 contained the main questionnaire, which included questions about social media use, news use, mental health, personality traits, conspiracy mentality, political views, political interest, political efficacy, and more. By linking data donations with responses to the post-survey at the level of individual respondents, it is possible to explore correlations between the mentioned variables and different YouTube usage patterns. Furthermore, questions about YouTube usage (incl. whether respondents let others use their accounts, how they select YouTube content, and how they watch YouTube content) make it possible to compare self-reports to the donated data. 1,068 respondents completed the post-survey questionnaire. However, inspection of the collected data showed that four respondents had unusually fast response rates (see criteria below), and five respondents who had passed the data validation turned out to not meet the data donation requirements. This results in a final sample of 1,064 respondents with valid complete responses to Survey 4 and 1,063 respondents with successful data donations. After completing Survey 4, respondents received a gift certificate of 150 Danish kroner (approx. 20 euros or 20 US dollars), which could be used for sendtanke.dk.

Due to the many steps of the data collection and the relatively low share of invited participants that were converted into valid leads, the field period was longer than originally expected, lasting from 15 August–26 November 2024, starting with a pilot period that ended on 20 September 2024.

The next sections explain the recruitment of respondents and each of the data collection steps in further detail and account for screening-out of respondents and drop-off in each step.

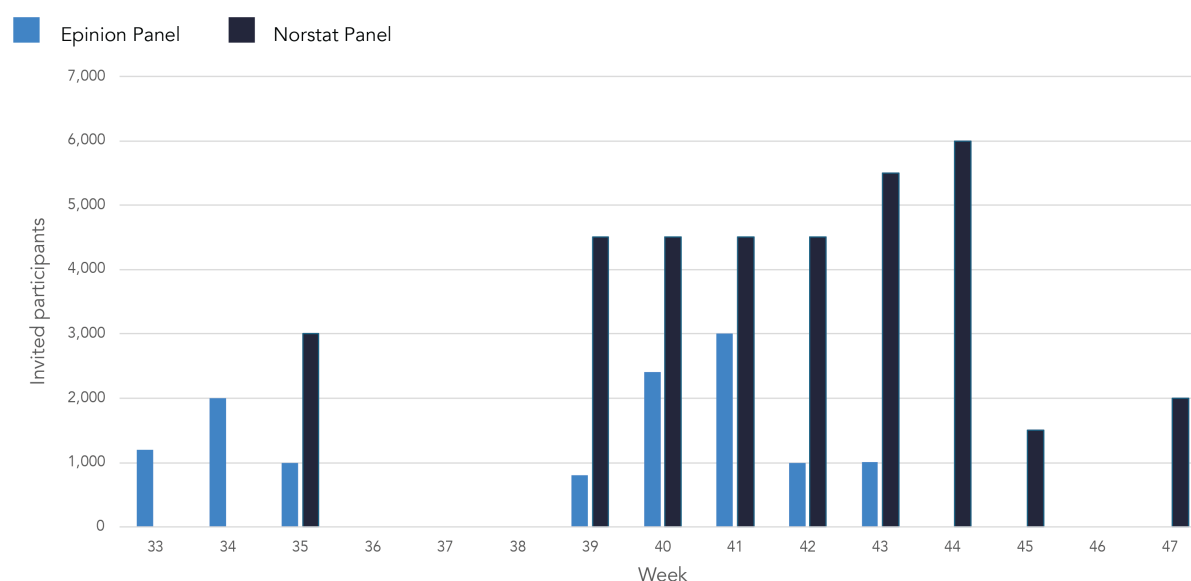
4. Recruitment

4.1 Panel description

The study was conducted with participants from two panels. The first is Epinion's own panel – called the qualitative panel by Epinion – which is an opt-in panel with panellists who agreed, in addition to participating in surveys, to be invited to studies that require more effort than traditional survey studies (e.g, focus groups). We refer to this panel as the Epinion panel. The second is Norstat's panel – with Norstat being a sub-contractor of Epinion for this study and running an opt-in panel with panellists who agreed to participate in surveys and more extensive studies. We refer to this panel as the Norstat panel.

Epinion started to recruit via the Epinion panel and continued to recruit via the Norstat panel with parallel fieldwork weeks (for a distribution of invitations across the fieldwork period, see Figure 2). Across the two panels, 48,405 panellists were invited to participate. In the Norstat panel, 36,000 individuals were invited (74% of the total invited), in the Epinion panel, 12,405 individuals were invited (26% of the total invited) to participate. In the sample that completed all parts of the study ($n = 1,068$), 155 (15%) were originally invited via the Epinion panel and 913 (85%) via the Norstat panel.

Figure 2 Distribution of invitations across the fieldwork period by panel, 2024



Source: Epinion

The gap in the fieldwork period can be explained by the evaluation of the pilot phase and resulting adjustments to the data collection (see section 7.1) that needed implementation before the actual start of the field-work period.

4.2 Incentives

Participants who completed the full study received a gift certificate of 150 Danish kroner (approx. 20 euros or 20 US dollars) for a website (sendentanke.dk) that allows the use of the voucher for different products. In addition, participants from the Norstat panel received coins from Norstat for participating in the screening survey (Survey 1) equalling approximately 5 Danish kroner (less than 1 euro/ US dollar) for usage in the Norstat incentive system.

4.3 Invitation material

Respondents were invited to participate via e-mail. The invitation introduced the overall purpose of the study, broadly described as gaining insights into Danes' use of YouTube and attitudes towards various societal issues. Furthermore, respondents were informed that the study consisted of several steps, yet with no detailed explanation hereof. At this stage, respondents were not informed that the study required data donation (see elaboration in section 5.3). The invitation also informed that the purpose of the first step of the study was to determine whether participants fit the target group, and that participants who fit the target group and completed the study would receive the 150 Danish kroner incentive (see section 4.2). After respondents opened the invitation link, they were asked for consent to participate, following information about how to get in touch with Epinion and the research team, their right to withdraw their consent at any time, possibilities of having their data deleted, and links for additional information about their rights.

For Norstat panelists, the procedure included an additional step. They received an invitation to the study in the Norstat system that asked them to start the screening survey (Survey 1) and informed them that it would take approximately five minutes to complete and that they would receive the five coins (see section 4.2 about incentives) for completing it, which could be used in the Norstat incentive system. If they followed the link in this invitation, they were redirected to Epinion's implementation of the screening survey. The first page of the screener gave respondents the same information provided to Epinion's panelists in the invitation e-mail along with the additional information about the study (as described above).

4.4 Reminders

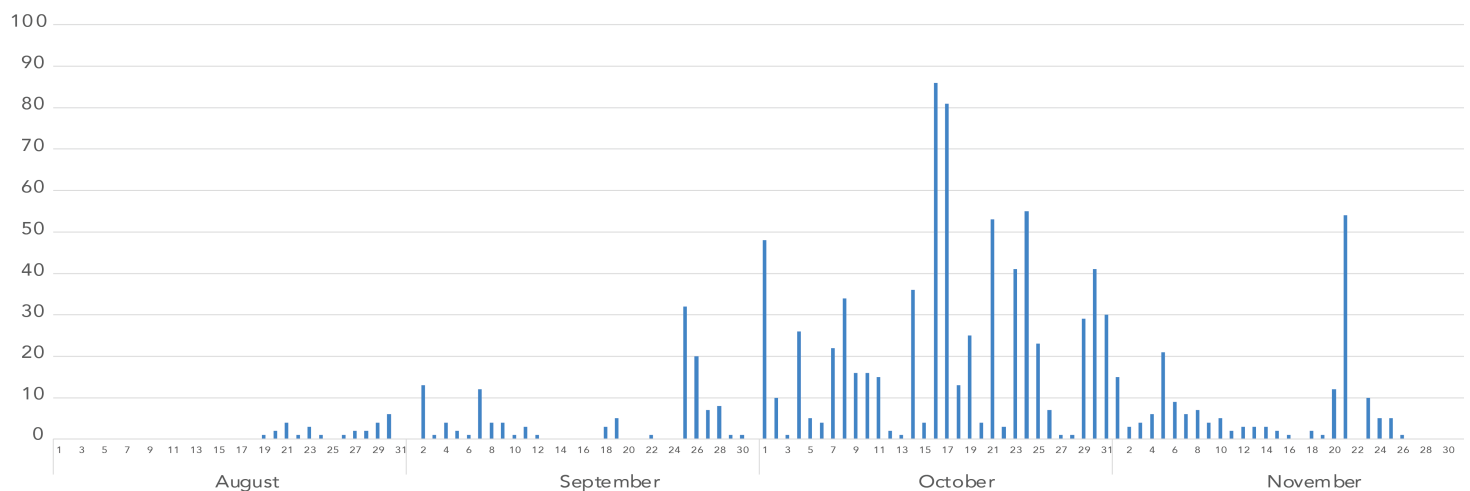
Panelists who did not open the link to the study they received in the invitation e-mail received up to two e-mail reminders. Reminders contained the same information as the invitations. First reminders were sent three to five days after the invitations, and second reminders were sent three to five days after the first reminders. Panelists who only partly completed Survey 1 to Survey 3 were contacted via phone or text message and encouraged to continue their participation. To meet the sample quotas, efforts to recruit panelists via reminders focused on underrepresented sub-groups – namely female panelists aged 55 or older – towards the end of the data collection period.

4.5 Completions over time

Throughout the fieldwork period, Epinion invited participants to the study continuously. In Figures 3 and 4, we show how validated completions developed over the fieldwork period by analysing the registered end times of Survey 4. While Figure 3 shows the number of completions per fieldwork day, Figure 4 shows the accumulated numbers.

Both figures show that completions increased noticeably at the end of September. This can be attributed to the end of the pilot phase of the study and the resulting adjusted inclusion criteria (see section 7.1). On some days, more completions are registered than on others. The peak on 21 November, for example, can potentially be explained by Epinion's effort to recruit the missing older female participants and sending out more invitations and reminders to this target group.

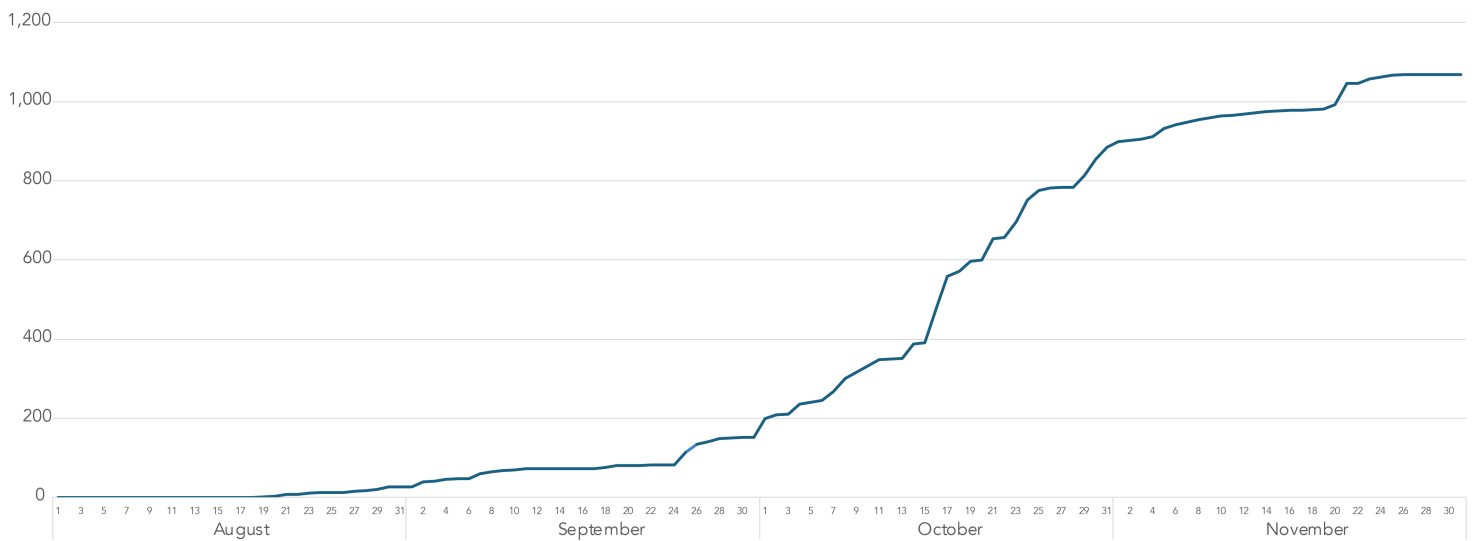
Figure 3 Number of completions by fieldwork day



Comments: $n = 1,068$. Fieldwork started 15 August 2024.

Source: computations by Aarhus University

Figure 4 Accumulated completions by fieldwork day



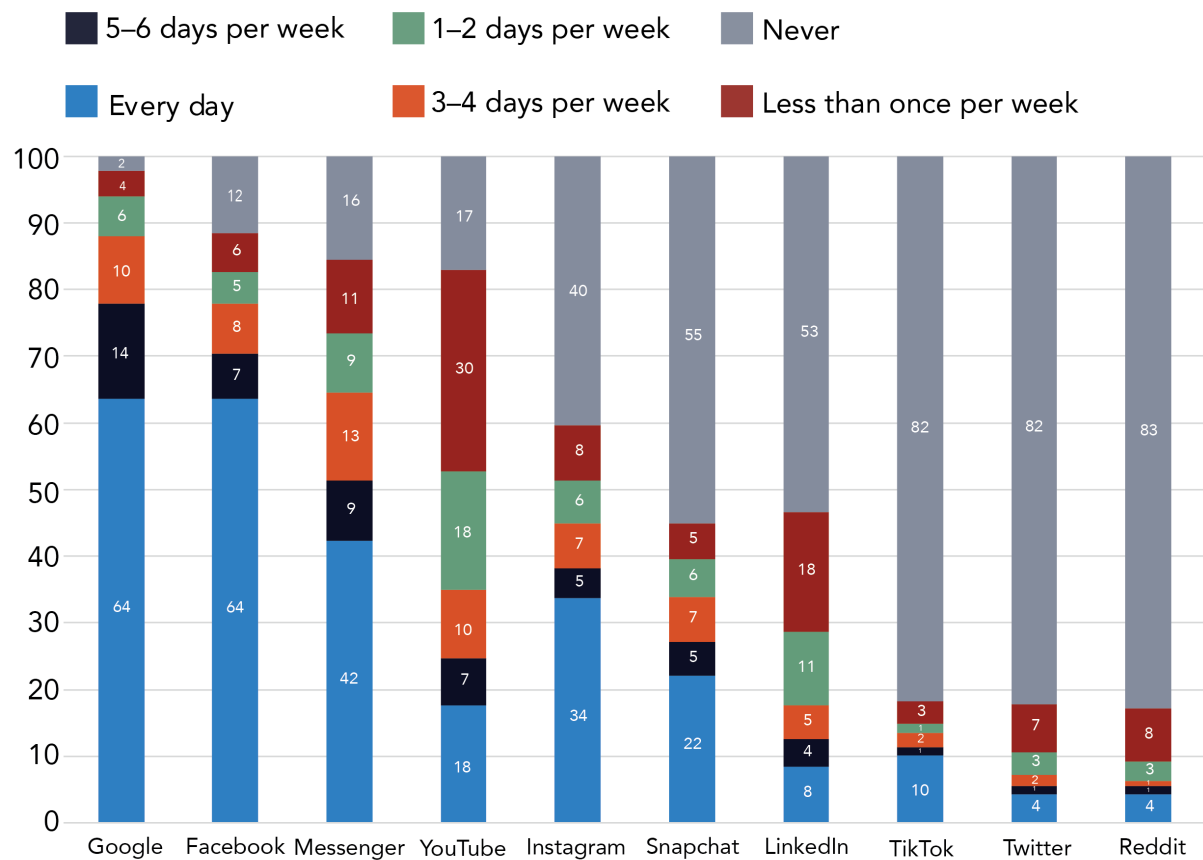
Comments: $n = 1,068$.

Source: computations by Aarhus University

4.6 Target population

In collaboration with Epinion, quotas for age, gender, and education were developed for the Danish adult YouTube user population. The basis for the computation was a nationally representative survey conducted by Epinion to measure social media usage among the adult Danish population. Figure 5 shows the social media usage measured in that survey. This distribution provided Epinion information about the size of the YouTube user population in Denmark. Approximately 50 per cent of the participants in this study reported weekly use of YouTube, with 18 per cent reporting daily usage.

Figure 5 Social media usage in the adult Danish Population (per cent)



Comments: $n = 953$. Danes, national representative selection.

Source: Epinion

In the present study, participants must have used YouTube at least every second week to be eligible to participate (see elaboration below). As such, we define the adult Danish YouTube user population as persons aged 18+ who use YouTube at least every second week. However, biweekly users could not be identified based on Epinion's social media usage survey, which distinguished only between weekly usage and less than weekly usage. To define the quotas employed for recruitment of participants in the present study, Epinion used the sociodemographic distributions (age, gender, and education) for the respondents who reported using YouTube at least once a week in the above-mentioned survey, as this was the best available approximation of our target population. This is also the benchmark we use for evaluating how successful the study was in obtaining a final sample representative of the adult Danish YouTube user population. These distributions are displayed in Table 1.

Table 1 Demographic distribution of 18+ YouTube users in Denmark

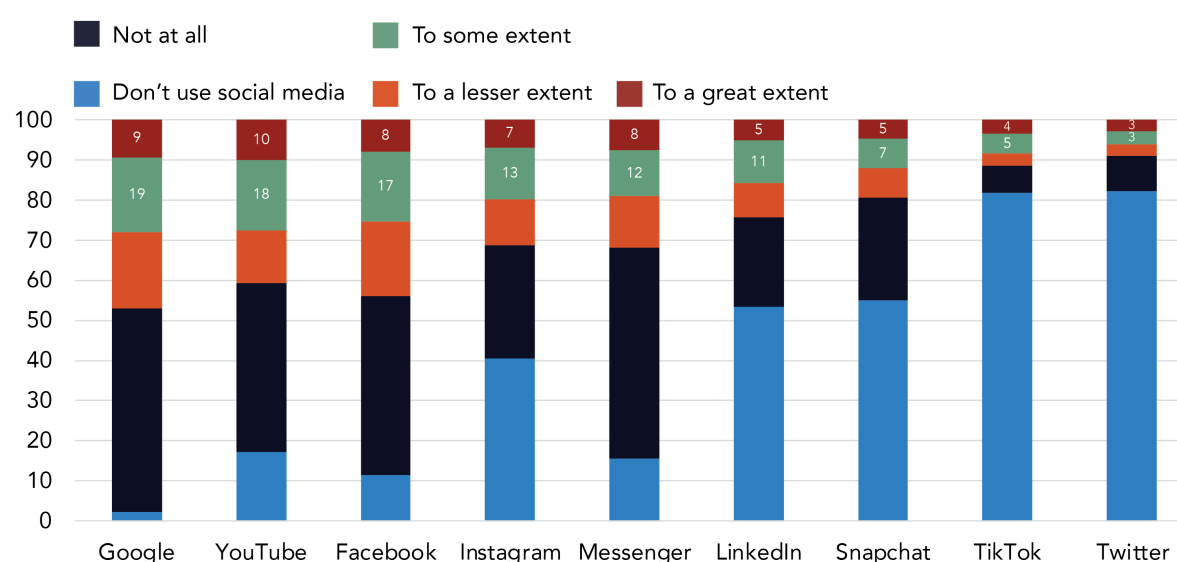
Quota	Group	%
Age	18–34	39
	35–54	33
	55 +	28
Gender	Female	45
	Male	54
Level of education	Lower-secondary school or vocational training	25
	Upper-secondary school or short-cycle higher education	26
	Medium-cycle or long-cycle higher education	49

Comments: $n = 488$. Danes who reported using YouTube at least weekly. Gender does not sum to 100% because 0.3% identified as non-binary (not shown in table) and because of rounding of decimals.

Source: Epinion

The survey conducted by Epinion also covered information about the willingness to donate data among YouTube users in Denmark. Results showed that there is a general willingness to donate YouTube or Google data among 28 per cent of the adult Danish population (see Figure 6). Epinion stated that additional analyses showed that women and men were equally willing to donate data and that younger respondents were more willing to donate. Epinion used this information to evaluate the feasibility of the study and for planning the recruitment process.

Figure 6 Willingness of Danes to donate social media data (per cent)



Comments: $n = 953$. Danes, national representative selection.

Source: Epinion

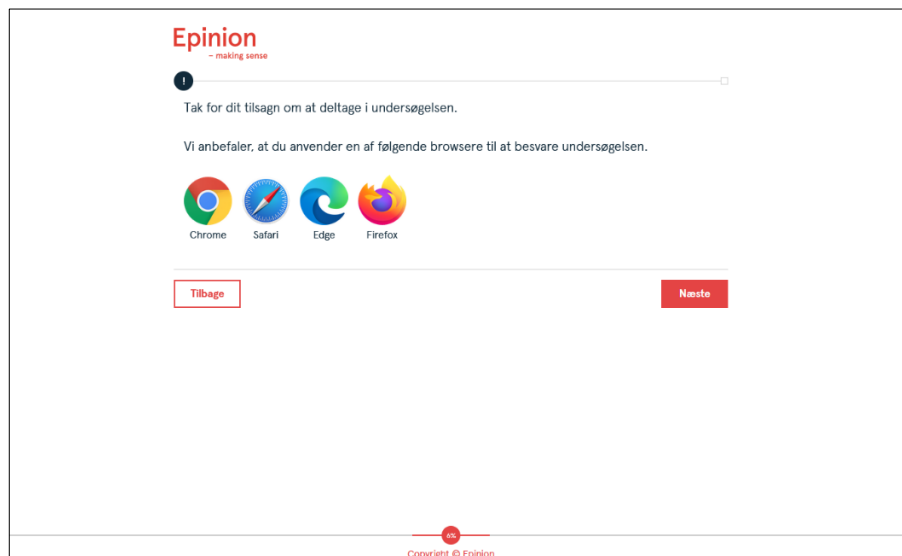
5. Survey 1: Screening participants into the study

The purpose of Survey 1 was to determine whether invited panellists fit the criteria for participating in the study. First, respondents were introduced to the overall purpose of the study (broadly described as gaining insights into Danes' use of YouTube and attitudes towards various societal issues). Furthermore, they were informed about the use and storage of their data and asked whether they consented to participate. Next, respondents who consented were presented with two screening questions about their use of YouTube to assess whether they would have sufficient YouTube data to participate in the study. To be eligible to participate, respondents must have reported using YouTube at least every second week and that they started using the platform before 2020. The latter was important for obtaining data allowing for analyses of changes in viewing patterns over time. Whether the respondents used YouTube frequently enough and had data going far enough back in time could also be (and was) assessed after the respondents donated their data. However, by using self-reported YouTube use to screen the participants at the beginning of the survey, we could minimise the time and effort spent by participants who would later find out that they had too little data to participate in the study. After the screened-in participants had uploaded their data (in Survey 3), it was validated that they had adequate data to participate in the full study (see explanation in section 7.1).

Survey 1 was set up so that respondents who had completed the survey could not re-enter it. However, four respondents apparently managed to bypass this mechanism, reflected in a survey start time that was clocked later than their survey end time. These respondents have not been included in the analysis below. In addition to this, the nine respondents that Epinion registered as duplicates have not been included in the analysis (to avoid that these respondents' answers would count twice).

At the start of the survey, respondents were encouraged to use Chrome, Safari, Edge, or Firefox browsers (see Figure 7), as we had tested the setup using these browsers. We tested the setup both on different mobile phones and across PC and Mac platforms.

Figure 7 Screenshot of instructions to use specific browsers for the study



5.1 Screening based on YouTube use

Table 2 displays the distribution of the respondents' answers to the first screening question about frequency of YouTube use. As the table shows, a bit more than two thirds (70.1%) reported using YouTube at least once every second week and were thus eligible to participate. A little less than one third (29.9%) reported using YouTube less frequently and were screened out ($n = 6,639$). In addition to this, 481 respondents did not provide an answer to the question and were screened out for this reason. It should be noted that the share of respondents who reported weekly use of YouTube is a bit higher than expected based on Epinion's prior survey on Danes' social media use (see Figure 5). This might be because some invited panellists who seldom or never use YouTube chose not to participate in the study because the invitation and consent form stated that it was about YouTube use.

Table 2 Screening question 1, frequency of YouTube use

		%	N
Frequency of YouTube use	Every day	20.1	4,462
	Every second day	12.5	2,788
	Twice a week	15.2	3,388
	Once a week	12.5	2,776
	Once every second week	9.8	2,186
	Once a month	10.7	2,376
	Less than once a month	19.2	4,263
Total		100.0	22,239

Comments: Originally, respondents had to use YouTube on a weekly basis to participate in the study. However, in the pilot phase of data collection, it became clear that this made it too difficult to achieve enough valid leads. Therefore, this screening criterion was adjusted to biweekly use (see more in section 7.1). All respondents who had initially been screened out because they only used YouTube every second week were re-invited to participate in the study. However, according to Epinion's records, 167 of them did not re-enter the study. As such, these respondents were also screened out in this step.

The distribution of answers to the second screening question about when respondents started using YouTube is displayed in Table 3. The 6.8 per cent ($n = 1,048$) who reported having started using the platform in 2020 or later were screened out. In addition to this, 180 respondents did not answer the question and were screened out for this reason. This includes the 167 who were originally screened out because they only used YouTube every second week and were later re-invited, but who did not re-enter the survey (see the comments under Table 2).

Table 3 Screening question 2, started using YouTube

	Year	%	N
Started using YouTube	2005–2009	38.9	5,991
	2010–2014	33.7	5,190
	2015–2019	20.7	3,191
	2020 or later	6.8	1,048
Total		100.1	15,420

Comments: Total sums to more than 100 per cent because percentages are rounded.

5.2 Screening based on sociodemographic profile

After the initial screening questions, respondents were asked about their age, gender, and highest level of education to screen in a sample of respondents that came as close as possible to the target population on these parameters. 513 respondents either did not answer the question about age or reported a date of birth indicating that they were

below 18 years old (making them ineligible to participate). Many reported an age of 0 or 1 years, indicating that these respondents might have been old enough to participate but simply typed in their birth date wrong. Even so, these respondents were screened out.

461 respondents did not answer the question about gender. According to Epinion, this includes 457 who were screened out because their age profile did not fit the sampling criteria. 137 respondents did not answer the question about their highest level of education. According to Epinion, this includes 133 who were screened out because their gender profile did not fit the sampling criteria. All these respondents were male and, as noted above, we needed older female respondents towards the end of the data collection to fill the sample quotas.

5.3 Screening based on consent to donating YouTube data

By design, the consent form described above that respondents were initially presented did not mention donation of YouTube data. Information about this was presented later in a second consent form, after respondents had answered questions about their YouTube use and sociodemographic profile. This makes it possible to analyse whether participants who were initially interested in participating and who use YouTube sufficiently to be eligible to participate ended up deciding not to after learning that they must donate data. Furthermore, it is possible to analyse what characterises this group of participants. In this second consent form, respondents were also informed that it would be possible to get help from a telephone hotline in case they experienced technical difficulties and were reminded that they would receive a gift certificate of 150 Danish kroner if they completed the study.

Table 4 displays the shares of respondents who did and did not consent to donating data, divided by age, gender, and level of education. A relevant first thing to note is that close to half of the respondents (49.2%) who were eligible to participate in the study at this point (based on the prior screening criteria) did not consent to participate after being informed that they had to donate their data. This indicates that donating data is something that can dissuade many potential participants. The respondents who did not consent to donating their data were screened out ($n = 6,346$), as were 362 respondents who did not answer the question (according to Epinion, 22 of these respondents were screened out because they did not meet the sociodemographic profile needed to fill the sample quotas).

Table 4 Consent to donating data

		No		Yes		Total	
		<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	1,186	39.7	1,800	60.3	2,986	100.0
	35–54	1,578	46.0	1,854	54.0	3,432	100.0
	55+	3,582	55.3	2,899	44.7	6,481	100.0
	Total	6,346	49.2	6,553	50.8	12,899	100.0
Gender	Female	2,950	50.1	2,939	49.9	5,889	100.0
	Male	3,386	48.5	3,592	51.5	6,978	100.0
	Total	6,336	49.2	6,531	50.8	12,867	100.0
Level of education	Lower-secondary	420	46.4	486	53.6	906	100.0
	Upper-secondary	672	43.0	891	57.0	1,563	100.0
	Vocational	1,544	53.5	1,341	46.5	2,885	100.0
	Short-cycle	613	51.3	583	48.7	1,196	100.0
	Medium-cycle	1,408	51.5	1,327	48.5	2,735	100.0
	Bachelor	593	45.9	700	54.1	1,293	100.0
	Long-cycle	1,096	47.2	1,225	52.8	2,321	100.0
	Total	6,346	49.2	6,553	50.8	12,899	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

The next relevant question is of course whether there are any systematic differences between participants who were willing and unwilling to participate when told that they must donate their data. The findings displayed in Table 4 suggest that this is the case. Most notably, there is a strong association between age and consent to participate. Among the youngest age group of 18–34-year-olds, 60.3 per cent consented to participate, compared with 44.7 per cent of the oldest age group (55+). A chi-square test showed that the association between age and consent is significant at the .05 level [χ^2 (2, n = 12,899) = 217.16, p < 0.001]. While the data does not reveal why respondents decided not to consent, the lower willingness to participate observed among the older respondents could be due to less confidence in their own technological skills in terms of requesting, downloading, and donating their data.

Table 4 also indicates a small gender difference, where larger shares of men than women consented to participate (51.5% vs. 49.9%). However, this difference is not significant (χ^2 [1, n = 12,867] = 3.15; p = 0.076). Finally, the table also shows differences between education groups: The biggest difference is found between those holding upper-secondary education (equivalent of high school) as their highest level of education, among whom 57 per cent consented, and those with a vocational education, among whom 46.5 per cent consented to participate. A chi-square test showed that the association between education and consent is significant [χ^2 (6, n =

12,899) = 65.65, $p < 0.001$]. However, it should be noted that there is no clear linear pattern where lower or higher education is related to greater or lesser willingness to participate.

5.4 Screening based on availability of stored YouTube data

Next, respondents who consented to donating their data were asked if they had a personal Google account and whether their settings for this account were set to save their YouTube data. Respondents were informed that having a Google account on which YouTube data was saved and having YouTube watch-history data going back to before 2020 was necessary for participating in the study.¹ To help respondents provide as accurate answers as possible, they were informed that if they used services like Gmail, Google Docs, or YouTube, they likely had a Google account. Furthermore, they were provided a visual six-step guide to identify whether their Google account was set to save their data (see the Appendix).

Beginning with the question of whether respondents had a Google account, Table 5 displays the shares of respondents who reported they did have a Google account, did not have a Google account, or did not know, divided by age, gender, and level of education. As the table displays, approximately one in five of the respondents (19.1%, $n = 1,244$) were screened out because they either did not have a Google account (12.0%) or did not know whether they did (7.1%). In addition to this, 56 respondents did not answer the question and were screened out for this reason.

¹ In the data validation process, we enforced the less strict criterion that respondents must have data dating back to before 2021 (see section 7.1). A stricter criterion was named in the instructions to respondents at this point to lower the risk that some would complete the first three survey steps and then be rejected during data validation due to data not dating far enough back in time.

Table 5 Google accounts by age, gender, and level of education

		Yes		No		Don't know		Total	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	1,608	89.9	120	6.7	60	3.4	1,788	100.0
	35–54	1,549	84.3	200	10.9	89	4.8	1,838	100.0
	55+	2,096	73.0	461	16.1	314	10.9	2,871	100.0
	Total	5,253	80.9	781	12.0	463	7.1	6,497	100.0
Gender	Female	2,341	80.6	315	10.8	249	8.6	2,905	100.0
	Male	2,893	81.0	463	13.0	214	6.0	3,570	100.0
	Total	5,234	80.8	778	12.0	463	7.2	6,475	100.0
Level of education	Lower-secondary	405	84.2	49	10.2	27	5.6	481	100.0
	Upper-secondary	773	87.3	67	7.6	45	5.1	885	100.0
	Vocational	1,037	77.9	181	13.6	113	8.5	1,331	100.0
	Short-cycle	464	79.9	73	12.6	44	7.6	581	100.0
	Medium-cycle	1,015	77.1	190	14.4	111	8.4	1,316	100.0
	Bachelor	590	85.0	60	8.6	44	6.3	694	100.0
	Long-cycle	969	80.1	161	13.3	79	6.5	1,209	100.0
	Total	5,253	80.9	781	12.0	463	7.1	6,497	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

The table also displays notable differences between age groups. Among the 18–34-year-olds, 89.9 per cent reported they have a Google account and only 3.4 per cent said that they do not know. In comparison, only 73.0 per cent of the oldest age group (55+) reported that they have a Google account, and 10.9 per cent do not know. A chi-square test showed that the association between age and having a Google account is significant [χ^2 (4, $n = 6,497$) = 233.21, $p < 0.001$]. This indicates that the need for having a Google account (and being able to determine whether one has an account) can create age skewness in the sample when working with data donation.

Regarding gender, Table 5 shows that a slightly higher share of men reported that they do not have a Google account (13.0% compared with 10.8%), whereas a slightly higher share of women reported that they do not know (8.6% compared with 6.0%). A chi-square test showed that the association between gender and answers to the question about having a Google account is significant [χ^2 (2, $n = 6,475$) = 20.94, $p < 0.001$]. However, it should be noted that the difference in how large shares of men and women who reported that they do have a Google account is marginal (81.0% vs. 80.6%). In other words, the need for having a Google account seemingly does not create gender skews in the sample. Finally, the table also shows some differences between education groups, with the biggest difference found between those with upper-secondary school as their highest level of education (87.3% reported having a Google account) and those with medium-cycle higher education (77.1% reported having a Google account). A chi-

square test showed that the association between education and answers to the question about having a Google account is significant [$\chi^2(12, n = 6,497) = 58.79, p < 0.001$]. However, as above, it should be noted that there is no clear pattern where lower or higher education is related to having a Google account.

Table 6 displays the shares of respondents who reported that their Google account settings were set for storing their data, reported that it was not, or reported that they did not know, divided by age, gender, and level of education. As displayed in the table, close to one in three (28.1%) of the respondents who did have a Google account reported that their YouTube data was not stored on the account (19.3%) or that they did not know whether this was the case (8.8%). These respondents also had to be screened out ($n = 1,267$). In addition to this, 742 respondents did not answer the question (according to Epinion, 70 of them were screened out due to not matching the socio-demographic profile needed to fill the sample quotas).

Table 6 Storage of YouTube data by age, gender, and level of education

		Yes		No		Don't know		Total	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	1,165	77.7	234	15.6	100	6.7	1,499	100.0
	35–54	1,030	74.9	255	18.5	90	6.5	1,375	100.0
	55+	1,049	64.1	380	23.2	208	12.7	1,637	100.0
	Total	3,244	71.9	869	19.3	398	8.8	4,511	100.0
Gender	Female	1,417	71.6	364	18.4	199	10.1	1,980	100.0
	Male	1,813	72.1	502	20.0	199	7.9	2,514	100.0
	Total	3,230	71.9	866	19.3	398	8.9	4,494	100.0
Level of education	Lower-secondary	265	72.0	69	18.8	34	9.2	368	100.0
	Upper-secondary	532	74.2	125	17.4	60	8.4	717	100.0
	Vocational	605	67.6	184	20.6	106	11.8	895	100.0
	Short-cycle	271	69.3	90	23.0	30	7.7	391	100.0
	Medium-cycle	571	69.0	182	22.0	74	8.9	827	100.0
	Bachelor	394	77.0	78	15.2	40	7.8	512	100.0
	Long-cycle	606	75.7	141	17.6	54	6.7	801	100.0
	Total	3,244	71.9	869	19.3	398	8.8	4,511	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

When looking at differences between age groups, a similar pattern to that described above can be observed. Among the youngest, a larger share reported that their YouTube data is stored on their Google account (77.7%) and a smaller share reported not knowing (6.7%). Among the oldest, only 64.1 per cent reported that their data is stored and 12.7 per cent reported that they do not know. A chi-square test showed

that the association between age and storing YouTube data is significant [$\chi^2 (4, n = 4,511) = 90.64, p < 0.001$]. In other words, the need for respondents to have data stored on their profiles (and being able to assess whether it is) can also create age skewness in the sample when working with data donation. The table also shows that slightly more men than women reported that their data is stored (72.1% vs. 71.6%) and that slightly more women reported that they do not know (10.1% vs. 7.9%). While the difference is significant [$\chi^2 (2, n = 4,494) = 7.19, p = 0.027$], it is small, indicating that the need for storing data does not introduce notable gender skewness in the sample. Finally, the table also displays differences between the education groups. Again, a chi-square test shows that they are significant [$\chi^2 (12, n = 4,511) = 35.77, p < 0.001$], but there is no clear pattern where lower or higher education is related to higher shares storing their data.

5.5 Drop-off during completion of the survey

Before finishing the survey, respondents were asked to share their contact information (name, phone, e-mail, and e-mail associated with their Google account). This was necessary for Epinion to be able to contact the respondents with the link for the next survey. 173 respondents did not provide their phone number, e-mail, or e-mail associated with their Google account. This indicates that these respondents might have dropped off because they did not want to give their personal contact information (but note that, according to Epinion, 17 of them were screened out at the end of the survey because they did not match the sociodemographic profile needed). In addition to this, four respondents did not have a survey end time, indicating that they exited the questionnaire after filling in their contact information but before formally completing the survey. This results in a final sample of 3,067 valid leads. These participants were e-mailed a link to participate in Survey 2.

6. Survey 2: Guiding participants to request their data

Out of the 3,067 valid leads, 341 (11.1%) did not have a survey start time for Survey 2. This indicates that they never started Survey 2 and leaves 2,726 respondents (88.9%) who did start Survey 2. Table 7 displays the shares of participants who did and did not enter Survey 2, divided by age, gender, and education. As the table shows, the youngest and oldest age groups were a bit less prone to start Survey 2 compared with the middle age group (88.1% and 87.3% vs. 91.2%), and men were a bit more inclined to start Survey 2 than women (90.1% vs. 87.2%). While the differences are small, chi-square tests show that they are significant [age: $\chi^2(2, n = 3,067) = 8.47, p = 0.015$] [gender: $\chi^2(1, n = 3,054) = 6.06, p = 0.014$]. The table also shows minor differences between the education groups, but a chi-square test showed that they are not significant [$\chi^2(6, n = 3,067) = 3.25, p = 0.777$].

Table 7 Drop-off between completing Survey 1 and starting Survey 2

		Completed Survey 1 but did not start Survey 2		Started Survey 2		Total	
		<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	130	11.9	966	88.1	1,096	100.0
	35–54	89	8.8	921	91.2	1,010	100.0
	55+	122	12.7	839	87.3	961	100.0
	Total	341	11.1	2,726	88.9	3,067	100.0
Gender	Female	171	12.8	1,170	87.2	1,341	100.0
	Male	170	9.9	1,543	90.1	1,713	100.0
	Total	341	11.2	2,713	88.8	3,054	100.0
Level of education	Lower-secondary	31	12.8	212	87.2	243	100.0
	Upper-secondary	54	10.9	440	89.1	494	100.0
	Vocational	63	11.1	507	88.9	570	100.0
	Short-cycle	21	8.1	238	91.9	259	100.0
	Medium-cycle	63	11.6	481	88.4	544	100.0
	Bachelor	43	11.3	337	88.7	380	100.0
	Long-cycle	66	11.4	511	88.6	577	100.0
	Total	341	11.1	2,726	88.9	3,067	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

6.1 Step-by-step guide to request data

When respondents entered Survey 2, they were informed that the second step consisted of a guide for how to request their YouTube data via Google’s Takeout solution.

Furthermore, they were informed that requesting the data would take 3–5 minutes, and they were reminded that they could get help from a telephone hotline if they ran into problems. Next, respondents were provided a link to Google’s Takeout solution and asked to log into their account. After this, they were instructed to follow three steps to order their data.

In the first step, they were instructed to deselect all data and then select “YouTube and YouTube Music”, as we only needed them to donate their YouTube data. Respondents also had the option of watching a video guide on how to do this, as displayed in Figure 8.

Figure 8 Screen shot of guide to requesting YouTube data, step 1



In the second and third steps, respondents were instructed to select JSON file format and to choose the largest possible file size. For both steps, a video guide was available to the respondents (see Figure 9). Respondents were also informed that they should only include the data types “subscriptions”, “history”, and – if possible – “comments”, and to deselect all other data categories.

Figure 9 Screenshot of guide to requesting YouTube data, steps 2 and 3

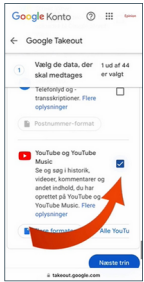


In the bottom of the three-step guide, respondents were asked whether they had requested their YouTube data. Answer categories were “yes” and “no, I need help ordering my data”.

Respondents who answered “yes” to having requested their data were presented with several follow-up questions to check that they had ordered their data correctly. First, respondents were asked whether they had made sure that they only ordered data from YouTube and YouTube Music. The screen showed an image of what box should be checked in the Google Takeout solution if they had done it correctly (see Figure 10). If respondents reported that they did not order this data or had ordered other data, they were redirected to the video guide instructing them how to order the correct data (step 1 in the guide above). After being redirected, respondents were also given the opportunity to ask for help (see below).

Figure 10 Screenshot of follow-up question, checking that respondents ordered correct data

Vi stiller dig nu en række spørgsmål for at sikre, at dine data er bestilt korrekt



Tilvalgte du kun data fra YouTube og YouTube Music?

☐ Ja

☐ Nej, jeg bestilte ikke data fra YouTube og YouTube Music

☐ Nej, jeg bestilte data fra andre steder end YouTube og YouTube Music

[Tilbage](#) [Næste](#)

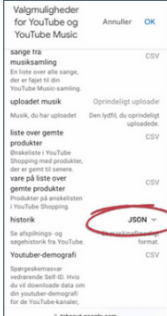
Copyright © Epinion

Respondents who confirmed that they had ordered the correct data were asked whether they had ordered the data in the correct file format (JSON) and shown a picture of what it should look like on Google Takeout if they had done it correctly (see Figure 11). If respondents reported that they did not order the data in the correct format, they were redirected to the video guide instructing them how to select the correct file format (step 2 in the guide above). After being re-directed, the respondents were also given the opportunity to ask for help (see below).

Figure 11 Screenshot of follow-up question, checking that respondents chose correct file format

Epinion
- making sense

Bestilte du data i formatet JSON?



☐ Ja

☐ Nej, jeg kom til at bestille i et andet format

[Tilbage](#) [Næste](#)

Copyright © Epinion

Respondents who confirmed they had ordered the data in the correct file format were informed that processing the request usually takes very short time and were encouraged to check their e-mail immediately, as they might already have received their data. They were also informed that, in rare cases, it could take up to five days. Finally, they were told that once they had received their data, they would be ready for the third step of the data collection process (Survey 3) and that a link for the survey had been sent to their e-mail address.

If respondents at any step of Survey 2 reported that they needed help, they were informed how they could contact Epinion via phone (in a certain time slot) or via e-mail. Respondents also had the option of asking to be called by Epinion at a specific time that suited them (see Figure 12).

Figure 12 Screenshot of help page, Survey 2

Epinion
- making sense

?

Du har skrevet, at du har brug for hjælp til at bestille data. Du kan ringe til [redacted] alle hverdage mellem 10 og 12 eller skrive en e-mail til [redacted]. Så vender vi tilbage hurtigst muligt.

Ellers kan du markere her, hvornår på dagen vi må ringe til dig og hjælpe dig med din databestilling.

☐ 09:00-12:00
☐ 12:00-15:00
☐ 15:00-18:00
☐ 18:00-20:00

Tilbage **Næste**

25%
Copyright © Epinion

6.2 Drop-off in Survey 2

Unintended by the research team at Aarhus University, Survey 2 and Survey 3 were set up so that respondents could re-enter these surveys whenever they wanted. This facilitated that respondents who initially had technical issues and needed assistance could go back to the survey and make a new attempt at requesting, downloading, and donating their YouTube data. However, if they did, their original data from these survey parts were overwritten. This means that we cannot use the variables where respondents had the opportunity to ask for help from the designated helpline to infer how many respondents made use of this (since all respondents who asked for help and successfully completed the survey steps afterwards do not figure as having asked for help in the survey data).

Moreover, the setup meant that if respondents who had successfully completed Survey 2 and Survey 3 accidentally opened the link to these surveys and pressed the “next” button on the first screen, their survey data for these survey parts were overwritten. As a result, some respondents who have in fact completed Survey 2 and Survey 3 have missing data on some or all variables in these survey parts. Some of these respondents can be identified, for example, because they have missing values on Survey 2 and/or Survey 3, but non-missing values on Survey 4 (and thus must have completed Survey 2 and Survey 3). However, it is not possible to distinguish respondents who opened Survey 2 but failed to complete it from respondents who completed Survey 2 but accidentally deleted their data by re-opening the link and never started Survey 3. The same goes for the drop-off between Survey 3 and Survey 4. This means that we cannot use the variables in Survey 2 and Survey 3 for analysing exactly at what steps of the technical process respondents dropped off. Nor can we see exactly how many respondents completed Survey 2 and Survey 3. However, we can infer approximately how many respondents dropped off in Survey 2 by comparing the number of respondents who started Survey 2 to the number of respondents who started Survey 3. Likewise, we can infer approximately how many respondents dropped off in Survey 3 by comparing the number of respondents who started Survey 3 to the number of respondents who started Survey 4 (see section 7.2.).

Of the 2,726 respondents who started Survey 2, 2,007 also have a start time for Survey 3. This indicates that 719 respondents dropped off during Survey 2 (26.4%). Table 8 displays how the drop-off is distributed across age, gender, and education. Looking first at age, more than four in five (82.5%) of the youngest participants who started Survey 2 also started Survey 3, indicating that less than 20 per cent of the youngest participants dropped off during the first technical part of the data collection where respondents had to request their YouTube data. Unlike this, only a little more than half (56.9%) of the oldest respondents who started Survey 2 also started Survey 3. This indicates that more than 40 per cent of the oldest respondents dropped off during the first technical part of the data collection process. The data does not show why respondents drop off, but the age difference could likely be explained by younger parts of the population being more technically confident and skilled. A chi-square test showed that the association between age and dropping off in Survey 2 is significant [$\chi^2(2, n = 2,726) = 177.62, p < 0.001$].

Table 8 Drop-off in Survey 2

		Started Survey 2 but not Survey 3		Started Survey 3		Total	
		<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	169	17.5	797	82.5	966	100.0
	35–54	188	20.4	733	79.6	921	100.0
	55+	362	43.1	477	56.9	839	100.0
	Total	719	26.4	2,007	73.6	2,726	100.0
Gender	Female	311	26.6	859	73.4	1,170	100.0
	Male	407	26.4	1,136	73.6	1,543	100.0
	Total	718	26.5	1,995	73.5	2,713	100.0
Level of education	Lower-secondary	59	27.8	153	72.2	212	100.0
	Upper-secondary	77	17.5	363	82.5	440	100.0
	Vocational	150	29.6	357	70.4	507	100.0
	Short-cycle	68	28.6	170	71.4	238	100.0
	Medium-cycle	176	36.6	305	63.4	481	100.0
	Bachelor	71	21.1	266	78.9	337	100.0
	Long-cycle	118	23.1	393	76.9	511	100.0
	Total	719	26.4	2,007	73.6	2,726	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

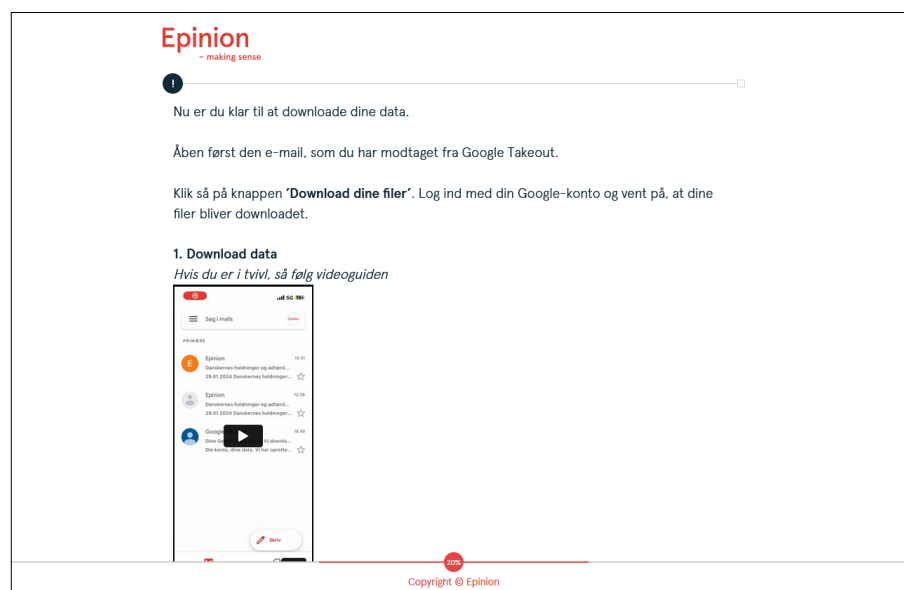
The table shows no differences between men and women in terms of dropping off in Survey 2 [$\chi^2(1, n = 2,713) = 0.014, p = 0.905$]. There are, however, substantial differences between some of the education groups. While a little less than two thirds (63.4%) of those with a medium-cycle higher education got through Survey 2 and started Survey 3, more than four in five (82.5%) of those with an upper-secondary education did so. A chi-square test showed that the association between level of education and drop-off is significant [$\chi^2(6, n = 2,726) = 54.93, p < .001$]. Yet, it should again be pointed out that there is no clear pattern where lower or higher level of education is related to higher drop-off rates.

7. Survey 3: Guiding participants to download and securely donate their data

Entering Survey 3, respondents were welcomed to the third part of the study for downloading and uploading their data and were again informed about the hotline in case they faced problems.

The respondents were then asked whether they received an e-mail from Google Takeout to the e-mail address that is linked to their Google account. If they negated, they were requested to check their e-mail and informed that there could be a potential time gap between data request and receiving the e-mail. They were asked whether their data request dated more than five days back, and if they affirmed, they were asked to request their data again; if they negated, they were informed again that it can take up to five days to receive the e-mails and to check their e-mails regularly. Once participants affirmed that they had received the e-mail, they were asked which device they planned to use for participating in the study. This was asked to allow customised instructions for how to download their data based on the system they were using. Next, they were instructed how to download the data; parts of this procedure are illustrated in Figure 13.

Figure 13 Instructions for downloading the YouTube files



Comments: The figure shows the first screen of the procedure.

In the next step, participants were asked to rename their downloaded zip files. With this procedure, Epinion ensured that it was possible to link the donated data with the survey data at the level of individual participants. The name that the file should be given was provided in bold letters in the instructions and a video guide was also available. Figure 14 shows the first screen of this procedure. The renaming of the files was validated in an additional question. In case respondents faced problems renaming the file, they were informed that the technical hotline would get in touch with them. In

case they affirmed they had renamed the downloaded file correctly, they were allowed to continue with the third step to upload their data.

Figure 14 Instructions for renaming downloaded data file



Comments: The above example is the first screen of the procedure and created for purposes of documentation; it does not indicate actual participation.

In the third and final step where respondents had to upload their data, they received instructions with a secure link for data upload and additional video guidance if needed. They were also informed again with whom they share the data and were then asked whether they had successfully shared the data via the secure link. If they negated, they were informed that the technical support would contact them for help. If they affirmed that they had successfully uploaded the data, they were informed that Epinion would check their uploaded data and that they – once data was validated – would receive an invitation to the final part of the study (Survey 4) via e-mail.

7.1 Applied data validation criteria

Once participants uploaded their data, the data was validated by Epinion. Participants passed validation if they uploaded the correct file type (JSON), named the files as instructed, included the correct documents in the upload (uploading the watch history was a requirement), and the file appeared to be from the correct person (match between the e-mail addresses linked to the uploaded data and the one registered at Epinion). Furthermore, validation required that the data met certain defined criteria for the content of the data. First, the watch history must include at least 26 entries within the last year (meaning biweekly YouTube usage, on average). Second, at least one of the uploaded documents must prove that participants used YouTube earlier than 2021 (at least one entry within the files donated by a respondent – i.e., in the comment, watch, or search file – must be from before 1 January 2021). The latter

criterion deviates from the one used in Survey 1, where respondents were screened out if they reported having started using YouTube in 2020 or later. The self-reported criterion applied in Survey 1 was stricter than the one applied in the data validation to minimise the risk that some respondents who reported adequate YouTube usage in the screener would end up being screened out at the point of the data validation, after having invested time and effort to donate their data. In case the applied criteria for the content were not met, respondents were informed that they did not match the target group of the study. If respondents did not meet the format criteria (e.g., the naming of the files), the support hotline tried to solve the technical issues. If the cases could not be solved, respondents were screened out of the study due to not passing validation.

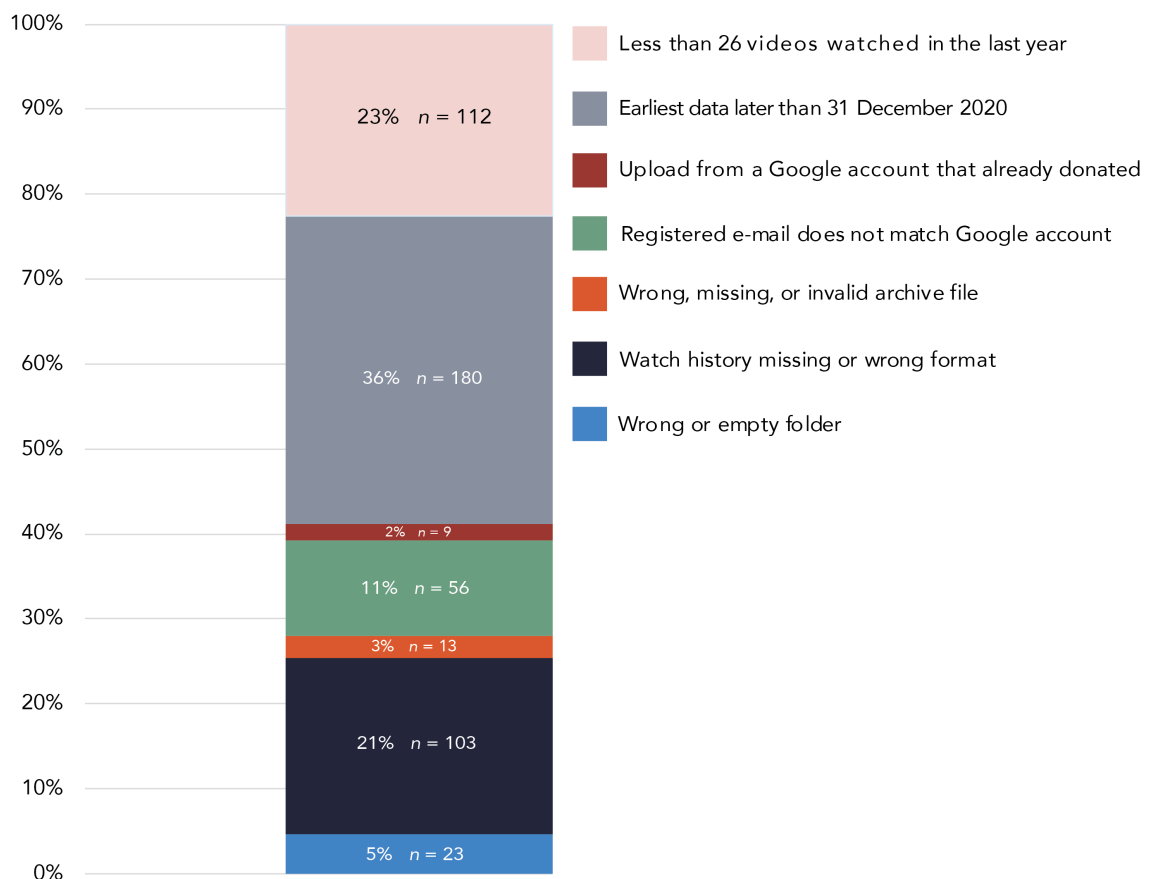
The criteria described above were applied throughout most of the fieldwork period. However, as earlier described, we started the fieldwork period with a pilot period (from 15 August–20 September 2024). At the point of data collection, to our knowledge no publications have addressed data donation studies at national scale in the Danish context, and those that applied scalable approaches in other countries (Wedel et al., 2024; Welbers et al., 2024) implemented different study designs (recruitment via a survey first). Other data donation studies (e.g., van Driel et al., 2022; Breuer et al., 2023) did not recruit a large number of participants to their study. Thus, in this context, we implemented a pilot period due to this limited pre-knowledge about how participants would respond to invitations to a data donation study in a Danish context, making reliable predictions about the fieldwork difficult.

In this pilot period, more strict validation criteria were applied – namely weekly YouTube usage, reflected in at least 52 entries in the watch history in the last year, and requirement that the watch history must reach back to before 1 January 2020 (matching the stricter criterion applied in Survey 1). In addition, respondents had to report weekly YouTube usage (rather than biweekly usage) in the screening question in Survey 1. After an evaluation of the pilot period, these criteria proved too strict – as many did not meet the validation criteria as briefly mentioned earlier. For example, inspection of the donated data showed that some participants who reported having started using YouTube before 2020 in Survey 1 had no entry reaching back to before 2020 in the watch-history file; however, their claim was proved by entries in the search or comment file. In addition, some respondents had rich prior histories, but sparse recent histories, and we wanted to account also for this kind of using pattern. These observations led to an adjustment of the criteria on 20 September 2024, resulting in the criteria accounted for above, where the pilot period was followed by the actual fieldwork period. This change can be seen in the number of completions over the fieldwork period (see Figure 4). Here, we see that completions increased once validation criteria were changed to the less strict criteria. To allow analysis of the data collected in the pilot period, the participants who failed the strict validation criteria in the pilot phase were re-invited to participate based on the new criteria, and they could continue participation in case they already succeeded in partly completing single steps of the study and being validated according to the new criteria. Thus, all who were invited over the full data collection period and fulfilled the new criteria had the chance to participate in the study.

As noted, Epinion validated the donated data based on the described criteria. We now turn to analysis of the reasons why some participants failed data validation. In general,

the reasons were documented in seven categories that can be broadly divided into two areas – data validation failed due to missing or incorrect file upload or data validation failed due to the actual content of the donated data not matching the criteria. Figure 15 provides an overview of the portions falling into each category for respondents who failed data validation.

Figure 15 Reasons for failing applied validation criteria for uploaded data registered by Epinion



Comments: N = 496

Source: Epinion (Layout: Aarhus University)

According to Epinion, 496 respondents passed the self-reported screening criteria and uploaded data but failed the applied data validation due to the documented seven reasons. Most respondents failed validation because they had not watched enough videos in the last year (23%), had no data entry from before 1 January 2021 (36%), or had not uploaded a watch-history file or, if so, in the wrong file format (21%). Other reasons include that they already donated data with the registered e-mail address (9 cases, which are treated as duplicates and not included in the analysis of screening out and drop-off), that they donated data from a Google account that did not match the e-mail address registered by Epinion, or that they did not upload a correct archive file, which is a file automatically generated and added to the data download package in the

downloading process and which was required for account validation by Epinion (e.g., for the identification of duplicates). A final reason for failing validation was that respondents only uploaded an empty folder or a folder in the wrong format, hindering the further processing of the files.

Although the collected data does not allow for an empirical analysis of *why* the 496 respondents passed the self-reported screening criteria used in Survey 1 but failed Epinion's data validation based on the applied criteria, there are various potential explanations for this discrepancy. Possible explanations include 1) that in Survey 1 respondents provided incorrect answers to the screening questions; 2) that participants use multiple accounts and, for example, reported on overall usage while the account they donated data from does not reflect this usage; and 3) that participants indeed had used YouTube far enough back in time to qualify but had since then switched to a new account, meaning that the data from the new account would not reflect usage going far enough back in time. In the first scenario, participants might recall their YouTube usage incorrectly. In the second, they might report correctly but their report cannot be confirmed because multiple accounts exist, and they donated data from an account that does not prove their report. In the third scenario, the reports cannot be confirmed due to a change in the account and the new account not matching the reported YouTube usage.

7.2 Drop-off in Survey 3

Of the 2,007 respondents who had a start time for Survey 3, 1,125 also had a start time for Survey 4. This indicates that 882 respondents (43.9%) dropped off during Survey 3 (including respondents with non-valid uploads). It is noticeable that the process of downloading and donating YouTube data thus seemingly caused higher drop-off rates than the process of requesting the data (where 26.4% dropped off). This is especially the case considering that the respondents who started Survey 3 are respondents who had sufficient technical skills to complete Survey 2. Table 9 displays how the drop-off is distributed across age, gender, and education groups. Again, there is a clear age difference. Close to two thirds of the youngest age group (63.7%) managed to go through Survey 3 and begin Survey 4, compared with less than half of the oldest age group (40.3%). While the data does not show why respondents dropped off, the difference could likely be due to lower technical skills among the older part of the population. A chi-square test showed that the association between age and drop-off is significant [$\chi^2(2, n = 2,007) = 68.57, p < 0.001$].

Table 9 Drop-off in Survey 3

		Started Survey 3 but not Survey 4		Started Survey 4		Total	
		<i>n</i>	%	<i>n</i>	%	<i>N</i>	%
Age	18–34	289	36.3	508	63.7	797	100.0
	35–54	308	42.0	425	58.0	733	100.0
	55+	285	59.7	192	40.3	477	100.0
	Total	882	43.9	1,125	56.1	2,007	100.0
Gender	Female	397	46.2	462	53.8	859	100.0
	Male	479	42.2	657	57.8	1,136	100.0
	Total	876	43.9	1,119	56.1	1,995	100.0
Level of education	Lower-secondary	75	49.0	78	51.0	153	100.0
	Upper-secondary	151	41.6	212	58.4	363	100.0
	Vocational	161	45.1	196	54.9	357	100.0
	Short-cycle	80	47.1	90	52.9	170	100.0
	Medium-cycle	145	47.5	160	52.5	305	100.0
	Bachelor	113	42.5	153	57.5	266	100.0
	Long-cycle	157	39.9	236	60.1	393	100.0
	Total	882	43.9	1,125	56.1	2,007	100.0

Comments: The total *N* for gender is smaller because respondents who reported “other” to the question of gender were not included in the analysis.

Turning to gender, slightly more men (57.8%) than women (53.8%) went through Survey 3 and began Survey 4, but a chi-square test showed that this difference is not significant [$\chi^2 (1, n = 1,995) = 3.26, p = 0.071$]. Some differences can also be observed across education groups but a chi-square test showed that they are also not significant [$\chi^2 (6, n = 2,007) = 7.65, p = 0.265$].

8. Survey 4: Demographic distribution and duration

After participants' uploaded data were validated by Epinion, they received an e-mail invitation to the last part of the study – Survey 4, which they could access via a personalised link. In the invitation, they were informed that the survey would take 15–20 minutes.

In total, 1,068 respondents completed this survey. Four participants were identified as “speeders”, leaving a sample of 1,064 valid completes. A speeder is defined as completing the survey faster than the median response time minus two standard deviations (5 minutes and 16 seconds) among the 95 per cent slowest response times; another applied definition is that respondents completed the survey faster than the median response time minus one standard deviation (9 minutes and 59 seconds) and selected “don't know” answers for more than 25 per cent of the questions.

The median time for completing Survey 4 (excluding the four speeders) was 16.5 minutes. The median time for participants completing Survey 4 for participation in the study (from starting Survey 1 to completing Survey 4) was 2.3 days.

Table 10 displays a comparison of the sociodemographic distribution of the 1,064 valid completes and the target population of adult Danish YouTube users (as estimated in the survey study conducted by Epinion; see section 4.6).

The largest deviance from the target population can be observed for the oldest age group (55+), which is underrepresented by 11 percentage points in the final sample. The 18–34-years-olds are overrepresented by 6 percentage points in the final sample, and the 35–54-year-olds are overrepresented by 5 percentage points in the final sample. Furthermore, male participants are overrepresented by 4 percentage points in the final sample. With regards to education, the differences are minor.

Table 10 Demographic distribution of completes in comparison to Epinion study of the Danish YouTube user population

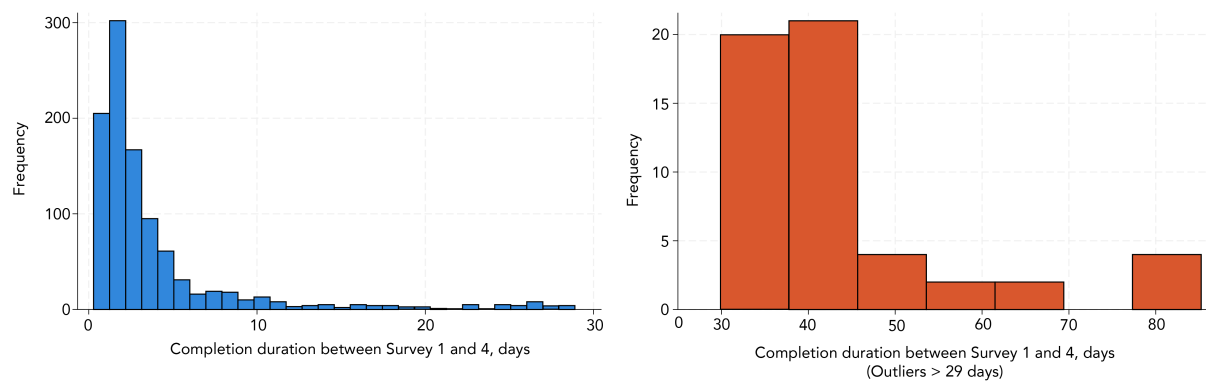
		Completed Survey 4		Epinion population study (%)	Difference (percentage points)
		Frequency	%		
Age	18–34	481	45.2	39	6
	35–54	402	37.8	33	5
	55+	181	17.0	28	-11
	Total	1064	100	100	
Gender	Female	439	41.5	45	-4
	Male	619	58.5	54	4
	Total	1058	100	99	
Level of education	Lower-secondary	72	6.8		
	Upper-secondary	202	19.0		
	Vocational	188	17.7		
	Short-cycle	86	8.1		
	Medium-cycle	150	14.1		
	Bachelor	146	13.7		
	Long-cycle	220	20.7		
	Total	1,064	100		
Broad education categories	Lower-secondary or vocational training	260	24.4	25	-1
	Upper-secondary school or short-cycle higher education	288	27.1	26	1
	Intermediate and long- term higher education	516	48.5	49	0

Comments: Information about the population study from Epinion: $n = 488$, and gender does not sum to 100% because 0.3% identified as non-binary (not shown in table) and because of rounding of decimals.

8.1 Completion times

Next, we describe the completion time of participation in the full study (starting Survey 1 and ending Survey 4), completion time for Survey 1, and completion time for Survey 4 in more detail. The distributions of completion times for each are shown in Figures 16, 17, and 18.

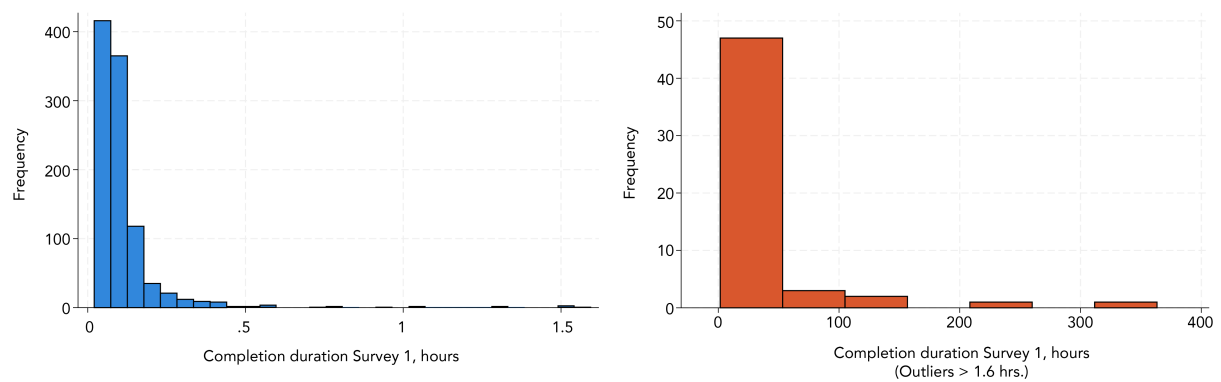
Figure 16 Completion duration between Survey 1 and Survey 4



Comments: The figure indicates only respondents who completed Survey 4: the quickest 95% (left) and the slowest 5% (right).

Source: computations by Aarhus University

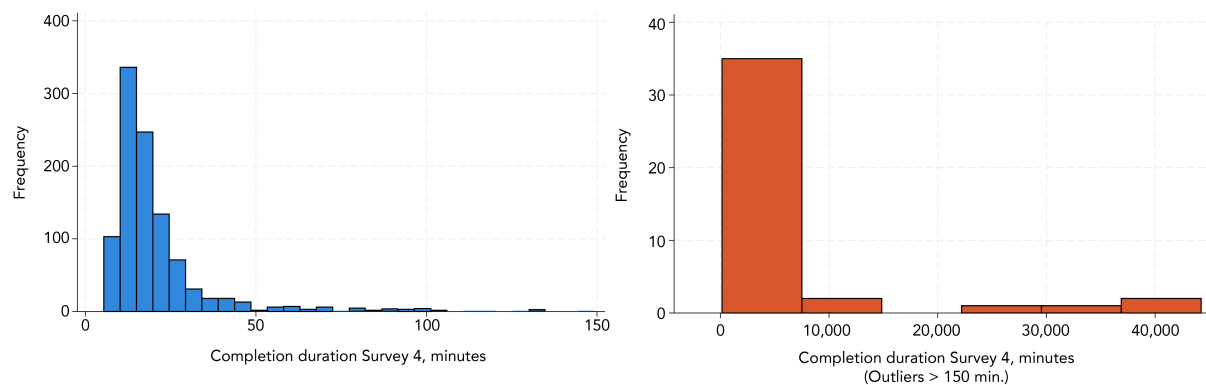
Figure 17 Completion duration Survey 1



Comments: The figure indicates only respondents that completed Survey 4: the quickest 95% (left) and the slowest 5% (right).

Source: computations by Aarhus University

Figure 18 Completion duration Survey 4



Comments: The figure indicates only respondents that completed Survey 4: the quickest 95% (left) and the slowest 5% (right).

Source: computations by Aarhus University

We compared the completion duration for gender, education, and age groups using two-sample *t*-tests with unequal variances. For the completion duration between starting Survey 1 and ending Survey 4 (i.e., completing the full study), no gender differences or age differences can be found. For education, the results show that for the participants who completed Survey 4, respondents belonging to the lowest education category spent significantly less time than respondents belonging to the middle or highest education category ($p < 0.05$). This difference cannot be found if only the quickest 95 per cent of the respondents are analysed.

Comparing completion durations for Survey 1 (starting Survey 1 and ending Survey 1) across demographics revealed that within the sample of participants who completed the full study, the youngest age group spent significantly less time completing Survey 1 than the middle and oldest age groups ($p < 0.05$). Other differences by gender or education were not significant.

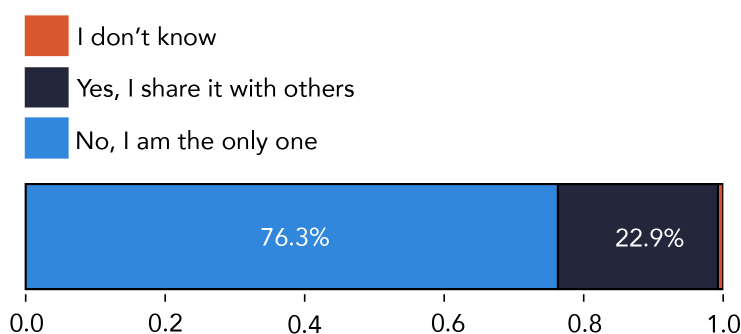
Regarding the duration of Survey 4 (starting Survey 4 and ending Survey 4), results showed that the oldest age group spent significantly less time than respondents in the other age groups ($p < 0.05$). This difference cannot be found if only the quickest 95 per cent of the respondents are analysed. This implies that the age difference is mainly driven by participants in the younger age groups who spent exceptionally long times.

9. The donated data files

The purpose of Survey 3 was to receive participants' YouTube-history files. Unlike data donation approaches that rely on tracking apps on smartphones or browser plugins, our method yields insights about platform usage across devices, apps, and browsers.

Participants' YouTube use is recorded as long as they are using a device or browser that is logged in to their Google account. This means, that their use is also recorded if they are not logged in with the goal of using YouTube, but rather their Gmail, Google Drive, or other services accessed through a Google account (logged in is default and the user actively needs to log out). Despite this advantage of the method, it also holds some limitations. On the one hand, participants could only donate the data of one account; thus, if a user uses two or more Google accounts, their donated data is necessarily incomplete, as it only reflects use recorded on one of the accounts. On the other hand, users may share usage of devices they are logged into with friends or family, such that their watch history represents more than just their own YouTube use. Most participants indicated that others do not use YouTube on the same device as themselves. Figure 19 shows their self-reported answers to shared use of account (asked in Survey 4) and illustrates that this limitation is only present for 23 per cent of the participants who donated valid data. Only 1 per cent indicated that they do not know.

Figure 19 Proportion of participants who state that others do or do not use YouTube on the same device as them



Comments: Question: Are there others who use YouTube on the same device as you?

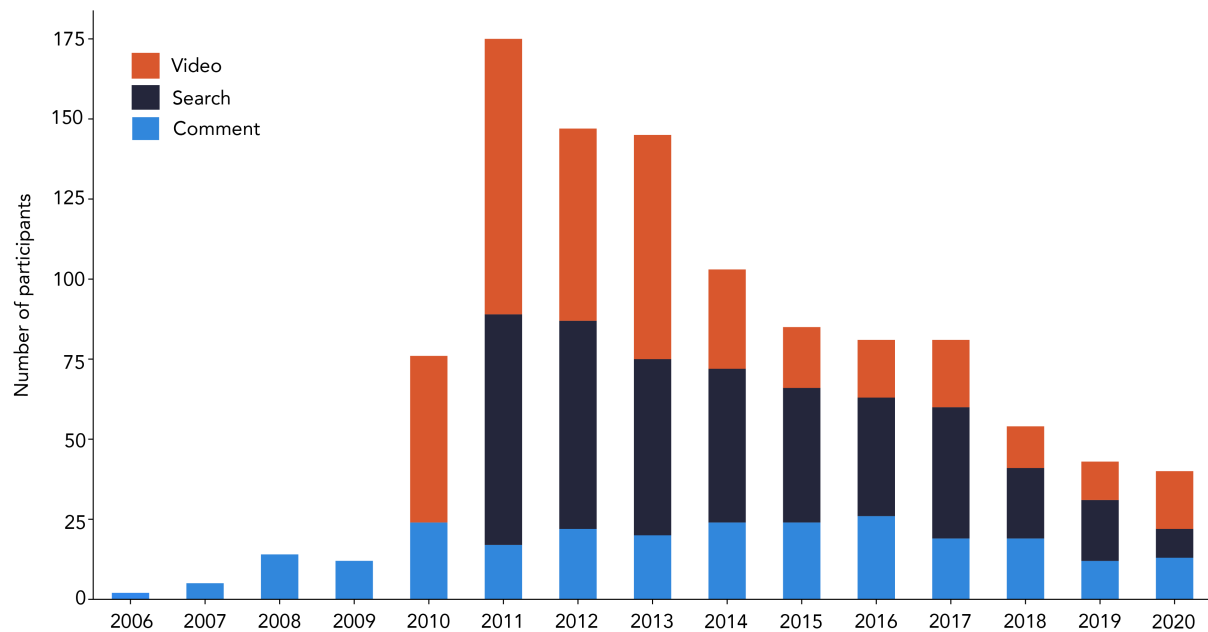
As a further limitation, YouTube videos can be reposted or embedded beyond the confines of the platform itself, for example, on other social media platforms such as Facebook, and participants' encounters with such off-platform videos are not reflected in their donated data.

YouTube provides data on different kinds of activity on the platform in separate datasets. Among others, users can request a log of their account's watched videos (watch history), search queries typed into YouTube's search bar (search history), comments posted (comment history), and channels subscribed to. We considered these data the most important regarding the scope of our study and, in accordance with data minimisation principles and the General Data Protection Regulation (GDPR), we limited the data we requested from our participants to these data types. Due to our focus on the content that Danes watched on YouTube, we asked participants to donate

their watch, search, comment, and subscription data but only required watch data in the validation process. Thus, in the validation process, we did not exclude participants who did not donate search, comment, or subscription data (see also section 7.1).

In general, the received logfiles of the study vary greatly in number of entries across the participants. As detailed previously, one of the objectives of the data collection was gathering data that allows analyses of YouTube usage over time. This is also reflected in the applied inclusion criteria (see a detailed description in sections 5 and 7.1). To provide an overview of the data received, we identified the earliest entry among each participant's donated files. Figure 20 illustrates the distribution of earliest entries in the participants' history files categorised by their origin: watch history, comment history, or search history. Furthermore, the figure reveals that the search, watch, and comment data YouTube provides upon a user's data-takeout request vary in terms of how far they extend back in time. Comment data extend all the way to 2006, the year after YouTube's creation. The earliest watch-data entries date back to 2010, and the earliest search-data entries to 2011. We cannot draw definitive conclusions as to why this is. However, the lack of a gradual increase in search and watch data suggests that the data has either not been logged or has been post hoc truncated.

Figure 20 Distribution of participants' oldest donated data point by year of creation



Among the participants who completed the full study, four were excluded from our sample because the analysis showed that they did not have any data entries prior to 2021 (which was required; see section 7.1). One further participant's watch-history data did not contain video IDs or URLs. Accordingly, this participant, too, was removed from further analysis. As Table 11 shows, all remaining participants have donated watch-history data (since this was a criterion for inclusion). Of these, 96.3 per cent had used YouTube search and donated their search data, 90.4 per cent were subscribed to one or more channels and donated subscription data, and 65.0 per cent had commented on either a video or a community post and donated their comment data. Due to the methodological limitations of the data donation approach, we are unable to tell whether participants who did not

submit all types of data withheld those data intentionally, coincidentally, or because they did not have these data in the first place (e.g., due to never having used YouTube’s comment function). However, we evaluate intentional deletion or refraining from uploading to be unlikely, as we did not at any point in the process indicate that it was voluntary to upload these files to participate and, furthermore, that it would require manually sorting out files from the uploaded folder.

Table 11 Data donation yield by type of donated data

	Donating participants		Data entries	
	<i>n</i>	%	<i>N</i>	Unique data entries
Watch history	1,063	100	18,828,115	7,358,698
YouTube videos	1,063	100	15,290,385	6,993,570
YouTube music	457	43.0	657,784	166,921
Advertisements & Other	978	92.0	2,879,946	201,658
Search history	1,027	96.6	5,580,756	1,625,093
Searches	1,024	96.3	2,609,670	1,449,625
Subscriptions	961	90.4	78,965	46,403
Comments	691	65.0	29,707	29,707

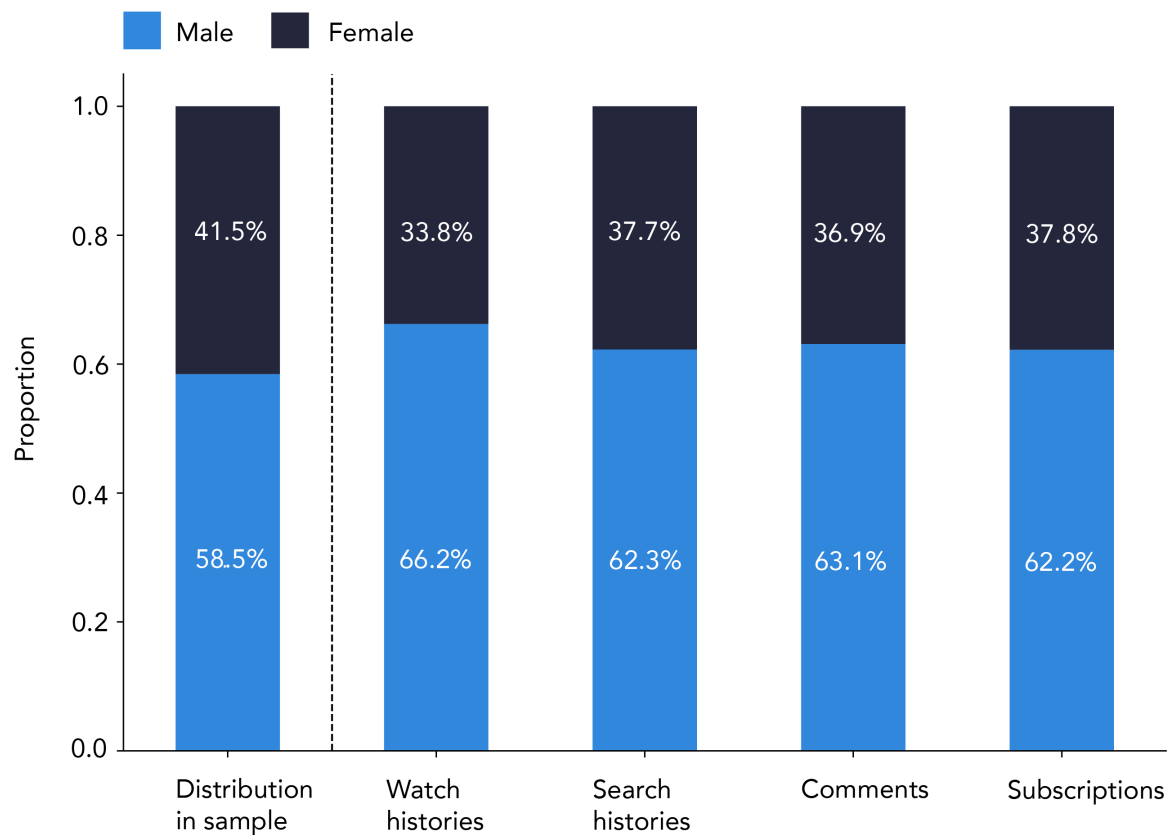
Comments: Three participants donated search-history files but there were no specific search queries in the files.

9.1 Sociodemographic profiles and the donated data

Gender

Participants who self-identified as male donated more data per person across all four data categories (see Figure 21 for an overview). Due to the low number of participants indicating their gender as “other” ($n = 6$), these are not depicted.

Figure 21 Distributions of contribution to the donated datasets divided by gender

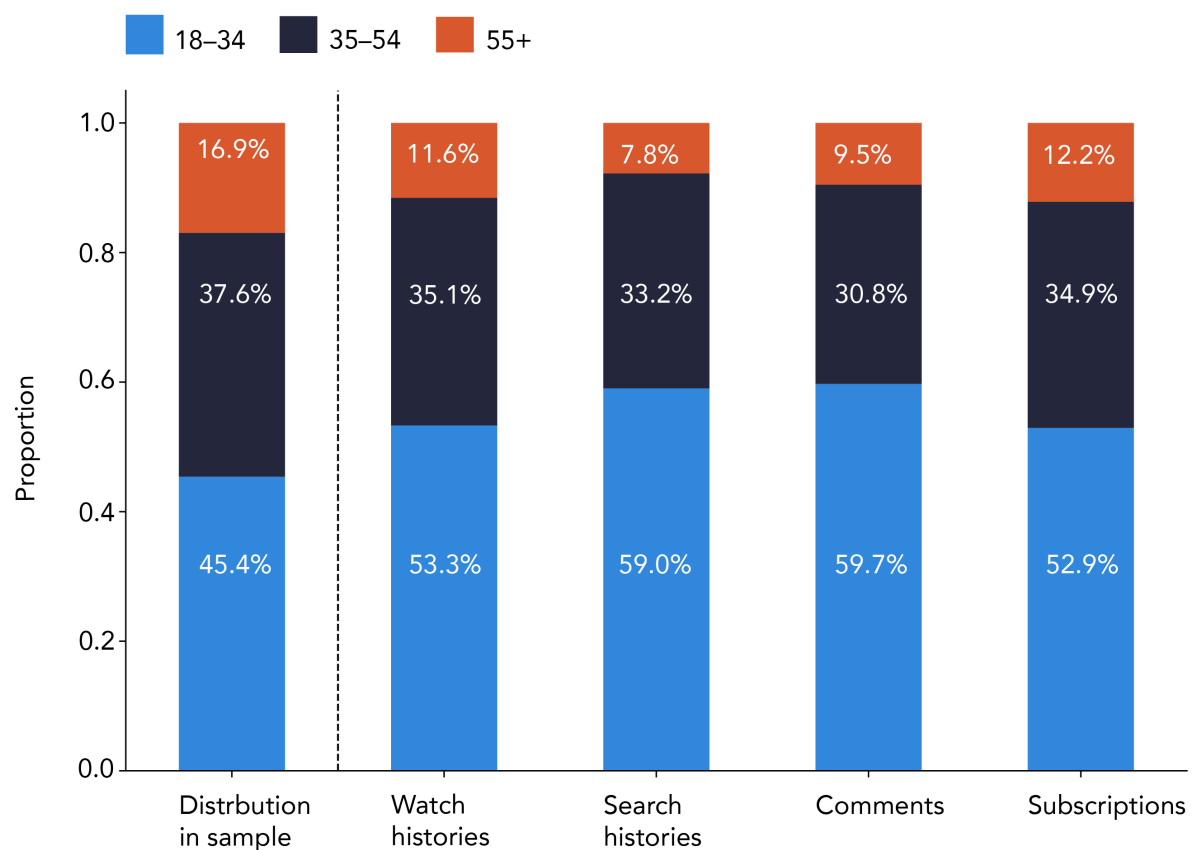


Comments: Left bar displays overall gender balance among all valid data donors ($n = 1,063$). In service of interpretability, six respondents who identified their gender as “other” have been excluded from the visualisation, leading to $n = 1,057$.

Age

Age as a demographic factor impacts the number of entries in relation to YouTube usage in our sample to a larger extent. As Figure 22 shows, entry points from the youngest group of participants make up a disproportionately large share of the donated data, which indicates that they use the platform more than the older age groups. This tendency is general across all data types but is particularly visible in the search and comment data.

Figure 22 Distribution of contribution to the donated datasets by age

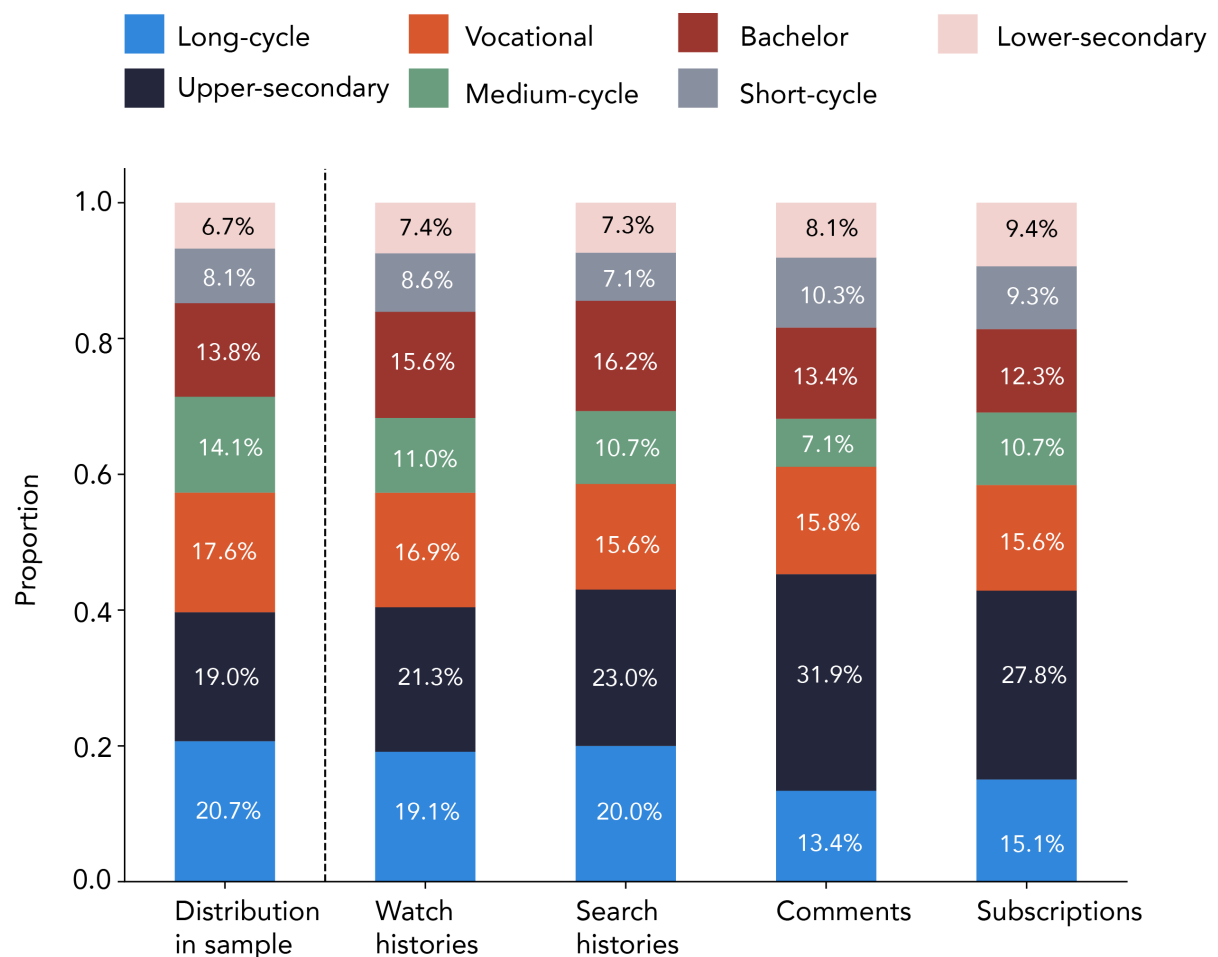


Comments: Left bar displays overall age balance among all participants.

Education

There are also differences between education categories among the participants regarding the number of data points they donated. Figure 23 illustrates that participants in the group with upper-secondary education contributed disproportionately more comment and subscription data points when compared to the general strata balance of the study (left bar). Note, however, that our pool of comment and subscription data is thinner than that of search and especially watch data (see Figure 20). Thus, peculiarities of individual participants stand to have a comparatively large effect on the overall distribution. Note also that the share of watch and search data donated by Bachelor-educated participants is substantially larger than their overall share in the set of participants and, conversely, that the medium-cycle educated participants have donated disproportionately less data across all data types.

Figure 23 Distribution of contribution to the donated datasets by education level



Comments: The left bar displays overall education balance among all valid data donors ($n = 1,063$).

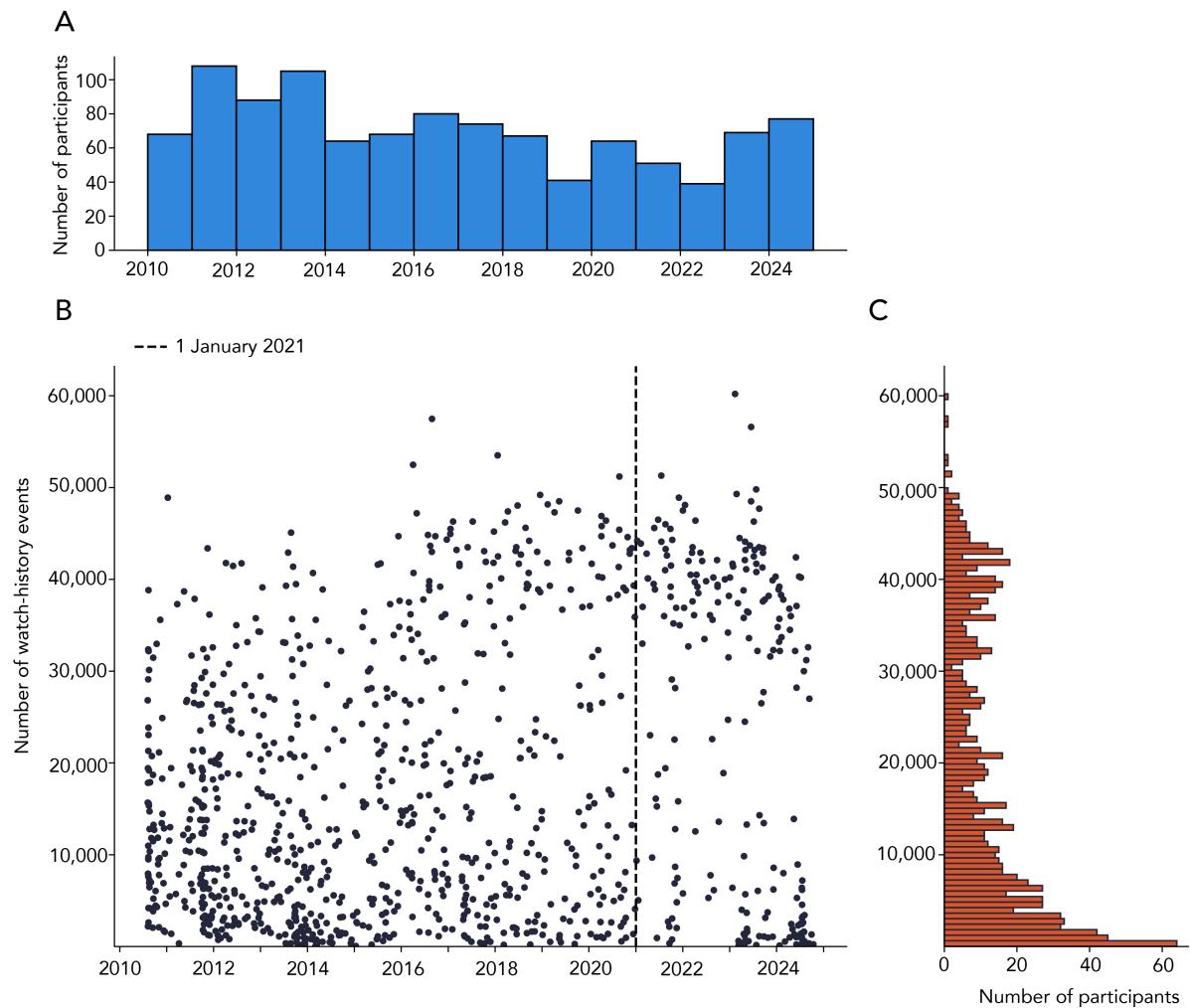
9.2 Viewing data

For a user requesting their watch history, YouTube generates a JSON file that contains a structured list of content requests that were made from that user's account to YouTube. These requests include videos, but also songs requested from the YouTube Music service, advertisements, and others (see Table 11 for an overview of their proportion in the data we obtained).

Each of the content-type requests features a date and time, indicating when the request was made. As such, each user's watch history is not just a representation of the content that the user selected and watched. On the one hand, they might not actively select the content, as in the case of algorithmically selected advertisement clips and content YouTube plays automatically, when playback begins after a prior video ended. On the other hand, the data exclusively contain start times, leaving uncertainty regarding whether a user has seen a video in full, partially, or interrupted playback right away.

There was substantial variation in the number of entries per watch history and in the timespan covered by the watch histories. Figure 24 illustrates this diversity. The age of the watch histories we obtained ranges from over ten years to a few months (see subplot A). There is a bias towards watch histories that contain comparatively few entries; however, there are also numerous extensive watch histories, some including more than 50,000 entries (see subplot C). On average, each watch history contains 17,712 entries. There seems to be no determining relationship between watch-history age and the number of its entries. This is illustrated in subplot B, which plots watch-history age against the number of their entries. It is worth noting that this is not an expected distribution from a representative sample of YouTube users. Our recruitment strategy and inclusion criteria, we believe, biased our data towards heavy YouTube users.

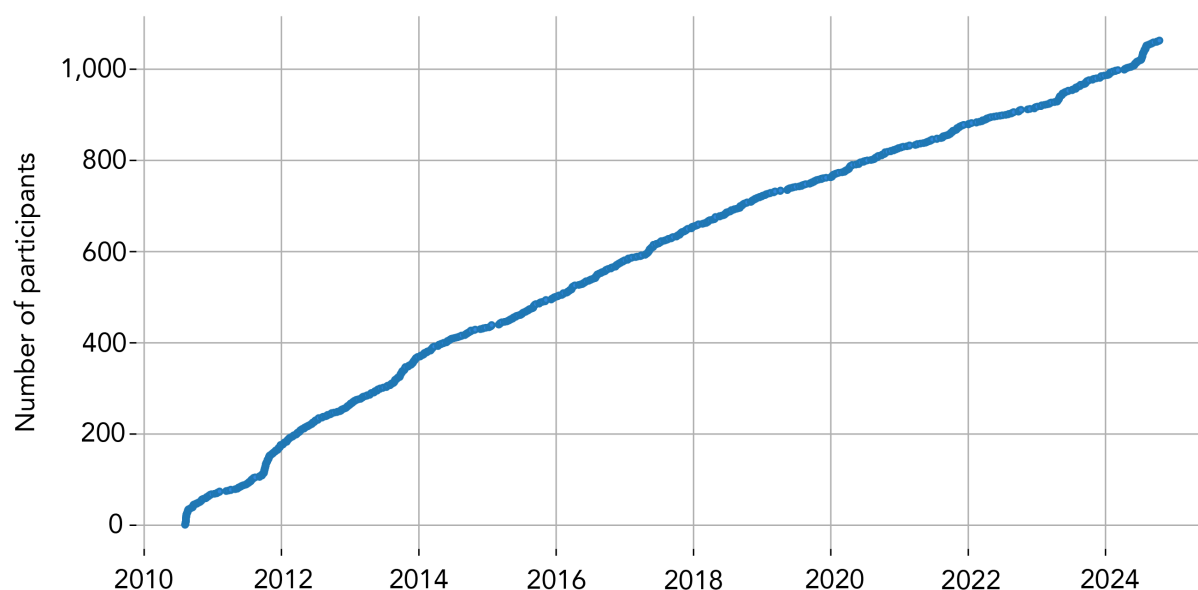
Figure 24 Overview of the donated watch-history data by watch-history age



Comments: A: Histogram of the distribution of earliest watch-history entries since the earliest entry in 2010. B: Scatterplot of the earliest watch-history entries plotted against the total number of entries in that participant's watch history (a dot represents one participant); participants were only included if their donated data was validated (with data going back to at least before 1 January 2021), indicated by the dashed line. C: Histogram of the distribution of number of watch-history entries across all watch histories.

The cumulative distribution of earliest watch-history entries (see Figure 25) does not feature substantial changes in slope and indicates a fairly even distribution of earliest data entries across participants between 2012 and 2024. Note that regardless of start date, all watch histories were required to feature recent activity to ensure collection of recent YouTube use data.

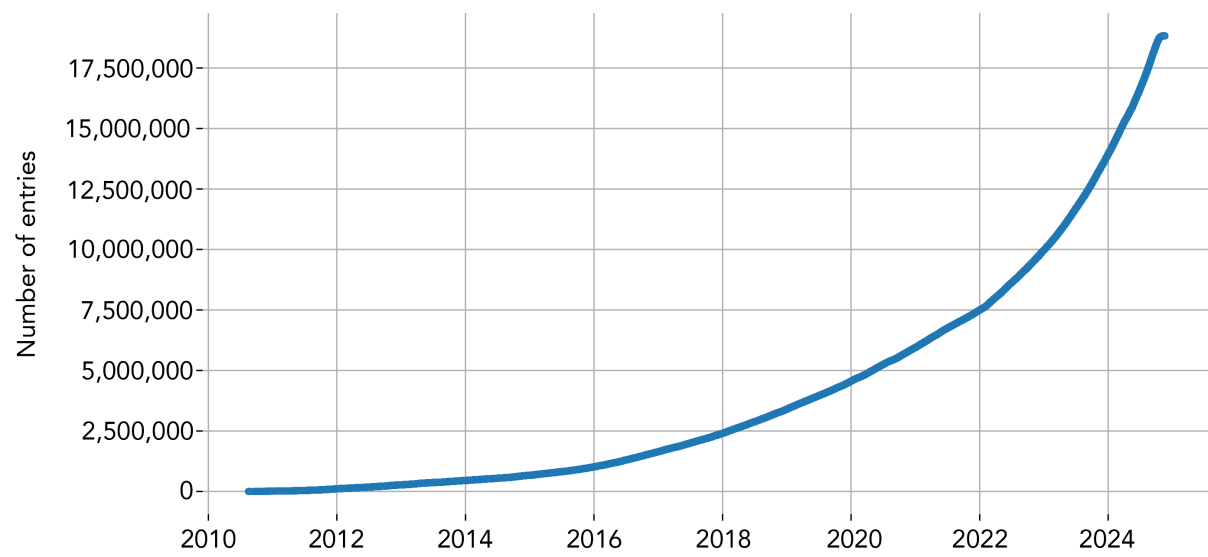
Figure 25 Cumulative distribution of earliest watch-history entries by participants over time



Comments: The curve does not feature substantial changes in slope, indicating an even distribution of watch-history start dates among the donating participants.

There is a heavy bias towards recent data among the donated data (see Figure 26). The reasons for this bias are likely manifold. For one, there were inclusion criteria that required eligible data donations to feature a certain number of recent data entries, while no requirements regarding the number of older data entries were made. Beyond that, YouTube users may have deleted or lost access to older YouTube accounts and replaced them with a new one. These biases have implications for how the data can be used in analyses. For one, analyses of past YouTube use draw on a diminishing number of participants and data points the further back in time they reach; for example, there are roughly half as many watch histories covering 2016 as there are covering 2024. Furthermore, the lower number of watch-data points in the past lead to an imbalance between the time periods if, for example, past and recent watching behaviour are compared. Roughly one third of all watch-history entries derive from the last two years before the survey was conducted.

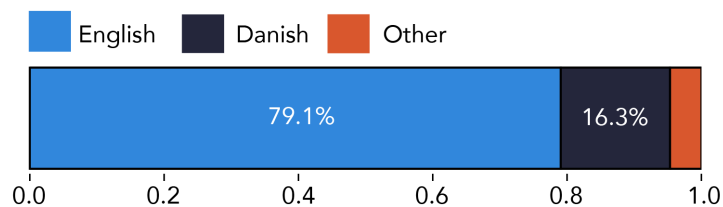
Figure 26 Illustration of the cumulative distribution of all watch-history entries over time



Distribution of languages

We estimated the distribution of languages in the obtained dataset with computational spoken language detection in a random sample ($n = 20,180$) of all watched videos, detected language in 54 per cent of them, and found that, with nearly 80 per cent, English is the dominant language. Danish, by contrast, makes up around 16 per cent of the videos with language (see Figure 27). These numbers are based on a stratified random sample of 20 videos per participant, taken from the last five years of their watch histories ($n = 1,009$; not all valid participants' watch histories featured sufficient videos in the relevant period). The videos were analysed using WhisperAI (github.com/openai/whisper). Not all videos feature spoken text. Furthermore, music, background noise, or low recording quality can make language detection difficult. Thus, we disregarded all videos where WhisperAI's confidence in the language detection was below 98 per cent (reducing the sample further to $n = 10,956$) – a threshold that reduced false language detection below 5 per cent in internal validation based on 200 randomly selected validated videos. WhisperAI is likely better at confidently recognising English than Danish, thus rendering underrepresentation of the Danish language here more likely. However, the obtained results align with the findings of our validation based on 200 randomly selected videos, where we found 14 per cent to be in Danish.

Figure 27 Distribution of detected languages among the videos in the donated watch histories



Comments: From the watch history of each participant who had sufficient videos in the past five years of their watch history ($N = 1,009$), a stratified random sample of 20 videos was taken. WhisperAI was used to determine the language of the resulting dataset of 20,180 videos. The figure depicts the language distribution of 10,956 videos where WhisperAI's language detection confidence was $> 0.98\%$.

The videos viewed the most by the participants

As Table 11 illustrates, participants viewed 15,290,385 videos, of which 6,993,570 were unique (determined by video URL). Thus, the donated data shows that participants at large do not view the same videos but rather have a unique video diet. According to our dataset, YouTube viewing consists of a lot of niche material only watched by a few participants. Table 12 shows the videos that are watched most frequently in the dataset.

Table 12 The 10 most frequently listed videos in the dataset

Rank	Video title	Number of entries	Number of participants who requested the video
1	Tommelfinger, tommelfinger, hvor er du Børnesang med fagter Popsi og Guitar-Krelle	2,718	167
2	Mike Rowland-Silver Wings Part 1.	2,414	1
3	Mike Rowland-Silver Wings Part 2.	2,369	1
4	Rejs dig op og sæt dig ned Musikmotorik Børnesang med fagter Popsi og Guitar-Krelle	2,335	136
5	Mike Rowland-Silver Wings Part 3.	2,301	1
6	Mike Rowland-Silver Wings Part 4.	2,194	1
7	Mike Rowland-Silver Wings Part 5.	2,153	1
8	Mike Rowland-Silver Wings Part 6.	2,096	1
9	Hjulene på bussen Børnesang med fagter Popsi og Guitar-Krelle	2,011	138
10	5. Fem små aber - Sangskattekisten	1,776	245

The list in Table 12 also indicates the weight of the niche content. The second and third (and rank 5, 6, 7, and 8) most viewed videos were watched by only one participant. Whereas the videos on rank 1, 4, 9, and 10 have been viewed by 10–20 per cent of the study participants. Note that this observation is likely exceptional. There are no other videos that have been watched by only one person among the top 100 most watched videos in our dataset, and the six videos that have been watched by only one participant have been watched by the same participant. Note also that apart from the videos watched by only one participant, all the most frequently watched videos are children's songs. This is most likely driven by children using YouTube on a device that is logged in to a parent's account.

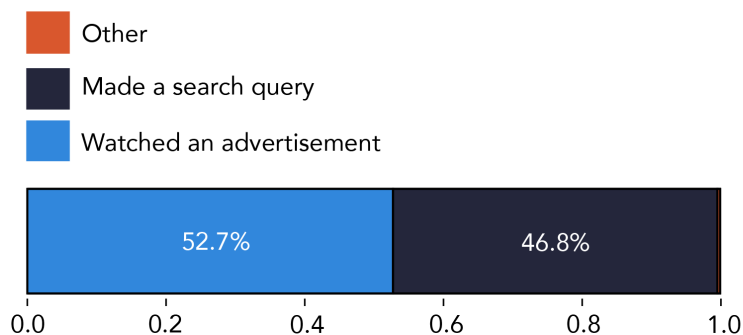
Table 13 shows the videos with the widest reach, meaning the highest number of participants playing a given video at least once, within our dataset. It shows that the top-viewed videos are solely music videos and mostly of English-speaking world-renowned artists, with only one Danish children's song in between (rank 8).

Table 13 The top 10 videos with the widest reach within our dataset

Rank	Video title	Number of participants (out of 1,063) whose watch history contains the video at least once
1	Rick Astley - Never Gonna Give You Up (Official Music Video)	325
2	Wham! - Last Christmas (Official Video)	320
3	PSY - GANGNAM STYLE(강남스타일) M/V	295
4	Mariah Carey - All I Want For Christmas Is You (Official Video)	266
5	Ed Sheeran - Shape of You (Official Music Video)	257
6	Mark Ronson - Uptown Funk (Official Video) ft. Bruno Mars	248
7	Passenger Let Her Go (Official Video)	245
8	5. Fem små aber - Sangskattekisten	245
9	Katy Perry – Roar	245
10	Queen – Bohemian Rhapsody (Official Video Remastered)	240

9.3 Search data

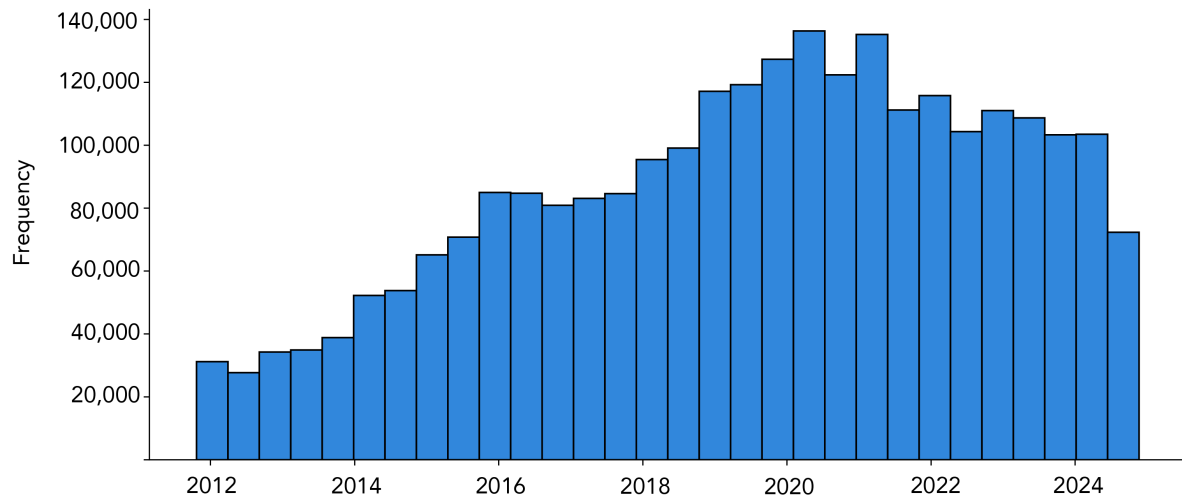
Like the watch histories, YouTube provides its users with their search history in JSON format. In aggregate, we received 5,587,646 search-history entries from 1,052 participants who donated their search histories. Closer inspection of the search data showed that it, too, contained more than just search queries that, with 2,610,550 entries, made up 46.78 per cent of the obtained search data. For instance, it is not immediately clear why the search log also contains 2,942,668 logs of watched advertisements (52.73%) and a total of 27,358 other data entries (0.39%). We have chosen to disregard these logfile instances for now. See Figure 28 for an overview.

Figure 28 Proportion of advertisements, search queries, and other entries across all donated search histories

Comments: The number of "other" entries is too small to make up a substantial part of the visualisation.

Figure 29 illustrates how the search queries are distributed across time. Interestingly, their number appears to fall since a peak around 2020.

Figure 29 Distribution of searches over time



Comments: The figure depicts only search query entries.

The most frequent searches

To understand what people are searching for, we query the dataset for the most frequent search queries, and Table 14 shows that the searches contain a mix of search terms from a Danish context (e.g., the Danish influencers Morten Münster and Fie Laursen), international celebrities (e.g., Taylor Swift), and international brands (e.g., Tesla, Minecraft).

Table 14 Top 10 most frequent searches

Rank	Search Query	Search Counts
1	pewdiepie	3,088
2	popsi+og+krelle	2,448
3	silver+wings+mike+rowland	2,017
4	taylor+swift	1,536
5	tesla	1,523
6	fie+laursen	1,383
7	morten+münster	1,308
8	minecraft	1,235
9	kristine+sloth	1,213
10	sidemen	1,151

The data draw a similar picture if we count search queries only once per participant to adjust for participants who frequently make the same searches, as seen in Table 15.

Table 15 Top 10 most common searches, with participants' searches counted only once

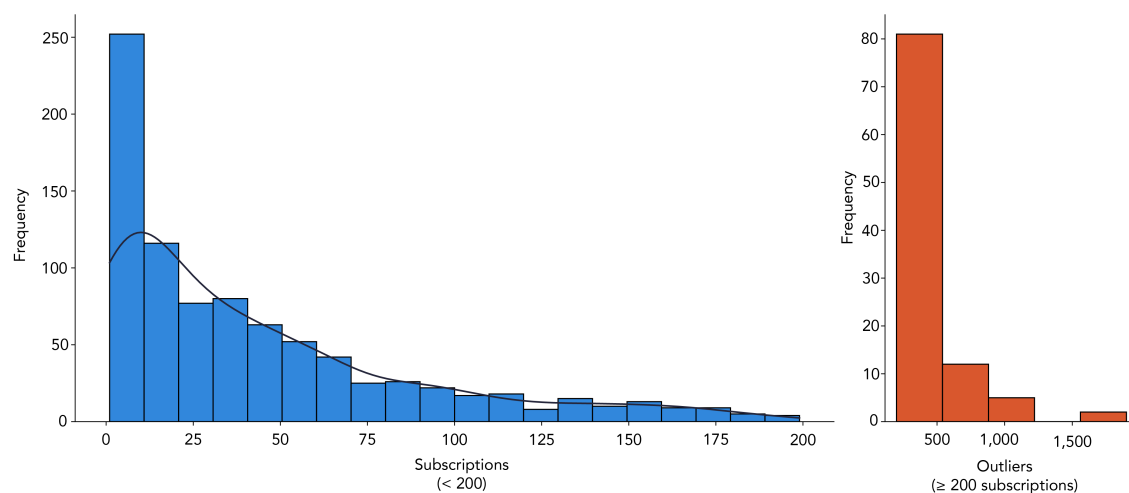
Rank	Search Query	Number of participants using that search query
1	pewdiepie	276
2	gangnam+style	275
3	taylor+swift	253
4	ed+sheeran	251
5	rasmus+seebach	242
6	justin+bieber	234
7	lukas+graham	226
8	baby+shark	214
9	mrbeast	205
10	morten+münster	204

Again, Table 15 shows that international celebrities (Taylor Swift, Ed Sheeran) and influencers (Pewdiepie) and famous videos (baby+shark) are common search terms for the most part, and only few national search terms appear (Rasmus Seebach, Lukas Graham, and Morten Münster). The search terms indicate a large interest in music on YouTube.

9.4 Subscription data

A YouTube subscription allows users to follow a channel and receive updates about new videos in their feed. It helps users keep track of creators without searching for them manually. Of the participants, 90.4 per cent shared what channels their account was subscribed to at the time of data retrieval and, thus, contributed to the subscription data set. Assuming that those participants who did not donate subscription data do not use the subscribe functionality of YouTube, we can conclude that the participants of our survey on average subscribe to 74.3 channels. However, the propensity to subscribe varies substantially between participants, as the distribution of numbers of subscriptions among our participants in Figure 30 indicates. While most participants subscribe to fewer than 200 channels, some subscribe to more than 500.

Figure 30 Distribution of number of subscriptions per user

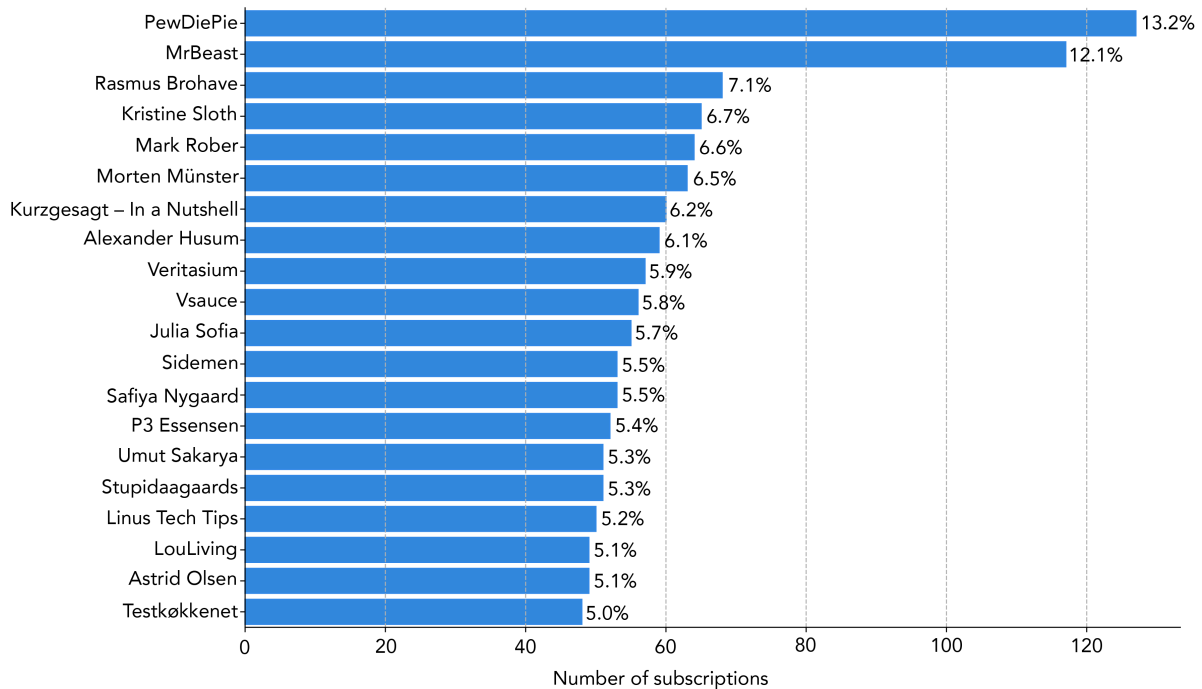


Comments: The data is displayed in two frames to reflect the main body of the data ($N = 862$ participants; left) and the outliers, with substantially more than average subscriptions ($N = 99$ participants; right).

Distribution of subscriptions on channels

Overall, participants of our survey subscribe to 46,403 different YouTube channels, indicating a substantial degree of variation between YouTube users. However, there is also evidence for some alignment, particularly regarding the most subscribed to channels. Figure 31 shows the 20 most subscribed to channels and illustrates that more than 10 per cent of our participants subscribed to the two most subscribed to YouTube channels: PewDiePie and MrBeast.

Figure 31 Top 20 channels by number of subscriptions



Comments: Notably, this data does not reflect that of the most watched videos, insofar as that the latter was dominated by music artists who are mostly absent here.

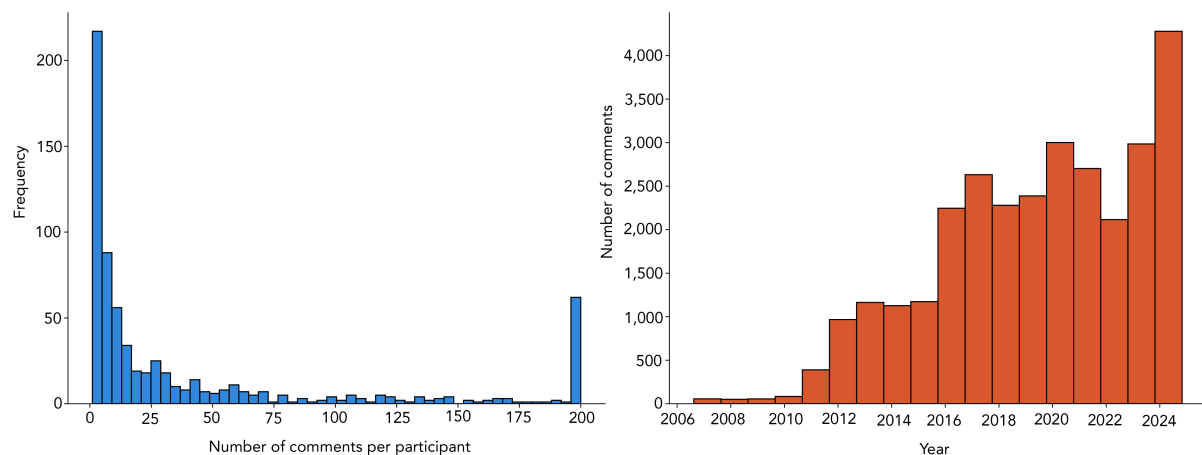
As with the search data, the subscription data we obtained indicates that YouTube is very much an international media platform. However, the share of Danish channels is more substantial here. More than half of the most subscribed to channels produce predominantly Danish content. The list of the channels that most respondents subscribe to also indicates that influencer and infotainment content plays a prominent role in terms of what content Danish YouTube users want to pay attention to. With *P3 Essensen*, one production by the Danish public service broadcaster DR is featured among the 20 most subscribed to channels.

9.5 Comment data

We asked the participants to donate their comment data to us. However, comments are potentially very personal and may easily be unique and thus allow re-identification of their author. Thus, following the principle of data minimisation and our goal to protect the identity of the study participants, Epinion removed the comment text from the data before we received it, and we focused our analysis on the less sensitive timestamps as well as the posts and videos that the participants in question have

commented on. YouTube affords the possibility to comment on videos and on so-called community posts – a social media style update that creators can share on their channel. These posts give creators the opportunity to engage their audience without posting a full video. Overall, 691 of our participants donated comment data, and we received no empty comment files. This indicates that YouTube did not generate a comment-history file for participants who never used the comment function. As discussed earlier, we cannot rule out that some participants intentionally withheld their comment data but consider it unlikely. We obtained a total of 29,707 comments. 29,182 of these were video-related comments and 525 were community-post-related comments. By default, YouTube appears to truncate a comment-history file, which it provides in the form of a “.xcls-file”, at 200 entries. This leads to a conspicuous peak in the distribution of number of comments per participant, illustrated in Figure 32. However, it is worth noting that the main share of respondents donated data that contained less than 200 comments, indicating that any limitation in the amount of data caused by a potential truncation of the comment data only pertains to 5.8 per cent of the participants in our sample.

Figure 32 Distribution of donated comment data



Comments: The left frame illustrates the number of comments donated by each participant. Note that participants who never commented are not represented here. The right frame shows the distribution of comments over time and depicts an increase in comments over time.

10. GDPR registration and ethical reflections

The “Social Media Influence” project funded by Independent Research Fund Denmark has been GDPR-registered at Aarhus University (no. 3251). The YouTube subproject that this study refers to has been ethically approved by the Internal Review Board at Aarhus University (no. 2024-038). Furthermore, Association of Internet Researchers’ research ethical guidelines 3.0 (Franzke et al., 2020) have been the primary backdrop of the ethical considerations of the data collection, storage, and publication.

In the data collection, Epinion acted as data processor while Aarhus University acted as data controller, as the research authorisation lies with the university team of researchers who decided how the study should be designed and carried out. During the data-collection process, donated YouTube data was stored on secure SharePoint servers at Epinion. The servers were logged so employees accessing servers were known at any point in time and external access required two-factor authentication, ensuring that data transfer with DATALAB researchers at Aarhus University, subsequently and during the process, could happen in a secure manner, where only named parties have access to the data. The access restriction was lifted when the project was delivered to DATALAB, and the data was deleted from Epinion’s servers when the project was finalised from Epinion’s side.

In terms of ethical considerations, we were careful to describe that participants could withdraw at any point in time and request that their data be deleted if they were not already anonymised. All data from Surveys 1–4 were anonymised. The data files did not in themselves contain personal data on the participants, as Epinion on our request carefully deleted all personal identifiable files, including comment texts, before handing over the data to the DATALAB research team at Aarhus University. If participants subscribed to their own channels, we would not be able to identify that this was in fact the channel of the person. Even though we have designed, tested, and evaluated the setup over the course of one year and know that the data is anonymised in a classic sense, we still promised the participants pseudonymisation to be sure to be able to live up to the contract installed with them. We also made sure to explicitly describe the potentially intimate character of the information that the participants disclosed to us by sharing their data in the informed consent text that participants were provided with when informed about the data donation (the second consent form discussed earlier). We also made sure to explicitly provide a careful description of places where the data may be shared, including password-required online academic repositories for reproduction purposes and aggregated, anonymised/pseudonymised data in publications, academic presentations, and broader knowledge dissemination, for example, policy work, public conferences, and media outlets.

In the analysis process, we used a unique key to join responses from Surveys 1–4 with the donated data files. Furthermore, we stored and computed all data on secure university or university cleared servers and high-performance computing facilities only accessible to members of the research team to prevent data breaches. For the analysis of the actual videos viewed by the participants, we retrieved and stored the videos at university or university-cleared servers with the purpose of datamining in accordance with copyright regulation (Danish: Ophavsretslov).

11. Discussion and conclusion

Overall, the study of YouTube usage in the Danish population as an illustrative case of data donation at national population scale showed that this method is indeed scalable. The project succeeded in retrieving valid donated data logfiles from 1,063 participants recruited using quotas to approximate a sample representative of the adult Danish YouTube user population (as estimated by a prior representative survey study) with a median total completion time of 2.3 days and with no gender, age, or education difference in completion time (excl. 5% outliers). Compared with the estimated adult Danish YouTube user population, the largest deviance is found for the oldest age group (55+), which is underrepresented in the final sample by 11 percentage points, despite oversampling the older age group in the recruitment process to meet the quota. The findings presented in this publication suggest that donating data requires technical confidence and skills that older age groups might not possess to the same extent as the younger age groups (see elaboration below), and this corresponds with findings from earlier studies (Ohme et al., 2021). The comparison between the final sample and the target population should be read with the reservation that we defined YouTube usage as biweekly usage but used the sociodemographic profile of weekly users as a benchmark, as this was the best data we had available.

The country context in which we scaled data donation potentially contributed to the success of the approach. This is difficult to assess as, to our knowledge, no similar studies have been conducted in other contexts. Future studies should therefore be aware of the specifics of the Danish case. In general, Denmark and the Nordic countries have a large penetration rate of social media, which means that a relatively large share of the population belongs to the target group. Furthermore, the higher level of social and institutional trust in Denmark and the Nordic countries compared with other countries could contribute to trust in scientists as well (Syvertsen et al., 2014), potentially increasing willingness to donate. Furthermore, the success of the present study might be influenced by the study design: Participants were requested to donate data *before* they answered additional questions in a post-survey questionnaire, ensuring that participants unwilling to donate were screened out early. One alternative would be to first recruit participants for a survey study in which an additional component at the end of the survey asks participants to donate data, with the risk that many respondents will be unwilling to donate (as applied by Welbers et al., 2024). While our study design allows for incentivising only donating participants, the latter approach would potentially require also incentivising participants who do not end up donating data. Future studies need to investigate how these differences in the design of data donation studies affect the occurrence of measurement errors according to, for example, the framework proposed by Boeschoten and colleagues (2022). Furthermore, for the Netherlands, Welbers and colleagues (2024) used the data donation method at scale, but partly carried out data collection at a music festival. Compared to this, the present study used a survey institute for recruitment solely, creating better grounds for obtaining a sample representative of the target population.

The scope of the overall project created some additional challenges, because the focus needed to be on users who had a biweekly YouTube use at minimum and who had data dating back to before 1 January 2021, due to a theoretical interest in rabbit holes and over-time cross-event analyses. For these and other reasons, 496 participants failed

the data validation. The main reasons for failing were that respondents did not watch enough videos (23%), did not have data reaching back to the required date (36%), or did not provide a correct watch history (21%); remaining participants failed due to other reasons. Thus, the extra requirements due to the scope of the project made the scaling of the data donation more difficult in terms of recruitment of participants. Furthermore, while the data validation procedure has the advantage of ensuring that only respondents who have uploaded the correct data are allowed to move on to the post-survey, it should be weighed against the ethical implications of disallowing participants who have gone through the data donation steps to complete the study and collect their reward. In the present study, requirements for the data were made clear to the participants throughout the data donation process.

Although scaling of data donation to national population level proved successful, the drop-off was large. When respondents who were eligible to participate and showed an initial interest in participating received information that participation required donation of data, close to half of the respondents declined to participate. The findings show that it is especially older participants who do not consent to participate when informed of the data donation, which could indicate that less willingness to donate data is related to insecurity regarding one's own technical skills. However, there is also a sizeable drop-off among the younger age groups following information of the data donation requirement.

Prior findings show that technical skills and concerns more strongly affect willingness to donate data than privacy concerns (Ohme et al., 2021). Future studies should shed more light on what explains the widespread reluctance to participate in donation studies reflected in our study, also including, for example, time demands, concerns over safe data storage, opacity regarding what information the data package contains, and low trust in scientists and research, as well as how such concerns can be mitigated.

In the technically demanding parts of the study, again especially older participants dropped off. This points to actual (and not just self-perceived) lack of technical skills as a central reason for drop-off. Furthermore, it must be considered that the older respondents who completed data donation studies might not be representative of their age group but reflect user patterns, attitudes, and so on among more technically skilled members of the older population. While some differences could be found between education groups, lower or higher education was not consistently associated with more or less willingness to participate or drop off at the stages requiring technological skills. This indicates that data donation does not render samples systematically skewed towards lower or higher education.

There is a higher drop-off rate in the step where respondents needed to download and share their data than in the step where they had to request their data. This suggests that it is particularly important to provide guidelines and assistance for finding the data package, downloading it in the correct format (in case the technical implementation requires specific formats), and uploading it to a safe server. However, the numbers are also influenced by the data validation process happening at this stage. Welbers and colleagues (2024), for example, have designed a built-in module for retrieving and validating data (see also Araujo et al., 2022). This might prevent this issue to a larger extent, albeit they did not receive as many donated files as our study, potentially due to other issues already discussed.

Data donation requires that data is stored. Findings from the study show that a sizeable share of respondents otherwise eligible and willing to participate were screened out because of this requirement. In other words, one of the limitations of the method is that there are parts of the population whose YouTube (and other social media) use cannot be captured using this method. This happens, for instance, when they do not have a Google account, are not signed in to their Google account (done by default though), or when their usage happens in a third-party environment (e.g., watching YouTube videos via Facebook or TikTok) where the YouTube URL does not travel with the content. Thus, data donation studies might need to be supplemented by a tracking element to collect more comprehensive data about the extent to which individuals watch YouTube. However, these methods have numerous other limitations, as they do not capture usage retrospectively or across different devices and browsers (Ohme et al., 2024).

Due to limitations in the technical setup, it was not possible to assess whether spending money on the extensive phone and e-mail hotline in this study made a difference in terms of securing completions. Future studies should shed light on this to determine whether putting resources into such a hotline is money well spent.

The aim of this publication was to provide a thorough description and discussion of the method used to scale data donation to national population level. In addition to this, initial analytical overviews of the donated data were presented. This overview supports the argument that data donation as a method only provides what is made available by platforms at any given time of the traced activities made by participants as users of a platform. For instance, the oldest comment data extends much later into the past than the oldest viewing or search data. Furthermore, by default, YouTube's internal data provision processes seem to provide only 200 comments per participant. Even though donated data gives a good glimpse – and maybe the best glimpse available at the time of writing – into what people do on platforms, researchers must be careful when drawing conclusions from logfile data about actual platform behaviour. To support the insights from the donated data, the study has collected self-reported survey data from the participants regarding their YouTube usage to increase the robustness of the interpretations. However, more future studies focusing on the validity of donated data logfiles are needed to evaluate the robustness of data donation studies across platforms, populations, and countries.

The initial analyses of the donated data paint a picture of YouTube usage in Denmark as involving predominantly the consumption of niche content. In total, the participants' watch histories contained 18,828,115 entries (15 million after excluding advertisements and other non-video entries). When excluding advertisements, other non-video entries, and duplicates, these include a total of 6,993,570 unique videos. Of the participants who successfully donated valid data, 96.3 per cent donated search data, 90.4 per cent donated subscription data, and 65.0 per cent donated comment data. This indicates that well over half of the participants at some point had engaged in dialogue around either videos or posts on videos. This is particularly interesting given that YouTube often is not included as a social media platform in national and international statistics, even though the platform provides opportunities for users to upload, like, share, and comment on videos and posts like other social media platforms (see, e.g., Danmarks Statistik, 2023; Statistics Norway, 2023).

There is a stark recency bias in the donated data. With increasing recency, the data coverage increases. Conversely, there is less data describing YouTube use in the more distant past. This is important to consider when comparing historic and current on-platform behaviour. Furthermore, it must be considered that, when analysing older data, one draws from a thinner, and thus less representative, pool of participants.

The demographics of the 1,063 participants who donated valid data show that males, young people, and those whose highest achieved education is either upper-secondary education or a bachelor's degree donate disproportionately more data, indicating that these groups also use YouTube more intensely than others. A similar picture is painted by the group of participants with upper-secondary education in the context of donated comment and subscription data. This group donated substantially more comment and subscription data than the other groups, indicating that they use these functionalities of YouTube more actively.

The analysis of videos viewed by the highest number of participants showed that the top 10 list contained only music videos, and strikingly, the viral videos in the sample had been viewed by less than a third of all participants. Again, this suggests that YouTube users have individualised rather than overlapping viewing patterns. Future studies within the project and beyond could thus take this central role of niche content as a starting point for analysis.

The estimated language distribution in our dataset shows that English dominates among the watched videos. The computational audio analysis of a random sample ($n = 20,180$) found 79.1 per cent of videos with detectable language to be English, whereas only 16.3 per cent were found to be Danish. This pattern is reflected in the participants' searches. The top search list across participants is dominated by internationally well-known musicians and influencers, with only few Danish musicians and influencers among them.

At the same time, however, the list of channels that most participants subscribe to shows that there is an interest in nationally produced content as well, as this list is dominated by a more evenly distributed mix of international and national influencers and infotainment YouTube channels. This list also suggests that influencers play a central role on YouTube in terms of what content Danish YouTube users choose to follow, whereas channels by established media organisations were almost not represented on the list (e.g., the public service broadcaster DR was represented once). More studies within and beyond this project as well as advancement of the methodologies to analyse video consumption at scale are needed to provide a more precise understanding of the videos, genres, and topics that Danes watch on YouTube and, thus, the role YouTube plays in Denmark's democratic processes.

No method can comprehensively represent actual YouTube usage and on-platform behaviour. Accordingly, this study has accounted for numerous limitations to the method of data donation regarding participant recruitment, data collection, and the content of the donated data. Highlighting these limitations, we hope, will foster the development of future research in the field and inspire the development of even more rigorous methodologies. Despite its limitations, this case study has proven that scaling

data donation to national population level is possible. This includes the incorporation of a questionnaire in the data donation process that allows the contextualisation of the donated data.

In comparison to other, less clear obligations to proprietary data access from platforms, the data portability rule in the General Data Protection Regulation (GDPR) is clear in that platforms must provide individuals requesting it with all data stored that relates to the individual. Albeit regulators should work towards platforms providing more complete files both on YouTube (e.g., comments) and other platforms, the relatively clear GDPR requirements might be the reason why it has broadened the scope and stabilised the method to a larger extent than previous and existing access points under the Digital Service Act. Thus, it is important for future method developments within academia that the standards for what should be provided in the Digital Service Act (and GDPR) becomes clearer and non-negotiable.

Additionally, in the Digital Service Act, systemic risk analysis is the main focal point of data access for research in especially Article 40. In other words, with data donation as a method, the contract is with the users rather than with the platforms, as the users decide to take out their data and share it with researchers. It does not allow the platforms nor the European Union to decide if the purpose of the research is legitimate; instead, this decision lies with the users in dialogue with an ethically cleared research team. As all users might not read informed consent contracts carefully – at least not in relation to social media (Bechmann, 2014) – this places pressure on the internal review boards. At the same time, it secures freedom of science from political interventions from both platforms and politicians. As such, the method cannot be overlooked, as it provides an opportunity for a data source of behavioural and multimodal communication data that provides more insights into the trends and challenges of the digital media landscapes, now also at national population scale.

12. References

- Abbate, J. (2000). *Inventing the internet*. The MIT Press.
- Abidin, C. (2016). Visibility labour: Engaging with Influencers' fashion brands and #OOTD advertorial campaigns on Instagram. *Media International Australia*, 161(1), 86–100. <https://doi.org/10.1177/1329878X16665177>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics*, 6(2), 2053168019848554. <https://doi.org/10.1177/2053168019848554>
- Araujo, T., Ausloos, J., Van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., Van De Velde, B., De Vreese, C., & Welbers, K. (2022). OSD2F: An open-source data donation framework. *Computational Communication Research*, 4(2), 372–387. <https://doi.org/10.5117/CCR2022.2.001.ARAU>
- Bechmann, A. (2014). Non-informed consent cultures: Privacy policies and app contracts on Facebook. *Journal of Media Business Studies*, 11(1), 21–38. <https://doi.org/10.1080/16522354.2014.11073574>
- Bechmann, A. (2020). Tackling disinformation and infodemics demands media policy changes. *Digital Journalism*, 8(6), 855–863. <https://doi.org/10.1080/21670811.2020.1773887>
- Bechmann, A., & Bowker, G. C. (2019). Unsupervised by any other name: Hidden layers of knowledge production in artificial intelligence on social media. *Big Data & Society*, 6(1), 1–11. <https://doi.org/10.1177/2053951718819569>
- Bechmann, A., & Vahlstrup, P. B. (2015). Studying Facebook and Instagram data: The Digital Footprints software. *First Monday*, 20(12), 1–13. <https://doi.org/10.5210/fm.v20i12.5968>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press. <https://doi.org/10.1093/oso/9780190923624.001.0001>
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bennett, W. L., Segerberg, A., & Walker, S. (2014). Organization in the crowd: Peer production in large-scale networked protests. *Information, Communication & Society*, 17(2), 232–260. <https://doi.org/10.1080/1369118X.2013.870379>
- Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, 4(2), 388–423. <https://doi.org/10.5117/CCR2022.2.002.BOES>

- Bowker, G., Baker, K., Millerand, F., & Ribes, D. (2010). Toward information infrastructure studies: Ways of knowing in a networked environment. In J. Hunsinger, L. Klastrup, & M. Allen (Eds.), *International handbook of internet research* (pp. 97–117). Springer. https://doi.org/10.1007/978-1-4020-9789-8_5
- Bradshaw, S., & Howard, P. N. (2018). *Challenging truth and trust: A global inventory of organized social media manipulation*. Oxford Internet Institute. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf>
- Brennen, S., Simon, F. M., Howard, P. N., & Nielsen, R. K. (2020). *Types, sources, and claims of COVID-19 misinformation*. Reuters Institute for the Study of Journalism, University of Oxford. <https://doi.org/10.60625/risj-awvq-sr55>
- Breuer, J., Kmetty, Z., Haim, M., & Stier, S. (2023). User-centric approaches for collecting Facebook data in the ‘post-API age’: Experiences from two studies and recommendations for future research. *Information, Communication & Society*, 26(14), 2649–2668. <https://doi.org/10.1080/1369118X.2022.2097015>
- Bruns, A., Bechmann, A., Burgess, J., Chadwick, A., Clark, L. S., Dutton, W. H., Ess, C. M., Gruz, A., Halford, S., Hermida, A., Hofmann, J., Howard, P., Katzenbach, C., Liang, H., Lewis, S. C., Peng, W., Puschmann, C., Qiu, J., Quinn, K., ... Zimmer, M. (2018). Facebook shuts the gate after the horse has bolted, and hurts real research in the process. *Internet Policy Review*. <https://policyreview.info/articles/news/facebook-shuts-gate-after-horse-has-bolted-and-hurts-real-research-process/786>
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(7), 1164–1180. <https://doi.org/10.1177/1461444812440159>
- Buning, M., et al. (2018). *A multi-dimensional approach to disinformation: Report of the independent high level group on fake news and online disinformation* (pp. 1–40). Publications Office of the European Union – EU Commission. <https://data.europa.eu/doi/10.2759/739290>
- Charquero-Ballester, M., Walter, J. G., Nissen, I. A., & Bechmann, A. (2021). Different types of COVID-19 misinformation have different emotional valence on Twitter. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211041279>
- Danmarks Statistik. (2023, August 17). *Danmark bruger sociale medier mest i EU* [Denmark uses social media the most in the EU]. Retrieved March 11, 2025, from <https://www.dst.dk/da/Statistik/nyheder-analyser-publ/nyt/NytHtml?cid=46771>
- de Vreese, C., & Tromble, R. (2023). The data abyss: How lack of data access leaves research and society in the dark. *Political Communication*, 40(3), 356–360. <https://doi.org/10.1080/10584609.2023.2207488>
- Edwards, P., Bowker, G., Jackson, S., & Williams, R. (2009). Introduction: An agenda for infrastructure studies. *Journal of the Association for Information Systems*, 10(5). <https://doi.org/10.17705/1jais.00200>
- EFCSN (European Fact-Checking Standards Network). 2025, January 22). *EFCSN statement on platforms’ reduced commitments to the code of practice on*

- disinformation. https://efcsn.com/news/2025-01-22_efcsn-statement-on-platforms-reduced-commitments-to-the-code-of-practice-on-disinformation/
- Eurostat. (2024, December 17). *Individuals using the internet for participating in social networks*. <https://doi.org/10.2908/tin00127>
- Franzke, A., Bechmann, A., Ess, C. M., & Zimmer, M. (2020). *Internet research: Ethical guidelines 3.0*. AoIR. The International Association of Internet Researchers. <https://aoir.org/reports/ethics3.pdf>
- Freelon, D. (2018). Computational research in the post-API age. *Political Communication*, 35(4), 665–668. <https://doi.org/10.1080/10584609.2018.1477506>
- Freelon, D., McIlwain, C., & Clark, M. (2018). Quantifying the power and consequences of social media protest. *New Media & Society*, 20(3), 990–1011. <https://doi.org/10.1177/1461444816676646>
- Gillespie, T. (2018). *Custodians of the internet*. Yale University Press. <https://yalebooks.yale.edu/book/9780300173130/custodians-internet>
- González-Bailón, S. (2017). *Decoding the social world*. The MIT Press. <https://mitpress.mit.edu/books/decoding-social-world>
- Gruzd, A., & Wellman, B. (2014). Networked influence in social media: Introduction to the special issue. *American Behavioral Scientist*, 58(10), 1251–1259. <https://doi.org/10.1177/0002764214527087>
- Hatmaker, T. (2021, August 4). *Facebook cuts off NYU researcher access, prompting rebuke from lawmakers*. TechCrunch. <https://techcrunch.com/2021/08/04/facebook-ad-observatory-nyu-researchers/>
- Horowitz, M., Walter, J., & Bechmann, A. (2024). *Comment on the draft delegated regulation on data access provided for in the digital services act*. EU Commission. https://pure.au.dk/ws/portalfiles/portal/416790933/DSA_Art_40_Datalab_Deca_Nordis_54_.pdf
- Howard, P. N. (2020). *Lie machines: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives*. Yale University Press.
- Hussain, M. N., Tokdemir, S., Agarwal, N., & Al-Khateeb, S. (2018). Analyzing disinformation and crowd manipulation tactics on YouTube. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 1092–1095. <https://doi.org/10.1109/ASONAM.2018.8508766>
- Kaplan, J. (2025, January 7). *More speech and fewer mistakes* [Video]. Meta. <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>
- Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Free Press.
- Kulturministeriet. (2021). *Mediernes Udvikling i Danmark: Internetbrug og sociale medier* [Media development in Denmark: Internet use and social media]. https://kum.dk/fileadmin/_mediernesudvikling/2021/Internetbrug_og_sociale_medier_2021.pdf

- Lazer, D., Hargittai, E., Freelon, D., Gonzalez-Bailon, S., Munger, K., Ognyanova, K., & Radford, J. (2021). Meaningful measures of human society in the twenty-first century. *Nature*, 595(7866), 189–196.
<https://doi.org/10.1038/s41586-021-03660-7>
- Marwick, A. E. (2013). *Status update: Celebrity, publicity, and branding in the social media age*. Yale University Press.
- McQuail, D., & Deuze, M. (2022). *McQuail's media and mass communication theory*. Sage.
- Mitchell, A., Gottfried, J., Stocking, G., Walker, M., & Fedeli, S. (2019, June 5). *Many Americans say made-up news is a critical problem that needs to be fixed*. Pew Research Center's Journalism Project.
<https://www.journalism.org/2019/06/05/many-americans-say-made-up-news-is-a-critical-problem-that-needs-to-be-fixed/>
- Moeller, L. A., & Bechmann, A. (2019). *Research data exchange solution* [Report for the EU Commission]. SOMA.
- Newman, N., Fletcher, R., Robertson, C. T., Ross Arguedas, A., & Nielsen, R. K. (2024). *Reuters Institute digital news report 2024*. Reuters Institute for the Study of Journalism, University of Oxford.
<https://doi.org/10.60625/RISJ-VY6N-4V57>
- NordMedia Network. (2024, May 16). *Media trends in the Nordics: Key insights from the 2023 Swedish and Norwegian media barometers*.
<https://nordmedianetwork.org/latest/news/media-trends-in-the-nordics-key-insights-from-the-2023-swedish-and-norwegian-media-barometers/>
- Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2024). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Communication Methods and Measures*, 18(2), 124–141.
<https://doi.org/10.1080/19312458.2023.2181319>
- Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS screen time function. *Mobile Media & Communication*, 9(2), 293–313.
<https://doi.org/10.1177/2050157920959106>
- Pfiffner, N. (2024). *Data donation as an approach to studying digital media use – Exploring opportunities, challenges, and applications* [Doctoral dissertation, University of Zurich, Switzerland]. <https://doi.org/10.5167/uzh-262837>
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. <https://doi.org/10.1177/1461444816661553>
- Quandt, T. (2018). Dark participation. *Media and Communication*, 6(4), Article 4.
<https://doi.org/10.17645/mac.v6i4.1519>
- Quandt, T. (2021). Can we hide in shadows when the times are dark? *Media and Communication*, 9(1), Article 1. <https://doi.org/10.17645/mac.v9i1.4020>

- Razi, A., Alsoubai, A., Kim, S., Naher, N., Ali, S., Stringhini, G., De Choudhury, M., & Wisniewski, P. J. (2022). Instagram data donation: A case study on collecting ecologically valid social media data for the purpose of adolescent online risk detection. *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–9.
<https://doi.org/10.1145/3491101.3503569>
- Rieder, B. (2013, May 2). Studying Facebook via data extraction: The Netvizz application. *WebSci '13: Proceedings of the 5th Annual ACM Web Science Conference*, 346–355. <https://doi.org/10.1145/2464464.2464475>
- Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to ‘ranking cultures’: Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68.
<https://doi.org/10.1177/1354856517736982>
- Rogers, R. (2018). Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication*, 12, 450–472.
<https://ijoc.org/index.php/ijoc/article/view/6407>
- Statista. (2024, October 17). *Social media usage in the Nordics: Statistics & facts*. <https://www.statista.com/topics/11607/social-media-usage-in-the-nordics/#topicOverview>
- Statistics Norway. (2023). *Norsk mediebarometer [Norwegian media barometer]*. <https://tinyurl.com/4bk3a6hv>
- Syvvertsen, T., Enli, G., Mjøs, O. J., & Moe, H. (2014). *The media welfare state: Nordic media in the digital era*. University of Michigan Press.
<https://doi.org/10.2307/j.ctv65swsg>
- The Brussels Times. (2023, May 27). *Twitter leaves EU voluntary Code of Practice against disinformation*. <https://www.brusselstimes.com/524844/twitter-leaves-eu-voluntary-code-of-practice-against-disinformation>
- Thomas, L. (2021, June 4). *Supreme Court ruling that limits hacking law supports U-M researcher*. Michigan News, University of Michigan.
<https://news.umich.edu/supreme-court-ruling-that-limits-hacking-law-supports-u-m-researcher/>
- Turner, F. (2008). *From counterculture to cyberculture: Stewart Brand, the whole earth network, and the rise of digital utopianism*. University of Chicago Press.
<https://press.uchicago.edu/ucp/books/book/chicago/F/bo3773600.html>
- Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, 63(2), 221–243.
<https://doi.org/10.1111/jcom.12024>
- van Driel, I. I., Giachanou, A., Pouwels, J. L., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of social media data donations. *Communication Methods and Measures*, 16(4), 266–282.
<https://doi.org/10.1080/19312458.2022.2109608>

- Walter, J., & Bechmann, A. (2023). *Delegated regulation on data access provided for the digital services act: Response to call for evidence from Nordic observatory for digital media and information disorder (NORDIS)*. SSOAR.
<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-86773-5>
- Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe.
<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Warofka, A. (2018, November 5). *An independent assessment of the human rights impact of Facebook in Myanmar*. Meta.
<https://about.fb.com/news/2018/11/myanmar-hria/>
- Wedel, L., Ohme, J., & Araujo, T. (2024). Augmenting data download packages – Integrating data donations, video metadata, and the multimodal nature of audio-visual content. *Methods, Data, Analyses*, 1–32.
<https://doi.org/10.12758/mda.2024.08>
- Wei, M., Stamos, M., Veys, S., Reitinger, N., Goodman, J., Herman, M., Filipczuk, D., Weinshel, B., Mazurek, M. L., & Ur, B. (2020). What Twitter knows: Characterizing ad targeting practices, user perceptions, and ad explanations through users’ own twitter data. *Proceedings of the 29th USENIX Conference on Security Symposium*, 145–162.
- Welbers, K., Loecherbach, F., Lin, Z., & Trilling, D. (2024). Anything you would like to share: Evaluating a data donation application in a survey and field study. *Computational Communication Research*, 6(2), 1–28.
<https://doi.org/10.5117/CCR2024.2.5.WELB>
- Zaleznik, D. (2021). Facebook and genocide: How Facebook contributed to genocide in Myanmar and why it will not be held accountable. *Systemic Justice Journal*, 1. <https://systemicjustice.org/article/facebook-and-genocide-how-facebook-contributed-to-genocide-in-myanmar-and-why-it-will-not-be-held-accountable/>

13. Appendix: Screenshots of Surveys 1–4

The Appendix contains documentation of the survey implementation of the different steps accounted for in this publication. All questions were implemented with mandatory answering. Questions were not labelled in the implementation. E-mails and phone numbers are redacted and not operating anymore. The screenshots are taken from the instructions provided for participants wanting to use Firefox or Google Chrome for their data donation.

Table of contents

13.1 Survey 1	81
13.1.1 Invitation screen and first consent	81
13.1.2 Instruction for browser usage	82
13.1.3 YouTube use: Frequency	82
13.1.4 YouTube use: Categorised starting year	83
13.1.5 Birth date	84
13.1.6 Gender	84
13.1.7 Screen-out screen	85
13.1.8 Highest educational level	86
13.1.9 Trust in institutions	87
13.1.10 Information about data donation and consent to data donation	88
13.1.11 Google account	89
13.1.12 Archiving YouTube data	90
13.1.13 Contact information	92
13.1.14 Instruction screen for Survey 2	92
13.2 Survey 2	93
13.2.1 Welcome screen to Survey 2	93
13.2.2 Instructions for requesting data	94
13.2.3 Information about support	95
13.2.4 Information if support needed and time selected	95
13.2.5 Validation of data request	96
13.2.6 Wrong data file download reported	97
13.2.7 Confirming correct data request	97
13.2.8 Reported wrong file format request	98
13.2.9 After confirming requesting correct data format: Request to check e-mail for data download options and invitation to Survey 3	98

13.3 Survey 3	99
13.3.1 Welcome screen to Survey 3	99
13.3.2 Confirming reception of e-mail from Google takeout	99
13.3.3 Timing of data request	100
13.3.4 Information on checking e-mails	100
13.3.5 Information on requesting data again	101
13.3.6 Respondent's device	101
13.3.7 Instructions data download	102
13.3.8 Instruction renaming downloaded file	103
13.3.9 Information reporting no successful renaming	104
13.3.10 Reporting successful renaming	105
13.3.11 Reporting unsuccessful upload	106
13.3.12 Reporting successful upload: Information about next steps	106
13.3.13 Upload link screenshot	107
13.3.14 Screen if upload error registered at Epinion	107
13.4 Survey 4	108
13.4.1 Welcome screen to Survey 4	108
13.4.2 Social media use	109
13.4.3 Number of followers, friends, and connections on social media platforms	110
13.4.4 Motivation for YouTube usage	111
13.4.5 Media usage	112
13.4.6 Following influencer?	113
13.4.7 Interest in politics	113
13.4.8 Voting intention	114
13.4.9 Personality traits	115
13.4.10 Political orientation	116
13.4.11 Political efficacy	117
13.4.12 Political participation	118
13.4.13 Interest in news	119
13.4.14 News sources	120
13.4.15 Naming of influencer	121
13.4.16 Social media platforms for following named influencer	122
13.4.17 Parasocial relationship	123
13.4.18 Beginning of following influencer	124
13.4.19 Frequency of content consumption of first influencer mentioned	124

13.4.20 Frequency and type of content consumed on YouTube	125
13.4.21 Device use by other users than participant?.....	126
13.4.22 YouTube usage on which platform?.....	126
13.4.23 YouTube use based on self-selection or recommendation?.....	127
13.4.24 YouTube viewing sessions prolonged?.....	127
13.4.25 Effects of YouTube use	128
13.4.26 Exposure to misinformation or false news	129
13.4.27 Type of misinformation exposed to	129
13.4.28 Confidence in identifying misinformation.....	130
13.4.29 Commenting behaviour on YouTube.....	131
13.4.30 Conspiracy mentality.....	132
13.4.31 Mental and physical health.....	133
13.4.32 Trust in people	134
13.4.33 Frequency political discussions.....	134
13.4.34 Political topics discussed	135
13.4.35 Experiment Control group 1.....	135
13.4.36 Experiment Control group 2.....	136
13.4.37 Experiment Treatment group 1	136
13.4.38 Support of presented political claim	137
13.4.39 Experiment Treatment group 2.....	137
13.3.40 Agreement with political issue	138
13.3.41 Political participation issue specific	139
13.4.42 Perception of presented scenario.....	140
13.4.43 Household composition	140
13.4.44 Social class indicator	141
13.4.45 Migration background	141
13.4.46 Subjective location	142
13.4.47 Final screen and debriefing of the treatment groups	143
13.4.48 Final screen	143

13.1 Survey 1

13.1.1. Invitation screen and first consent

Epinion

– making sense



Epinion gennemfører i øjeblikket en undersøgelse for Aarhus Universitet. Formålet med undersøgelsen er at få viden om danskernes brug af YouTube og holdninger til forskellige samfundsspørgsmål.

Denne første del af undersøgelsen har til formål at afklare, om du tilhører målgruppen for undersøgelsen og samtidig ønsker at deltage.

Epinion vil dele dine besvarelser med projektgruppen på Aarhus Universitet i anonymiseret form. Dine anonymiserede besvarelser kan indgå i publikationer i videnskabelige tidsskrifter og i videnskabelige databaser, kan deles med projektets samarbejdspartnere ved andre universiteter, og kan indgå i øvrig forskningsformidling i et overordnet format, som ikke afslører din identitet.

Epinion sletter dit navn og lignende kontaktoplysninger senest 3 måneder efter arbejdet med resultaterne er afsluttet. Aarhus Universitet vil gemme og analysere dine data så længe, det er påkrævet af forskningens formål, og i overensstemmelse med ansvarlig forskningspraksis. I undersøgelsen er Aarhus Universitet dataansvarlige og Epinion er databehandler. Hvis du vil vide mere om dine rettigheder og om, hvordan Aarhus Universitet og Epinion behandler dine data, kan du læse mere om deres privatlivspolitik her:

Aarhus Universitets privatlivspolitik:

<https://www.au.dk/om/profil/privatlivspolitik>

0%

Copyright © Epinion

Epinions privatlivspolitik:

<https://epinionglobal.com/da/privatlivspolitik-for-epinion>

Du kan til enhver tid henvende dig for at få slettet, begrænset eller berigtiget dine oplysninger.

Du kan også kontakte Epinions databeskyttelsesrådgiver på [redacted] eller på telefon [redacted]

Hvis du har spørgsmål til undersøgelsen, kan du kontakte den forskningsansvarlige, professor Anja Bechmann, ved at sende en e-mail til [redacted] med "YouTube studie 2024" i emnefeltet.

Din deltagelse i undersøgelsen er frivillig. Du kan til enhver tid trække dit samtykke tilbage, hvorefter vi sletter dine oplysninger, hvis de ikke allerede er fuldt anonymiseret.

Ved at gå videre bekræfter du, at du giver samtykke til ovenstående.

Ja tak, jeg ønsker at deltage

0%

Copyright © Epinion

13.1.2 Instruction for browser usage



Tak for dit tilsagn om at deltage i undersøgelsen.

Vi anbefaler, at du anvender en af følgende browsere til at besvare undersøgelsen.



Chrome



Safari



Edge



Firefox

Tilbage

Næste

6%

Copyright © Epinion

13.1.3 YouTube use: Frequency



Hvor ofte bruger du YouTube på en typisk måned? Vælg venligst det svar, der passer bedst på dig.

- ☐ Hver dag
- ☐ Hver anden dag
- ☐ To gange om ugen
- ☐ En gang om ugen
- ☐ En gang hver anden uge
- ☐ En gang om måneden
- ☐ Jeg bruger YouTube mindre end én gang om måneden

Tilbage

Næste

13%

Copyright © Epinion

13.1.4 YouTube use: Categorised starting year

Epinion
- making sense

?

Angiv venligst det årstal, hvor du **begyndte at bruge YouTube**. Hvis du er i tvivl, så giv dit bedste bud.

- ☐ 2005 til 2009
- ☐ 2010 til 2014
- ☐ 2015 til 2019
- ☐ 2020 eller senere

Tilbage

Næste

20%

Copyright © Epinion

13.1.5 Birth date

(Year, Month, and Day as drop-down selection)

Epinion
- making sense



Hvad er din fødselsdag?

2025-01-01

Tilbage

Næste

26%

Copyright © Epinion

13.1.6 Gender

Epinion
- making sense



Er du?

- ☐ Mand
☐ Kvinde
☐ Andet

Tilbage

Næste

33%

Copyright © Epinion

13.1.7 Screen-out screen

(If respondents do not meet inclusion criteria)

Epinion

– making sense

Du passer desværre ikke ind i målgruppen for undersøgelsen. Vi beklager ulejligheden.

[These lines are only displayed in TEST mode]

The survey is redirecting to:

<https://survey-epinion.dk/>

100%

Copyright © Epinion

13.1.8 Highest educational level

(if respondents match screening criteria (instead of 13.1.7))

Epinion
- making sense



Hvad er din højeste gennemførte uddannelse?

- ☐ Grundskole (folkeskole, privatskole, efterskole)
- ☐ Gymnasial uddannelse (HTX, HHX, STX, HF)
- ☐ Erhvervsfaglig (f.eks. tømrer, frisør, kontorassistent)
- ☐ Kort videregående (f.eks. datamatiker, laborant)
- ☐ Mellemlang videregående/professionsbachelor (f.eks. teknikumingeniør, lærer)
- ☐ Bachelor (f.eks. første del af en lang videregående uddannelse)
- ☐ Lang videregående (f.eks. gymnasielærer, økonom, jurist)
- ☐ Ph.d. eller tilsvarende

Tilbage

Næste

40%

Copyright © Epinion

13.1.9 Trust in institutions (Randomisation of items)

Epinion
— making sense

?

Hvor stor eller lille tillid har du til...

	Meget stor tillid	Ret stor tillid	Ret lille tillid	Meget lille tillid	Ved ikke
... det danske retssystem	☐	☐	☐	☐	☐
... danske politikere i almindelighed	☐	☐	☐	☐	☐
... sociale medier generelt	☐	☐	☐	☐	☐
... sundhedspersonale i Danmark	☐	☐	☐	☐	☐
... politiet i Danmark	☐	☐	☐	☐	☐
... den danske regering	☐	☐	☐	☐	☐
... danske nyhedsmedier generelt	☐	☐	☐	☐	☐
... YouTube	☐	☐	☐	☐	☐
... de offentlige myndigheder i Danmark	☐	☐	☐	☐	☐
... forskning	☐	☐	☐	☐	☐
... Folketinget	☐	☐	☐	☐	☐

Tilbage

Næste

46%

Copyright © Epinion

13.1.10 Information about data donation and consent to data donation (If consent is not given here, respondents are screened out; see 13.1.7)



Din videre deltagelse i resten af undersøgelsen indebærer, at du deler dine data fra YouTube med Epinion. Formålet med det er at opnå detaljeret indsigt i danskernes YouTube-brug. Efterfølgende vil du blive bedt om at besvare et spørgeskema på cirka 15-20 minutter. Som tak for din deltagelse i den fulde undersøgelse vil du modtage et gavekort fra Epinion på **150 kr.** til sendentanke.dk.

Igenennem undersøgelsens 4 dele vil vi guide dig til, hvordan du bestiller og deler dine YouTube-data sikkert med Epinion. **Hvis du har spørgsmål undervejs, er der en hotline, du kan ringe til og få hjælp.**

Epinion anvender sendentanke.dk, til at udsende gavekort på vores vegne. Til brug for udsendelsen – og for at [Sendentanke.dk](https://sendentanke.dk) kan yde kundeservice, når du benytter gavekortet – deler vi oplysninger om dit navn og e-mail til [Sendentanke.dk](https://sendentanke.dk). Dine oplysninger bliver udelukkende anvendt til dette formål, og altså ikke solgt til tredjepart eller brugt til nogen form for markedsføring.

Hvis du deltager i resten af undersøgelsen, vil Epinion dele dine YouTube-data og dine besvarelser med projektgruppen på Aarhus Universitet i pseudonymiseret form. Dine YouTube-data og dine besvarelser kan indgå i publikationer i videnskabelige tidsskrifter og i videnskabelige databaser, kan deles med projektets samarbejdspartnere ved andre universiteter, og kan indgå i øvrig forskningsformidling i et overordnet format, **som ikke afslører din identitet.**

Dine YouTube-data kan indeholde detaljer om dig, dine interesser og dine forbrugsmønstre. For at forhindre utilsigtet brug, pseudonomiseres dine data så snart de er sammenkoblet med dine svar på spørgeskemaet, og de vil blive opbevaret på en sikker server, hvor kun autoriseret personale har adgang. Undersøgelsen, dataindsamlingen, og behandlingen af data **er godkendt af Forskningsetisk Komité på Aarhus Universitet.**

I øvrigt gælder samme vilkår, som du tidligere [har givet samtykke til \(klik her for at se informationen\)](#):

Tilsagn om at deltage i resten af undersøgelsen:

- ☐ Ja, jeg vil gerne deltage i resten af undersøgelsen, hvilket indebærer, at jeg deler mine YouTube-data og besvarer spørgeskemaet
- ☐ Nej, jeg vil ikke deltage i resten af undersøgelsen

Tilbage

Næste

85/96

Copyright © Epinion

13.1.11 Google account

(All that give consent in 13.1.10; if respondents select “no” or “don't know” here, they are screened out; see 13.1.7)

Epinion

– making sense

?

Har du en personlig Google-konto?

En Google-konto er en konto til Googles tjenester som f.eks. Gmail, Google Docs eller YouTube. Hvis du bruger disse tjenester, har du sandsynligvis en Google-konto. Du skal have været og stadig være logget ind på din Google-konto, når du bruger YouTube. Ellers gemmes dine data ikke. Det er nødvendigt, at du har været logget ind på din Google-konto og at den har din visningshistorik tilbage til 2019, for at du kan deltage i undersøgelsen og modtage dit gavekort.

☐ Ja

☐ Nej

☐ Ved ikke

Tilbage

Næste

60%

Copyright © Epinion

13.1.12 Archiving YouTube data

(If respondents select here not to archive YouTube data they are screened out; see 13.1.7)

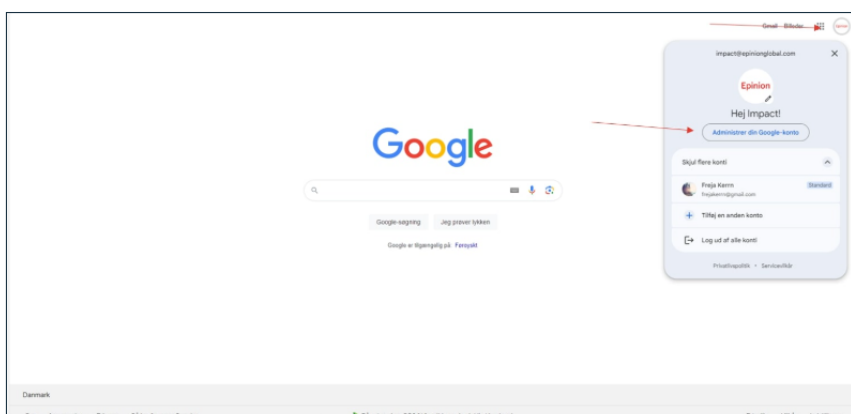
Epinion
- making sense



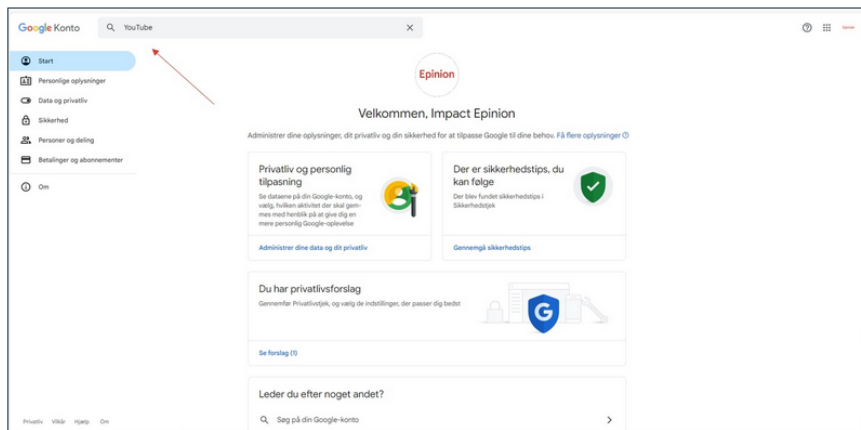
Gemmes dine YouTube-data på din Google-konto?

Det er nødvendigt, at dine YouTube-data gemmes på din Google-konto, for at kunne deltage i undersøgelsen. For at tjekke dette kan du:

- 1) Åben Google i din browser
- 2) Tryk på cirklen oppe i højre hjørne
- 3) Vælg 'Administrer din Google-konto'



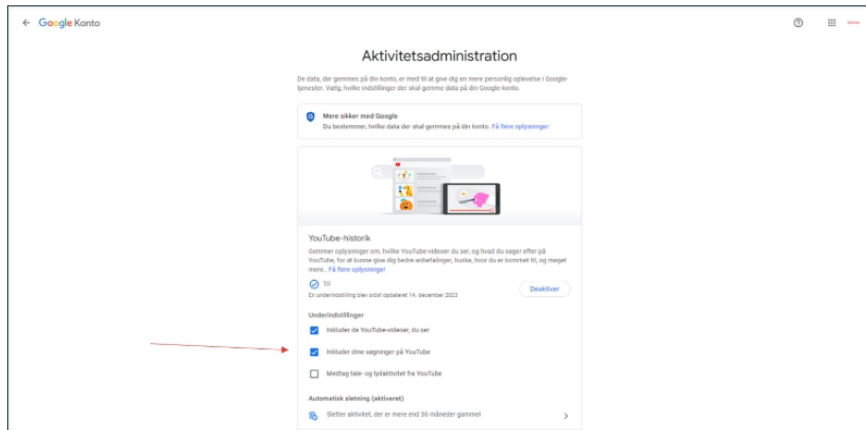
4) Skriv YouTube i søgefeltet



5) Klik på 'YouTube-historik'

6) Tjek om der er flueben i

- Inkluder de YouTube-videoer, du ser
- Inkluder dine søgninger på YouTube



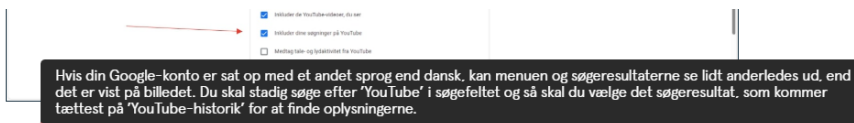
Viser din egen skærm ikke det, der er vist på billedet? [Klik her.](#)

Gemmer din Google-konto dine **YouTube-data for dine søgninger og de videoer, du ser?**

- ☐ Ja, jeg tillader at YouTube gemmer min søgningshistorik og min videohistorik
- ☐ Nej
- ☐ Ved ikke

Tilbage
Næste

60%
 Copyright © Epinion



Viser din egen skærm ikke det, der er vist på billedet? [Klik her.](#)

13.1.13 Contact information



For at vi kan invitere dig til undersøgelsen, skal vi bede om følgende oplysninger:

Vi bruger oplysningerne til at invitere dig til de næste dele af undersøgelsen og til at kunne kontakte dig i forbindelse med undersøgelsen, hvis dette er relevant.

Notér dit fornavn:	
Notér dit efternavn:	
Notér dit mobilnummer:	
Notér din e-mail:	
Bekræft din e-mail:	
Notér den e-mail, der er tilknyttet din google-konto:	

Tilbage

Næste

75%

Copyright © Epinion

13.1.14 Instruction screen for Survey 2



Tak for din besvarelse af undersøgelsens første del. Du er nu klar til at gå videre til næste del af undersøgelsen, som du finder her:

[Link til næste del af undersøgelsen.](#)

Trin 1
Tilmelding

Trin 2
Bestil Youtube-data

Trin 3
Hent og send data

Trin 4
Spørgeskema

Vi har også sendt linket til næste del af undersøgelsen på din e-mail [redacted] så du kan tilgå den, når du er klar.

Linket indeholder en guide, der hjælper dig med at bestille dine YouTube-data. Dette er undersøgelsens anden del. Hvis du oplever udfordringer med at bestille data, er der en telefonisk hotline, du kan ringe til.

Når dine data er bestilt, vil du modtage en e-mail med et link til undersøgelsens tredje del, som er en videoguide, der viser, hvordan du sikkert **uploader dine YouTube-data** hos Epinion. Når dine data er uploadet, modtager du et link til **et spørgeskema, som er fjerde og sidste del**.

Når spørgeskemaet er besvaret, modtager du en e-mail med et gavekort på **150 kr.** til sendtanke.dk.

100%

Copyright © Epinion

13.2 Survey 2

13.2.1 Welcome screen to Survey 2

Epinion
- making sense



Velkommen til den anden del af Aarhus Universitets undersøgelse af danskernes brug af YouTube.

I denne del guider vi dig til at bestille dine YouTube-data fra Google via Googles Takeout-løsning. Det tager kun 3-5 minutter at bestille disse data, og hvis du har brug for hjælp undervejs, er der en hotline, du kan ringe til.

Næste

0%

Copyright © Epinion

13.2.2 Instructions for requesting data

Epinion
- making sense



For at bestille dine YouTube-data på Google Takeout skal du gå ind på <https://google.com/takeout> og logge ind med din Google-konto, hvis du ikke allerede er logget ind. Når du er logget ind, kan du følgende nedenstående tre skridt.

Er flere forskellige Google-konti logget ind på dit device? [Klik her](#).

1. Først skal du tilvælge data fra YouTube og YouTube Music

Klik først på knappen 'Fravælg alle' og tjek, at der står, at 0 er valgt.

Rul herefter ned til bunden af sektionen og sæt et flueben i boksen ud for 'YouTube og YouTube Music'.

Hvis du er i tvivl, så følg videoguiden

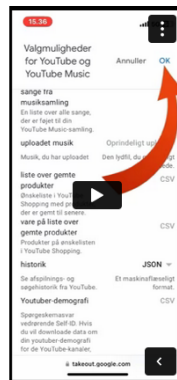


2. Så skal du vælge det rigtige filformat

Klik nu på knappen 'Flere formater' og rul ned til 'historik'. Klik på rulle-menuen til højre og vælg 'JSON'. Tryk derefter på 'OK'.

Klik så på knappen ved siden af hvor der står 'Alle YouTube-data medtages'. På denne skærm skal du først fravælge alle bokse. Dernæst skal du sætte flueben ud for 'abonnementer', 'historik' og, hvis du har mulighed for det, 'kommentarer'. Når du har gjort det, skal du trykke på 'OK'. Tryk på knappen 'Næste trin'.

Hvis du er i tvivl, så følg videoguiden

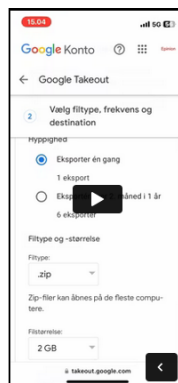


3. Vælg filstørrelse

Til sidst ruller du ned til 'Filstørrelse', og klikker på rulle-menuen nedenunder. Vælg venligst den størst mulige filstørrelse f.eks. 50 GB.

Tryk da på knappen, 'Opret eksport'. Nu har du bestilt dine data.

Hvis du er i tvivl, så følg videoguiden



Har du bestilt dine YouTube-data?

- ☐ Ja
- ☐ Nej, jeg har brug for hjælp til at bestille data

Tilbage

Næste

Copyright © Epinion

13.2.3 Information about support

(Following selecting need for help in 13.2.2)

Epinion
- making sense



Du har skrevet, at du har brug for hjælp til at bestille data. Du kan ringe til [redacted] alle hverdage mellem 10 og 12 eller skrive en e-mail til [redacted]. Så vender vi tilbage hurtigst muligt.

Ellers kan du markere her, hvornår på dagen vi må ringe til dig og hjælpe dig med din databestilling.

☐ 09:00-12:00

☐ 12:00-15:00

☐ 15:00-18:00

☐ 18:00-20:00

Tilbage

Næste

25%

Copyright © Epinion

13.2.4 Information if support is needed and time selected

(Time selected in 13.2.3)

Epinion

- making sense

Tak. En af vores medarbejdere vil snarest muligt forsøge at kontakte dig i det valgte tidsrum.

100%

Copyright © Epinion

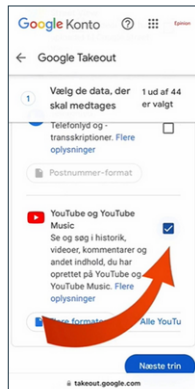
13.2.5 Validation of data request

(fter selecting yes to having ordered data in 13.2.2)

Epinion
- making sense



Vi stiller dig nu en række spørgsmål for at sikre, at dine data er bestilt korrekt



Tilvalgte du kun data fra YouTube og YouTube Music?

- ☐ Ja
- ☐ Nej, jeg bestilte ikke data fra YouTube og YouTube Music
- ☐ Nej, jeg bestilte data fra andre steder end YouTube og YouTube Music

Tilbage

Næste

25%

Copyright © Epinion

13.2.6 Wrong data file download reported

(Selecting either of the no options in 13.2.5; if one then selects here need help, it will go to 13.2.3 and 13.2.4)

Epinion

— making sense



Da vi kun skal bruge data fra YouTube og Youtube Music, skal vi bede dig bestille dine data igen. Videoen nedenfor viser, hvordan du bestiller dine data hos Google Takeout



Har du bestilt dine data?

- ☐ Ja
- ☐ Nej, jeg har brug for hjælp til at bestille data

Tilbage

Næste

13.2.7 Confirming correct data request

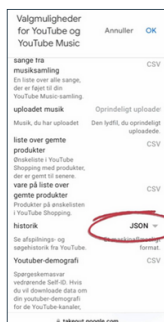
(After selecting having ordered the data correctly: Yes in 13.2.5 or 13.2.6)

Epinion

— making sense



Bestilte du data i formatet JSON?



- ☐ Ja
- ☐ Nej, jeg kom til at bestille i et andet format

Tilbage

Næste

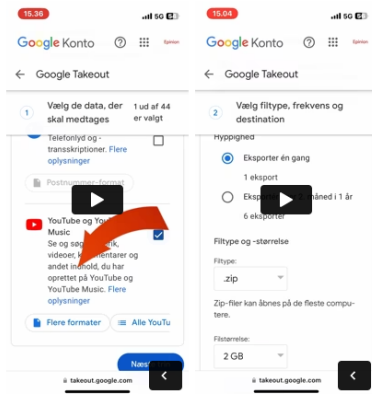
13.2.8 Reported wrong file format request

(Selecting no to having ordered the data in a JSON format in 13.2.7; Selecting need for help here will redirect as before to 13.2.3 and 13.2.4)

Epinion
- making sense



Da vi skal bruge data i formatet JSON, skal vi bede dig bestille dine data igen. Videoen nedenfor viser, hvordan du bestiller dine data hos Google Takeout



Har du bestilt dine data?

- ☐ Ja
- ☐ Nej, jeg har brug for hjælp til at bestille data

Tilbage

Næste

13.2.9 After confirming requesting correct data format: Request to check e-mail for data download options and invitation to Survey 3

(After having made sure to order the data correctly; Yes in 13.2.8)

Epinion
- making sense

Tak for at du har bestilt dine data.

Google Takeout behandler nu din bestilling.

Du modtager en e-mail på den e-mailadresse, der er tilknyttet din Google-konto, når dine data er klar. **Vi anbefaler, at du tjekker allerede nu** - det tager nemlig normalt ganske kort tid, før dine data er klar. I sjældne tilfælde kan det dog tage op til 5 dage afhængig af, hvor meget du har brugt YouTube.

Når dine data er klar, er du klar til tredje og næstsidste del af undersøgelsen, som du finder her:

Link til næste del af undersøgelsen.

Vi har også sendt linket til næste del af undersøgelsen på din e-mail [redacted] så du kan tilgå den, når du er klar.

Trin 1
Tilmelding

Trin 2
Bestil Youtube-data

Trin 3
Hent og send data

Trin 4
Spørgeskema

Husk, at hvis du har brug for hjælp, kan du ringe til [redacted] alle hverdage i tidsrummet 10 til 12 eller skrive en e-mail til [redacted]. Så vender vi tilbage hurtigst muligt.

13.3 Survey 3

13.3.1 Welcome screen to Survey 3



Velkommen til den tredje og næstsidste del af Aarhus Universitets undersøgelse af danskernes brug af YouTube.

I denne del guider vi dig til at hente og uploade dine YouTube-data

Hvis du har brug for hjælp undervejs, er der en hotline, du kan ringe til.

Næste

0%

Copyright © Epinion

13.3.2 Confirming reception of e-mail from Google takeout



Når dine filer er klar, vil du modtage en e-mail fra Google Takeout på den e-mailadresse, der er tilknyttet din Google-konto, hvor du kan downloade dine samlede filer.

Har du modtaget en e-mail fra Google Takeout om, at dine filer er klar til download?

☐ Ja

☐ Nej, jeg har ikke modtaget e-mailen fra Google Takeout endnu

Tilbage

Næste

0%

Copyright © Epinion

13.3.3 Timing of data request

(Following selecting no in 13.3.2)

Epinion
- making sense



Du modtager en e-mail på den e-mailadresse, der er tilknyttet din Google-konto, når dine data er klar. Vi anbefaler, at du tjekker allerede nu - det tager nemlig normalt ganske kort tid, før data er klar. Det kan dog i sjældne tilfælde tage op til 5 dage afhængig af, hvor meget du har brugt YouTube.

Hvis der er gået mere end 5 dage, anbefaler vi dig at bestille dine data igen.

Er der gået mere end 5 dage?

- ☐ Nej, ikke endnu
☐ Ja, det er længere siden

Tilbage

Næste

13%

Copyright © Epinion

13.3.4 Information on checking e-mails

(Following selecting no in 13.3.3)

Epinion
- making sense

Du modtager en e-mail på den e-mailadresse, der er tilknyttet din Google-konto, når dine data er klar. Vi anbefaler, at du tjekker allerede nu - det tager nemlig normalt ganske kort tid, før data er klar. Det kan dog i sjældne tilfælde tage op til 5 dage afhængig af, hvor meget du har brugt YouTube.

100%

Copyright © Epinion

13.3.5 Information on requesting data again

(Following selecting yes in 13.3.3)

Epinion

- making sense

Du bedes venligst bestille dine data igen. [Klik her for at komme til bestillingsguiden.](#)

Hvis du har spørgsmål, kan du ringe til [redacted] alle hverdage i tidsrummet 10 til 12 eller skrive en e-mail til [redacted]. Så vender vi tilbage hurtigst muligt.

100%

Copyright © Epinion

13.3.6 Respondent's device

(Following selecting yes in 13.3.2)

Epinion

- making sense

?

Hvilken type device bruger du til at gennemføre undersøgelsen?

- ☐ Computer med Windows styresystem (f.eks. af mærkerne Lenovo, HP, Dell eller Acer)
- ☐ Computer med MacOS styresystem (f.eks. MacBook eller iMac)
- ☐ iPad eller iPhone
- ☐ Anden tablet eller smartphone
- ☐ Jeg er i tvivl om hvilket device, jeg bruger til at gennemføre undersøgelsen

Tilbage

Næste

13%

Copyright © Epinion

13.3.7 Instructions for data download

(Here: selecting “Windows computer” in 13.3.6)



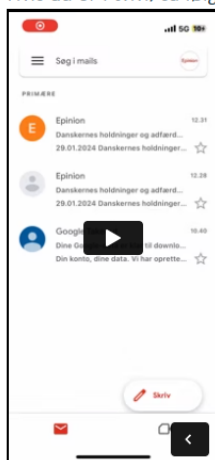
Nu er du klar til at downloade dine data.

Åben først den e-mail, som du har modtaget fra Google Takeout.

Klik så på knappen ‘**Download dine filer**’. Log ind med din Google-konto og vent på, at dine filer bliver downloadet.

1. Download data

Hvis du er i tvivl, så følg videoguiden



Når du har downloadet dine data, skal du klikke på ‘**Næste**’.

Tilbage

Næste

20%

Copyright © Epinion

13.3.8 Instructions renaming downloaded file

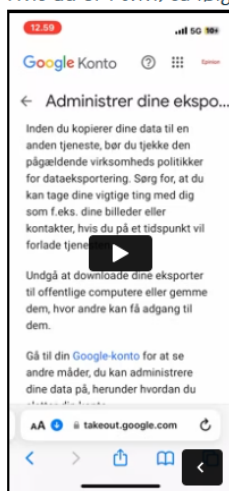


Nu skal du omdøbe filen. På en computer med Windows gør du dette ved at højreklikke på filen og trykke 'omdøb'. Din fil skal navngives **zrwg20.zip**

Du finder din downloadede .zip fil, 'takeout-... .zip', ved enten at **klikke på det færdige download i øverste højre hjørne af din browser eller ved at åbne den fra mappen 'Overførsler'** (på engelsk Downloads) fra Stifinderen.

2. Omdøb filen

Hvis du er i tvivl, så følg videoguiden



Har du navngivet filen **zrwg20.zip**?

- ☐ Ja, min fil hedder **zrwg20.zip**
- ☐ Nej, jeg kan ikke omdøbe filen

26%

Copyright © Epinion

13.3.9 Information reporting no successful renaming (Following selecting no in 13.3.8)

Epinion
- making sense



Du har skrevet at du ikke kan omdøbe din fil. En medarbejder ringer til dig hurtigst muligt for at hjælpe dig videre.

Hvornår på dagen må vi ringe til dig?

- ☐ 09:00-12:00
- ☐ 12:00-15:00
- ☐ 15:00-18:00
- ☐ 18:00-20:00

Tilbage

Næste

53%

Copyright © Epinion

13.3.10 Reporting successful renaming (Following selecting yes in 13.3.8)



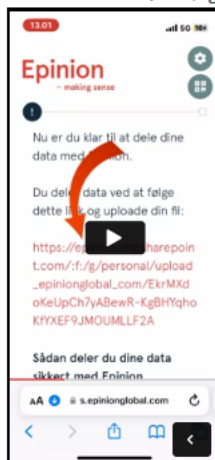
Nu er du klar til at dele dine data med Epinion.

Du deler data ved at følge dette link og uploade din fil. Her skal du også skrive til fornavn og efternavn:

Link til at uploade YouTube-data

3. Upload data

Hvis du er i tvivl, så følg videoguiden



Infoboks: Data deles via en sikret forbindelse. Det betyder, at det kun er Epinions projektmedarbejdere, der kan se de data, der uploades via linket. Du kan derfor ikke se hverken dine egne eller andres data, når de er uploadet til mappen. Data deles først med Aarhus Universitet, når de er pseudonymiserede.



Har du delt dine data med Epinion?

- ☐ Ja, jeg har delt
- ☐ Nej, jeg kan ikke dele filen

Tilbage

Næste

35%

Copyright © Epinion

13.3.11 Reporting unsuccessful upload (Following selecting no in 13.3.10)

Epinion
- making sense

?

Du har skrevet, at du ikke kan dele din fil. En medarbejder ringer til dig hurtigst muligt for at hjælpe dig videre.

Hvornår på dagen må vi ringe til dig?

☐ 09:00-12:00

☐ 12:00-15:00

☐ 15:00-18:00

☐ 18:00-20:00

Tilbage

Næste

80%

Copyright © Epinion

13.3.12 Reporting successful upload: Information about next steps (Following selecting yes in 13.3.10)

Epinion
- making sense

Tak for at du har uploadet dine data.

I morgen vil en medarbejder hos Epinion tjekke, at data er modtaget. Derefter får du adgang til den sidste del af undersøgelsen.

Vi sender dig linket til det sidste spørgeskema i en e-mail til din e-mailadresse

Trin 1
Tilmelding

Trin 2
Bestil Youtube-data

Trin 3
Hent og send data

Trin 4
Spørgeskema

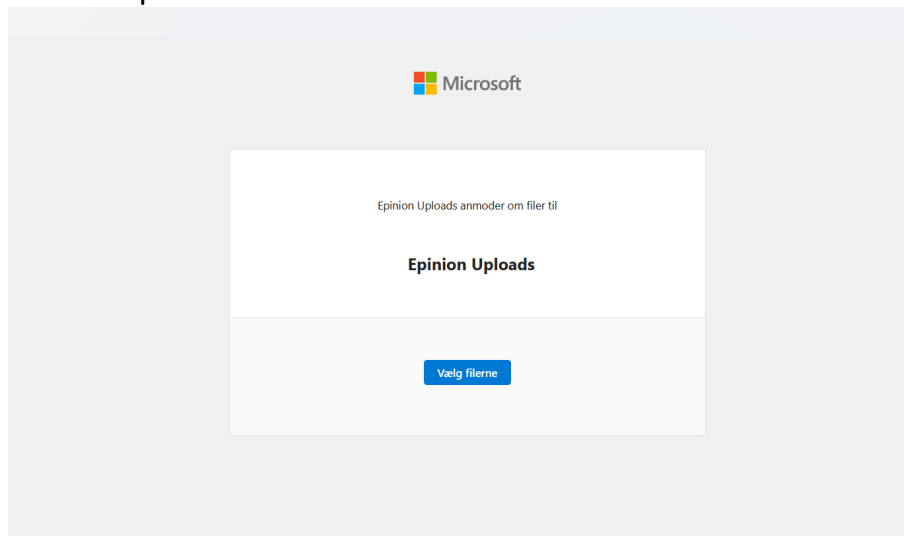
Hvis du har spørgsmål, kan du ringe til alle hverdage i tidsrummet 10 til 12 eller skrive en e-mail til . Så vender vi tilbage hurtigst muligt.

100%

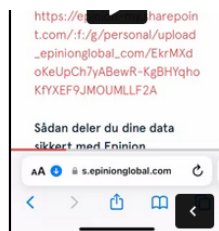
Copyright © Epinion

106

13.3.13 Upload link screenshot



13.3.14 Screen if upload error registered at Epinion



Infoboks: Data deles via en sikret forbindelse. Det betyder, at det kun er Epinions projektmedarbejdere, der kan se de data, der uploades via linket. Du kan derfor ikke se hverken dine egne eller andres data, når de er uploadet til mappen. Data deles først med Aarhus Universitet, når de er pseudonymiserede.



Har du delt dine data med Epinion?

Det ser ud til, at filen **fyzx49.zip** ikke har nået vores lager. Der kan være et problem med din upload. Kan du tjekke filnavnet eller prøve at uploade filen igen?

Hvis du har brug for hjælp, kan du ringe til XX XX XX i tidsrummet XX til XX eller skrive en e-mail til [redacted]

- ☒ Ja, jeg har delt
- ☐ Nej, jeg kan ikke dele filen

Tilbage

Næste

33%

Copyright © Epinion

13.4. Survey 4

13.4.1 Welcome screen to Survey 4

Epinion
- making sense



Velkommen til fjerde og sidste del af Aarhus Universitets undersøgelse af danskernes brug af YouTube.

I denne del vil vi bede dig besvare et spørgeskema. Det tager cirka 15-20 minutter at besvare.

Næste

13.4.2 Social media use (Randomisation)

Epinion
- making sense

?

I løbet af de sidste 7 dage, hvor mange dage har du brugt...?

	Alle 7 dage	6 dage	5 dage	4 dage	3 dage	2 dage	1 dag	Jeg har brugt platformen, men ikke inden for de sidste 7 dage	Jeg har aldrig brugt platformen	Ved ikke
... Twitch	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Rumble	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... WhatsApp	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... TikTok	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Snapchat	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Netflix	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Facebook	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... YouTube	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... 4Chan	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Meta Threads	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Messenger	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Pinterest	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Reddit	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Instagram	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Discord	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... X (tidligere Twitter)	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Telegram	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
... LinkedIn	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

2%

Copyright © Epinion

13.4.3 Number of followers, friends, and connections on social media platforms

(If selecting using the platforms → question not implemented for all social media platforms from 13.4.2, documentation here shows all social media platforms for which the question was asked; same order of items as in 13.4.2)

Epinion
- making sense



Hvor mange ...?

	0-499	500-999	1.000-4.999	5.000-9.999	10.000 eller flere	Ved ikke
... følgere har du på TikTok	☐	☐	☐	☐	☐	☐
... følgere har du på Snapchat	☐	☐	☐	☐	☐	☐
... venner har du på Facebook	☐	☐	☐	☐	☐	☐
... følgere har du på YouTube	☐	☐	☐	☐	☐	☐
... følgere har du på Instagram	☐	☐	☐	☐	☐	☐
... medlemmer har du på Discord	☐	☐	☐	☐	☐	☐
... følgere har du på X (tidligere Twitter)	☐	☐	☐	☐	☐	☐
... forbindelser har du på LinkedIn	☐	☐	☐	☐	☐	☐

Tilbage

Næste

13.4.4 Motivation for YouTube usage (Randomisation)

Epinion
- making sense



Hvor enig eller uenig er du i nedenstående udsagn? 1 er helt enig og 7 er helt uenig.

Jeg bruger YouTube...

	1 - Helt enig	2	3	4	5	6	7 - Helt uenig	Ved ikke
... som inspiration eller til at informere mig om nye tendenser	☐	☐	☐	☐	☐	☐	☐	☐
... i forbindelse med min hobby	☐	☐	☐	☐	☐	☐	☐	☐
... til at følge med i nyheder og aktuelle begivenheder	☐	☐	☐	☐	☐	☐	☐	☐
... til underholdning	☐	☐	☐	☐	☐	☐	☐	☐
... til at lære	☐	☐	☐	☐	☐	☐	☐	☐
... til at dele mit eget indhold	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

13.4.5 Media usage

(Randomisation after item 2)

Epinion
- making sense

?

I løbet af de sidste 7 dage, hvor mange dage har du brugt følgende medier som kilde til nyheder, enten i deres traditionelle formater eller online...?

	Alle 7 dage	6 dage	5 dage	4 dage	3 dage	2 dage	1 dag	Jeg har brugt mediet, men ikke inden for de sidste 7 dage	Jeg har aldrig brugt mediet
... DR	☐	☐	☐	☐	☐	☐	☐	☐	☐
... TV2	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Jyllands-Posten	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Den Korte Avis	☐	☐	☐	☐	☐	☐	☐	☐	☐
... BT	☐	☐	☐	☐	☐	☐	☐	☐	☐
... MSNBC	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Fox News	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Al Jazeera	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Ekstra Bladet	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Politiken	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Solidaritet	☐	☐	☐	☐	☐	☐	☐	☐	☐
... Berlingske	☐	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

9%

Copyright © Epinion

13.4.6 Following influencers?

Epinion
- making sense

?

I de følgende spørgsmål er vi interesserede i at høre om, hvilke personer, folk følger på sociale medier, som ikke er politikere.

Er der bestemte personer, som ikke er politikere, du følger, eller hvis indhold du ofte ser, på sociale medier?

- ☐ Ja
☐ Nej

Tilbage

Næste

9%

Copyright © Epinion

13.4.7 Interest in politics (If selecting no in 13.4.6)

Epinion
- making sense

?

Vil du sige, at du er meget interesseret i politik, noget interesseret, kun lidt eller slet ikke interesseret i politik?

- ☐ Meget interesseret
☐ Noget interesseret
☐ Kun lidt interesseret
☐ Slet ikke interesseret
☐ Ved ikke

Tilbage

Næste

11%

Copyright © Epinion

13.4.8 Voting intention

Epinion
- making sense



Hvad ville du stemme, hvis der var valg til Folketinget i morgen?

- ☐ Socialdemokratiet
- ☐ Radikale Venstre
- ☐ Det Konservative Folkeparti
- ☐ SF - Socialistisk Folkeparti
- ☐ Liberal Alliance
- ☐ Moderaterne
- ☐ Dansk Folkeparti
- ☐ Venstre, Danmarks Liberale Parti
- ☐ Danmarksdemokraterne - Inger Støjberg
- ☐ Enhedslisten - De Rød-Grønne
- ☐ Alternativet
- ☐ Et andet parti
- ☐ En kandidat uden for partierne
- ☐ Jeg ville stemme blankt
- ☐ Jeg ville ikke stemme
- ☐ Ved ikke

Tilbage

Næste

13.4.9 Personality traits

Epinion
- making sense

?

Hvor enig eller uenig er du i nedenstående udsagn?

Jeg ser mig selv som en, der...

	Helt enig	Delvist enig	Hverken eller	Delvist uenig	Helt uenig	Ved ikke
... er reserveret	☐	☐	☐	☐	☐	☐
... generelt er tillidsfuld	☐	☐	☐	☐	☐	☐
... har tendens til at være doven	☐	☐	☐	☐	☐	☐
... er afslappet, god til at håndtere stress	☐	☐	☐	☐	☐	☐
... har få kunstneriske interesser	☐	☐	☐	☐	☐	☐
... er udadvendt, social	☐	☐	☐	☐	☐	☐
... har tendens til at finde fejl ved andre	☐	☐	☐	☐	☐	☐
... udfører et grundigt stykke arbejde	☐	☐	☐	☐	☐	☐
... nemt bliver nervøs	☐	☐	☐	☐	☐	☐
... har en god fantasi	☐	☐	☐	☐	☐	☐

Tilbage

Næste

100%

Copyright © Epinion

13.4.10 Political orientation

Epinion
- making sense



I politik taler man ofte om venstre og højre. Hvor vil du placere dig selv på denne skala?

0 - Venstre	1	2	3	4	5	6	7	8	9	10 - Højre	Ved ikke
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

100%

Copyright © Epinion

13.4.11 Political efficacy (Randomisation)

Epinion
– making sense



Hvor enig eller uenig er du i følgende udsagn? 1 er helt enig og 7 er helt uenig.

	1 - Helt enig	2	3	4	5	6	7 - Helt uenig	Ved ikke
Jeg betragter mig selv som godt rustet til at deltage i politik og i samfundet	☐	☐	☐	☐	☐	☐	☐	☐
Undertiden er politik så indviklet, at folk som mig ikke rigtig kan forstå, hvad der foregår	☐	☐	☐	☐	☐	☐	☐	☐
Politiske partier er kun interesseret i at få min stemme, ikke i mine holdninger	☐	☐	☐	☐	☐	☐	☐	☐
Folk som mig har ikke noget at skulle have sagt, når det kommer til, hvad regeringen beslutter	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

22%

Copyright © Epinion

13.4.12 Political participation (Randomisation)

Epinion
— making sense

?

Hvor ofte har du deltaget i følgende aktiviteter i løbet af det seneste år?

	Meget ofte	Oft	Sommetider	Sjældent	Aldrig	Ved ikke
Kontaktet en politiker, journalist eller interesseorganisation via e-mail eller sociale medier for at påpege en bestemt problemstilling, du er optaget af	☐	☐	☐	☐	☐	☐
Bevidst ændret adfærd i din hverdag for at bakke op om en bestemt sag, f.eks. spist vegetarisk af hensyn til klimaet eller købt varer lokalt for at støtte lokale forretningsdrivende	☐	☐	☐	☐	☐	☐
Fulgt, blevet medlem af, eller opstartet en side, kanal eller gruppe på en social medieplatform for at støtte en bestemt sag eller en problemstilling, du er optaget af	☐	☐	☐	☐	☐	☐
Indsamlet eller doneret penge til et bestemt formål	☐	☐	☐	☐	☐	☐
Signaleret din holdning ved at gå med tøj eller bruge ting, der udtrykker et bestemt budskab, f.eks. en t-shirt, et net, en badge/nål, en kuglepen, en indkøbsmønt, eller klistermærker til din computer/mobil	☐	☐	☐	☐	☐	☐
Deltaget i en online protest eller bevægelse, f.eks. ved at tilføje en bestemt ramme til dit profilbillede eller bruge et bestemt hashtag på sociale medier	☐	☐	☐	☐	☐	☐

Tilbage

Næste

25%

Copyright © Epinion

13.4.13 Interest in news

Epinion
— making sense



Vil du sige, at du er særdeles interesseret i nyheder, meget interesseret, noget interesseret, ikke særligt interesseret, eller slet ikke interesseret i nyheder?

- ☐ Særdeles interesseret
- ☐ Meget interesseret
- ☐ Noget interesseret
- ☐ Ikke særligt interesseret
- ☐ Slet ikke interesseret
- ☐ Ved ikke

Tilbage

Næste

25%

Copyright © Epinion

13.4.14 News sources (Ranked, randomisation)

Epinion
— making sense



Hvor vigtige vil du sige, at disse medier er for dig som kilde til nyheder? Rangér venligst medierne hvor 1 er mest vigtig og 6 er mindst vigtig.

3 TV-kanalers nyhedssider online (f.eks. dr.dk, tv2.dk)

1 Sociale medier

5 TV og streaming af TV

4 Avisers nyhedssider online

2 Trykte aviser

6 Radio/podcast

Tilbage

Næste

20%

Copyright © Epinion

13.4.15 Naming of influencers (If selecting yes in 13.4.6)

Epinion
- making sense



Hvilke personer er der tale om? Skriv venligst navnet på mindst én og op til tre.

Du bestemmer selv, om du skriver deres fulde navn eller deres profilnavn.

Det er vigtigt, at du kun skriver navnet på én person i hvert tekstfelt.

- ☒ 1.
- ☒ 2.
- ☒ 3.

Tilbage

Næste



Copyright © Epinion

13.4.16 Social media platforms for following named influencer (Randomisation)

Epinion
— making sense

?

På hvilke sociale medier følger eller ser du indhold fra A, B og C? Kryds alle sociale medier af, hvor du følger/ser vedkommendes indhold.

	A	B	C
Twitch	☐	☐	☐
Rumble	☐	☐	☐
TikTok	☐	☐	☐
Snapchat	☐	☐	☐
Facebook	☐	☐	☐
YouTube	☐	☐	☐
4Chan	☐	☐	☐
Pinterest	☐	☐	☐
Reddit	☐	☐	☐
Instagram	☐	☐	☐
Discord	☐	☐	☐
X (tidligere Twitter)	☐	☐	☐
Telegram	☐	☐	☐
LinkedIn	☐	☐	☐
Ved ikke	☐	☐	☐

Tilbage

Næste

15%

Copyright © Epinion

13.4.17 Parasocial relationship (Randomisation, first mention in 13.4.15)

Epinion

– making sense

?

Hvis du tænker på A, hvor enig eller uenig er du da i nedenstående udsagn? 1 er helt enig og 7 er helt uenig.

	1 - Helt enig	2	3	4	5	6	7 - Helt uenig	Ved ikke
Jeg føler medlidenhed med vedkommende, når vedkommende begår fejl	☐	☐	☐	☐	☐	☐	☐	☐
Hvis der går længere tid end normalt, hvor vedkommende ikke poster noget, savner jeg indholdet	☐	☐	☐	☐	☐	☐	☐	☐
Vedkommende får mig til at føle mig godt tilpas, præcis som med en ven	☐	☐	☐	☐	☐	☐	☐	☐
Jeg ser frem til at se nyt indhold fra vedkommende	☐	☐	☐	☐	☐	☐	☐	☐
Jeg vil gerne møde vedkommende i virkeligheden	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

13.4.18 Beginning of following influencer (First mentioned in 13.4.15)

Epinion
- making sense

?

Hvornår begyndte du at følge eller se indhold fra influencer1?

- ☐ For mindre end en måned siden
- ☐ For 1-6 måneder siden
- ☐ For 7-12 måneder siden
- ☐ For 1-3 år siden
- ☐ For mere end 3 år siden
- ☐ Ved ikke

Tilbage

Næste

95%

Copyright © Epinion

13.4.19 Frequency of content consumption of first influencer mentioned (Note: After this question, it continues with 13.4.6–13.4.14)

Epinion
- making sense

?

Hvis du tænker på en typisk måned, hvor ofte ser du indhold fra influencer1?

- ☐ Flere gange om dagen
- ☐ Hver dag
- ☐ 5-6 dage om ugen
- ☐ 3-4 dage om ugen
- ☐ 1-2 dage om ugen
- ☐ Hver anden uge
- ☐ En gang om måneden
- ☐ Sjældnere
- ☐ Ved ikke

Tilbage

Næste

20%

Copyright © Epinion

13.4.20 Frequency and type of content consumed on YouTube (randomisation)

Epinion
- making sense

?

Hvor ofte ser du følgende type af indhold på YouTube?

	Meget ofte	Ofte	Sommetider	Sjældent	Aldrig	Ved ikke
Religion og spiritualitet	☐	☐	☐	☐	☐	☐
Hobby og DIY (do it yourself)	☐	☐	☐	☐	☐	☐
Nyheder	☐	☐	☐	☐	☐	☐
Skønhed og mode	☐	☐	☐	☐	☐	☐
Gaming	☐	☐	☐	☐	☐	☐
Sjove videoer (f.eks. skøre eksperimenter, pranks, fails)	☐	☐	☐	☐	☐	☐
Politik	☐	☐	☐	☐	☐	☐
Viden og læring	☐	☐	☐	☐	☐	☐
Sundhed og fitness	☐	☐	☐	☐	☐	☐
Vlogs og personlige videoer	☐	☐	☐	☐	☐	☐
Musik	☐	☐	☐	☐	☐	☐
Underholdning	☐	☐	☐	☐	☐	☐
Natur og dyr	☐	☐	☐	☐	☐	☐
Sport	☐	☐	☐	☐	☐	☐
Kunst og kultur	☐	☐	☐	☐	☐	☐
Film og serier	☐	☐	☐	☐	☐	☐
Teknologi	☐	☐	☐	☐	☐	☐
Indhold til børn	☐	☐	☐	☐	☐	☐
Andet, angiv venligst hvad: 	☐	☐	☐	☐	☐	☐

Tilbage

Næste

43%

Copyright © Epinion

13.4.21 Device use by other users than participant?

Epinion
- making sense



Er der andre, der bruger YouTube på samme computer/tablet/mobil som dig?

- ☐ Nej, jeg er den eneste, der bruger YouTube på den computer/tablet/mobil, hvor jeg bruger YouTube
- ☐ Ja, der er andre end mig, der bruger YouTube på den samme computer/tablet/mobil som mig (f.eks. barn eller partner)
- ☐ Ved ikke

Tilbage

Næste

43%

Copyright © Epinion

13.4.22 YouTube usage on which platform?

Epinion
- making sense



Hvordan ser du YouTube-videoer? Vælg alle relevante svarmuligheder.

- ☐ Gennem Instagram
- ☐ Gennem YouTube-appen
- ☐ Gennem LinkedIn
- ☐ Som Shorts (YouTube's egen version af TikTok/reels)
- ☐ Gennem Snapchat
- ☐ Gennem TikTok
- ☐ Gennem min internet-browser eller via link
- ☐ Gennem Facebook
- ☐ Andet, angiv venligst:
-
- ☐ Ved ikke

Tilbage

Næste

45%

Copyright © Epinion

13.4.23 YouTube use based on self-selection or recommendation?

Epinion

— making sense

?

Hvad beskriver bedst, hvordan du finder det indhold, du ser på YouTube?

- ☐ Jeg bruger søgefunktionen til at finde indhold og ignorerer anbefalinger
- ☐ Nogle gange følger jeg anbefalinger, og nogle gange bruger jeg søgefunktionen til at finde indhold
- ☐ Jeg ser mest indhold, jeg får anbefalet af YouTube
- ☐ Jeg ser mest indhold, hvor nogen har anbefalet det til mig/ har sendt et link til mig
- ☐ Jeg ser primært indhold fra kanaler, jeg følger
- ☐ Ved ikke

Tilbage

Næste

47%

Copyright © Epinion

13.4.24 YouTube viewing sessions prolonged?

Epinion

— making sense

?

Hvad beskriver bedst din måde at bruge YouTube på?

- ☐ Efter jeg har set det indhold, jeg ville, forsætter jeg med at se lignende indhold, som dukker op eller bliver anbefalet til mig af YouTube
- ☐ Jeg ser normalt kun, hvad jeg har planlagt og stopper så med at se YouTube
- ☐ Ved ikke

Tilbage

Næste

50%

Copyright © Epinion

13.4.25 Effects of YouTube use (Randomisation)

Epinion
- making sense

?

I hvilken grad vil du sige, at din YouTube-brug...

	I høj grad	I nogen grad	I lav grad	Slet ikke	Ved ikke
... gør dig i tvivl om, hvad der er sandt og falsk?	☐	☐	☐	☐	☐
... påvirker dine holdninger (f.eks. til en bestemt sag)?	☐	☐	☐	☐	☐
... får dig til at ønske, at du var bedre til bestemte ting?	☐	☐	☐	☐	☐
... gør dig nedtrykt eller deprimeret?	☐	☐	☐	☐	☐
... gør det sværere for dig at koncentrere dig om gøremål (f.eks. studie eller arbejde)?	☐	☐	☐	☐	☐
... påvirker dine handlinger (f.eks. ændrer din måde at gøre bestemte ting på)?	☐	☐	☐	☐	☐
... gør dig vred?	☐	☐	☐	☐	☐
... gør dig stresset?	☐	☐	☐	☐	☐
... giver dig mindre søvn?	☐	☐	☐	☐	☐
... er en belønning til dig selv for at komme igennem dagen?	☐	☐	☐	☐	☐
... får dig til at føle, at du bedre kan klare det, du skal?	☐	☐	☐	☐	☐
... får dig til at føle dig glad?	☐	☐	☐	☐	☐

Tilbage

Næste

54%

Copyright © Epinion

13.4.26 Exposure to misinformation or false news

Epinion

- making sense

?

Hvor ofte tror du, at du personligt har været eksponeret for misinformation eller falske nyheder...

	Meget ofte	Ofte	Sommetider	Sjældent	Aldrig	Ved ikke
... på YouTube?	☐	☐	☐	☐	☐	☐
... generelt?	☐	☐	☐	☐	☐	☐

Tilbage

Næste

54%

Copyright © Epinion

13.4.27 Type of misinformation exposed to

(If reporting not “Rarely”, “Never”, or “Don't know” in 13.4.26 (Item 1 or 2), randomisation)

Epinion

- making sense

?

Hvilke(t) emne(r) handlede pågældende misinformation eller falske nyheder om? Vælg alle relevante svarmuligheder.

- ☐ Sundhed
- ☐ Klima
- ☐ Kriminalitet
- ☐ Ligestilling
- ☐ Vacciner
- ☐ Krige og konflikter i verden
- ☐ Svindel (f.eks. hvor man franarres penge)
- ☐ Politik
- ☐ Immigration
- ☒ Økonomi
- ☐ Andet, angiv venligst hvad

Tilbage

Næste

57%

Copyright © Epinion

13.4.28 Confidence in identifying misinformation

Epinion
- making sense

?

Hvor sikker eller usikker føler du dig på, at du kan genkende misinformation eller falske nyheder, når du møder det?

- ☐ Meget sikker
- ☐ Noget sikker
- ☐ Ikke så sikker
- ☐ Slet ikke sikker
- ☐ Ved ikke

Tilbage

Næste

50%

Copyright © Epinion

13.4.29 Commenting behaviour on YouTube (Randomisation)

Epinion
- making sense



Hvor ofte bruger du YouTube's kommentarfunktion...

	Meget ofte	Ofte	Sommetider	Sjældent	Aldrig	Ved ikke
... til at rose en, der har lavet en video?	☐	☐	☐	☐	☐	☐
... til at udtrykke en politisk holdning?	☐	☐	☐	☐	☐	☐
... til at kommentere på en video, fordi den, der har lavet den, opfordrede til det?	☐	☐	☐	☐	☐	☐
... til at deltage i en samtale med en, der har lavet en video, eller med andre brugere?	☐	☐	☐	☐	☐	☐
... til at påpege forkert information i en video?	☐	☐	☐	☐	☐	☐
... til at stille spørgsmål?	☐	☐	☐	☐	☐	☐

Tilbage

Næste

63%

Copyright © Epinion

13.4.30 Conspiracy mentality (Randomisation)

Epinion
– making sense

?

I det følgende vil du blive præsenteret for nogle udsagn om samfundet. Angiv venligst hvor enig eller uenig du er i hvert udsagn.

	1 - Helt enig	2	3	4	5	6	7 - Helt uenig	Ved ikke
Jeg tror, at der er hemmelige organisationer, der har stor indflydelse på politiske beslutninger	☐	☐	☐	☐	☐	☐	☐	☐
Jeg tror, at statslige myndigheder nøje overvåger alle borgere	☐	☐	☐	☐	☐	☐	☐	☐
Jeg tror, at der sker mange vigtige ting i verden, som offentligheden aldrig får noget at vide om	☐	☐	☐	☐	☐	☐	☐	☐
Jeg tror, at politikere for det meste ikke fortæller os de sande motiver for deres handlinger	☐	☐	☐	☐	☐	☐	☐	☐
Jeg tror, at begivenheder, der på overfladen ikke ser ud til at hænge sammen, er resultatet af hemmelige aktiviteter	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

65%

Copyright © Epinion

13.4.31 Mental and physical health

Epinion
- making sense



I de seneste fire uger, hvor ofte er det sket...

	Meget ofte	Ofte	Sommetider	Sjældent	Aldrig	Ved ikke
... at du følte dig stresset eller presset på tid?	☐	☐	☐	☐	☐	☐
... at du følte dig nedtrykt og trist?	☐	☐	☐	☐	☐	☐
... at du følte dig afslappet og velafbalanceret?	☐	☐	☐	☐	☐	☐
... at du følte dig fuld af energi?	☐	☐	☐	☐	☐	☐
... at du havde stærke fysiske smerter?	☐	☐	☐	☐	☐	☐
... at du følte dig ensom?	☐	☐	☐	☐	☐	☐

Tilbage

Næste

68%

Copyright © Epinion

13.4.32 Trust in people

Epinion

- making sense



Synes du, at man kan stole på de fleste mennesker, eller synes du, at man ikke kan være forsigtig nok i omgangen med andre mennesker?

- ☐ Stole på de fleste
- ☐ Kan ikke være forsigtig nok
- ☐ Ved ikke

Tilbage

Næste

65%

Copyright © Epinion

13.4.33 Frequency political discussions

Epinion

- making sense



Hvor ofte sker det, at du...

	Meget ofte	Ofte	Sommetider	Sjældent	Aldrig	Ved ikke
... diskuterer politik med andre?	☐	☐	☐	☐	☐	☐
... forsøger at overbevise andre om dine politiske synspunkter?	☐	☐	☐	☐	☐	☐
... deler information om politiske emner med andre?	☐	☐	☐	☐	☐	☐

Tilbage

Næste

68%

Copyright © Epinion

13.4.34 Political topics discussed

(If answering yes (meget ofte, ofte, or sommetider) to any item from question 13.4.33)

Epinion

- making sense

?

Hvilke(t) politiske emne(r) er det, du diskuterer eller deler information om?

Tilbage

Næste

65%

Copyright © Epinion

13.4.35 Experiment Control group 1

(KONTROL1-group) ("Next" button occurs after 10 seconds)

Epinion

- making sense

!

Forestil dig, at du ser et indslag fra DR.

Indslaget handler om et politisk forslag om at indføre kønskvoter, så virksomheder i fremtiden skal have mindst 40 procent kvinder i deres bestyrelse. Baggrunden for forslaget er, at der i dag er færre kvinder end mænd i virksomheders ledelse. Indslaget forklarer, at argumenterne for forslaget er, at kønskvoter kan være med til at øge ligestillingen mellem mænd og kvinder og skabe mere diversitet i virksomheders ledelse.

Tilbage

70%

Copyright © Epinion

13.4.36 Experiment Control group 2

(“Next” button occurs after 10 seconds)

Epinion
— making sense



Forestil dig, at du ser et indslag fra TV2

Indslaget handler om et politisk forslag om at indføre kønskvoter, så virksomheder i fremtiden skal have mindst 40 procent kvinder i deres bestyrelse. Baggrunden for forslaget er, at der i dag er færre kvinder end mænd i virksomheders ledelse. Indslaget forklarer, at argumenterne for forslaget er, at kønskvoter kan være med til at øge ligestillingen mellem mænd og kvinder og skabe mere diversitet i virksomheders ledelse.

Tilbage

Næste

70%
Copyright © Epinion

13.4.37 Experiment Treatment group 1

(TREATMENT1-group; “Next” button occurs after 10 seconds)

Epinion
— making sense



Forestil dig, at du ser en video, influencer 1 har delt, som handler om noget andet end influencer 1 normalt taler om.

I videoen taler influencer 1 om et politisk forslag influencer 1 har hørt om i medierne. Forslaget handler om at indføre kønskvoter, så virksomheder i fremtiden skal have mindst 40 procent kvinder i deres bestyrelse. Baggrunden for forslaget er, at der i dag er færre kvinder end mænd i virksomheders ledelse. Influencer 1 forklarer, at argumenterne for forslaget er, at kønskvoter kan være med til at øge ligestillingen mellem mænd og kvinder og skabe mere diversitet i virksomheders ledelse.

Tilbage

75%
Copyright © Epinion

13.4.38 Support of presented political claim (If treatment group 1 or 2)

Epinion
- making sense

?

Hvad tror du, at influencer 1 synes om det omtalte politiske forslag?

- ☐ Synes, det er en god idé
- ☐ Synes, det er en dårlig idé
- ☐ Ved ikke

Tilbage

Næste

77%

Copyright © Epinion

13.4.39 Experiment Treatment group 2 (TREATMENT2-group, followed by 13.4.38)

Epinion
- making sense

!

Forestil dig, at du ser en video, influencer 1 har delt, som handler om noget andet end influencer 1 normalt taler om.

I videoen taler influencer 1 om et politisk forslag influencer 1 har hørt om i medierne. Forslaget handler om at indføre kønskvoter, så virksomheder i fremtiden skal have mindst 40 procent kvinder i deres bestyrelse. Baggrunden for forslaget er, at der i dag er færre kvinder end mænd i virksomheders ledelse. Influencer 1 forklarer, at argumenterne for forslaget er, at kønskvoter kan være med til at øge ligestillingen mellem mænd og kvinder og skabe mere diversitet i virksomheders ledelse.

I videoen siger influencer 1, at ligestilling er en vigtig sag og fortæller om at have skrevet under på en underskriftsindsamling, der støtter ligestilling.

Tilbage

77%

Copyright © Epinion

13.4.40 Agreement with political issue (Randomisation)

Epinion
- making sense



Hvor enig eller uenig er du i følgende udsagn? 1 er helt enig og 7 er helt uenig.

	1 - Helt enig	2	3	4	5	6	7 - Helt uenig	Ved ikke
Kampen for ligestilling er en vigtig dagsorden	☐	☐	☐	☐	☐	☐	☐	☐
MeToo-bevægelsen er gået for vidt	☐	☐	☐	☐	☐	☐	☐	☐
Når kvinder forlanger ligebehandling, er de i virkeligheden bare ude efter særbehandling	☐	☐	☐	☐	☒	☐	☐	☐
Kvinder oplever ofte kønsdiskrimination	☐	☐	☐	☐	☐	☐	☐	☐
Der bør indføres kønskvoter i virksomheders bestyrelser i Danmark	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste

77%

Copyright © Epinion

13.4.41 Political participation issue specific (Randomisation)

Epinion
- making sense



Hvor sandsynligt eller usandsynligt vil du sige, det er, at du ville gøre følgende? 1 er meget sandsynligt og 7 er meget usandsynligt.

	1 - Meget sandsynligt	2	3	4	5	6	7 - Meget usandsynligt	Ved ikke
Boykotte produkter fra virksomheder, fordi de ikke støtter op om ligestilling	☐	☐	☐	☐	☐	☐	☐	☐
Deltage i en demonstration til fordel for ligestilling	☐	☐	☐	☐	☐	☐	☐	☐
Reagere på, kommentere på, eller dele et opslag om ligestilling på sociale medier	☐	☐	☐	☐	☐	☐	☐	☐
Skrive under på en underskriftsindsamling online, der handler om ligestilling	☐	☐	☐	☐	☐	☐	☐	☐
Købe produkter fra virksomheder, fordi de støtter op om ligestilling	☐	☐	☐	☐	☐	☐	☐	☐

Tilbage

Næste



Copyright © Epinion

13.4.42 Perception of presented scenario (If in treatment group 1 or 2)

Epinion
- making sense



Hvor sandsynligt eller usandsynligt vil du sige, det er, at influencer 1 ville lave en video som den, du blev bedt om at forestille dig?

- ☐ Det kunne sagtens ske
- ☐ Det kunne måske godt ske
- ☐ Det ville nok ikke ske
- ☐ Det ville aldrig ske
- ☐ Ved ikke

Tilbage

Næste

84%

Copyright © Epinion

13.4.43 Household composition

Epinion
- making sense



Kan du sige, hvem der er del af den husholdning, du bor i? Vælg alle svarmuligheder, der beskriver din situation.

- ☐ Din partner (f.eks. din kone, din mand, din kæreste)
- ☐ Barn eller børn under 15 år, som bor hos dig hele tiden
- ☐ Barn eller børn over 15 år, som bor hos dig hele tiden
- ☐ Barn eller børn under 15 år, men kun noget af tiden (f.eks. pga. skilsmisse)
- ☐ Barn eller børn over 15 år, men kun noget af tiden (f.eks. pga. skilsmisse)
- ☐ Andre tætte relationer, f.eks. familiemedlemmer og/eller venner
- ☐ Andre personer
- ☐ Jeg bor alene
- ☐ Ved ikke

Tilbage

Næste

77%

Copyright © Epinion

13.4.44 Social class indicator

Epinion
- making sense



Er det i løbet af det sidste år hændt, at din husstand havde svært ved at betale de løbende udgifter til mad, husleje, transport mv.?

- ☐ Ja, ofte
- ☐ Ja, nogle gange
- ☐ Ja, sjældent
- ☐ Nej, aldrig
- ☐ Ved ikke

Tilbage

Næste

79%

Copyright © Epinion

13.4.45 Migration background

Epinion
- making sense



Er du født i Danmark?

- ☐ Ja, jeg er født i Danmark
- ☐ Nej, jeg er ikke født i Danmark

Er dine forældre født i Danmark?

- ☐ Begge mine forældre er født i Danmark
- ☐ Kun en af mine forældre er født i Danmark
- ☐ Ingen af mine forældre er født i Danmark
- ☐ Ved ikke

Tilbage

Næste

81%

Copyright © Epinion

13.4.46 Subjective location

Epinion
— making sense



Vil du sige, at du bor i ...?

- ☐ Et landområde eller landsby
- ☐ En lille eller mellemstor by
- ☐ En større by eller storby
- ☐ Ved ikke

Tilbage

Næste

84%

Copyright © Epinion

13.4.47 Final screen and debriefing of the treatment groups

Epinion

– making sense

Mange tak for din deltagelse i undersøgelsen!

Resultaterne fra undersøgelsen vil bidrage med ny forskningsviden om den danske befolknings brug af YouTube og holdninger til forskellige samfundsspørgsmål.

Undervejs i spørgeskemaet, du netop har besvaret, er du blevet bedt om at forestille dig et bestemt scenarie. Dette scenarie er fiktivt og fremstillet til brug i spørgeskemaundersøgelsen med det formål at opnå indsigt i, om og hvordan indhold på sociale medier kan påvirke politiske holdninger og handlinger.

Din deltagelse i undersøgelsen er frivillig. Du kan til enhver tid trække dit samtykke tilbage, hvorefter vi sletter dine oplysninger, hvis de ikke allerede er fuldt anonymiseret. Det kan du gøre ved at kontakte [REDACTED]
Hvis du har spørgsmål til undersøgelsen, kan du kontakte den forskningsansvarlige professor Anja Bechmann ved at sende en e-mail til [REDACTED] med "YouTube studie 2024" i emnefeltet.

[These lines are only displayed in TEST mode]

The survey is redirecting to:

<https://survey-epinion.dk/>

100%

Copyright © Epinion

13.4.48 Final screen

Epinion

– making sense

Mange tak for din deltagelse i undersøgelsen!

Resultaterne fra undersøgelsen vil bidrage med ny forskningsviden om den danske befolknings brug af YouTube og holdninger til forskellige samfundsspørgsmål.

Din deltagelse i undersøgelsen er frivillig. Du kan til enhver tid trække dit samtykke tilbage, hvorefter vi sletter dine oplysninger, hvis de ikke allerede er fuldt anonymiseret. Det kan du gøre ved at kontakte [REDACTED]
Hvis du har spørgsmål til undersøgelsen, kan du kontakte den forskningsansvarlige professor Anja Bechmann ved at sende en e-mail til [REDACTED] med "YouTube studie 2024" i emnefeltet.

[These lines are only displayed in TEST mode]

The survey is redirecting to:

<https://survey-epinion.dk/>

100%

Copyright © Epinion

Explore Nordicom's reports,
books, statistics, and more at
www.nordicom.gu.se



NORDICOM



DATLAB
CENTER FOR DIGITAL SOCIAL RESEARCH



DANMARKS FRIE
FORSKNINGSFOND
INDEPENDENT RESEARCH
FUND DENMARK



Nordicom is funded with
support from the
Nordic Council of Ministers

This publication was produced by DATALAB – Center for Digital Social Research at Aarhus University,
and published in collaboration with Nordicom at the University of Gothenburg.