Beyond the Black Box: A Multimodal Approach to Understanding In-App Communication

Patrick Zerrer / Paul Pressmann / Cornelius Puschmann / Philipp Krieter*

Mobile devices are increasingly central as sources of up-to-date information, making the precise recording of information behavior on these devices more relevant for research. Established methods of automated data collection are reaching their limits when capturing in-app communication, such as political content within social media applications. Based on two case studies, we present two approaches that enable the identification of exposure to relevant content and, to some extent, the collection of in-app content. The first case study focuses on identifying relevant exposure across different apps on a mobile device using app tracking and screen recordings. The second case study focuses on linking exposure to seen content and deriving respective content features to obtain an enriched dataset. We discuss the advantages and limitations of both approaches and present conceptual frameworks for processing and analyzing such data.

Key words: mobile tracking, screen recording, mobile sensing, android logs, in-app tracking

1. Introduction

Technological and media changes are contributing to the ongoing fragmentation of media diets (Hasebrink & Popp, 2006). This includes the emergence of new media formats, the growing relevance of social media platforms and messaging services, and the widespread adoption of smartphones (Behre et al., 2024). These developments coincide with the increasing datafication of everyday media and communication behaviors (Breiter & Hepp, 2018). Against this backdrop, automated observation of smartphone use through mobile tracking offers a promising approach for drawing insights into media diets and media-mediated interpersonal communication (Ohme et al., 2024).

Automated mobile tracking offers several advantages as a data collection method. The increasing fragmentation of media and information use makes accurate measurement through self-reporting challenging (Araujo et al., 2017; Parry et al., 2021; Scharkow, 2016). While some researchers have raised justified criticisms of automated tracking (Bosch et al., 2024), others highlight its benefits, noting that it allows for more precise measurement (Stier, Breuer, et al., 2020) and the exploration of duration, frequency, sequence, and content engagement (Adam et al., 2024; Christner et al., 2022). In addition, automated tracking

* Dr. Patrick Zerrer, University of Bremen, ZeMKI Center for Media, Communication and Information Research, Linzer Str. 4,28359 Bremen, Germany, pzerrer@uni-bremen.de, https://orcid.org/0000-0002-8827-1336;

Paul Pressmann, M.A., University of Bremen, ZeMKI Center for Media, Communication and Information Research, Linzer Str. 4, 28359 Bremen, Germany, pressmann@uni-bremen.de, https://orcid.org/0009-0008-1271-4429;

Prof. Dr. Cornelius Puschmann, University of Bremen, ZeMKI Center for Media, Communication and Information Research, Linzer Str. 4, 28359 Bremen, Germany, puschmann@uni-bremen.de, https://orcid.org/0000-0002-3189-0662;

Philipp Krieter, Düsseldorf University of Applied Sciences, Faculty of Media, Münsterstraße 165, 40476 Düsseldorf, Germany, philipp.krieter@hs-duesseldorf.de.

DOI: 10.5771/1615-634X-2026-1-31

provides a comprehensive view of user behavior across the entire smartphone, rather than being limited to individual platforms or apps. However, most current mobile tracking methods cannot capture in-app user behavior, such as watching a short news video on Instagram. Given that a substantial portion of relevant smartphone use occurs within apps, recording these usage episodes represents a central methodological challenge (Muise et al., 2024).

In this paper, we address two key research questions regarding the recording of user behavior within apps. First, how can research-relevant exposure within apps be identified (RQ1)? Second, how can the in-app content viewed by users be recorded (RQ2)? Answering RQ1 requires measuring general smartphone usage behavior and detecting events that are instrumental for addressing the research objectives. Addressing RQ2 involves a more detailed examination of the received and relevant audio-visual content by linking smartphone usage data with extracted content through a combination of scraping and automated content analysis. Additionally, we develop research principles for working with mobile tracking and screen recording data, drawing on both the literature and our case studies. We argue that this approach offers advantages over contemporary data donation strategies, which often place an undue burden on participants and are prone to coverage gaps and omissions.

2. Current state and challenges of automated tracking

The increasing differentiation and digitalization of media consumption have contributed to a growing use of automated tracking methods in communication research (Guess et al., 2020; Jürgens & Stark, 2022; Maier et al., 2025; Stier et al., 2021; Wojcieszak et al., 2023). This trend is largely driven by the recognition that traditional data collection methods, such as surveys and media diaries, often struggle to capture media usage with precision (e.g. Parry et al., 2021).

Many studies in communication science and related disciplines employ mobile or desktop tracking for data collection, each method offering distinct advantages and drawbacks (Adam et al., 2024; Stier et al., 2021; Stier, Kirkizh et al., 2020). All tracking approaches share the primary goal of recording exposure—that is, capturing a 'person's contact with relevant content along with the time at which it occurs. Some methods further combine exposure data with additional sources, using unique idenitifiers to ensure correct allocation. For instance, web content can be collected and linked to an individual participant's exposure. Studies employing different automated tracking methods include proxy-based tracking, browser-based (mobile) tracking, and app tracking (Clemm von Hohenberg et al., 2024).

One method for collecting online exposure data is the proxy approach. In this method, the proxy acts as a digital intermediary that records outgoing and incoming internet connections, providing information about accessed URLs and apps used (Christner et al., 2022). Proxy server-based tools intercept 'participants' requests and store a complete record of the content accessed (Menchen-Trevino & and Karr, 2012). In this way, the proxy functions as an invisible intermediary connecting the 'user's device to the internet, routing all traffic through itself.

Another approach for collecting 'users' exposure is through browser extensions, which record visited URLs—either in full or reduced to the domain—along with the time and often the duration of each visit (Stier, Kirkizh et al., 2020; Wojcieszak et al., 2023). This enables assessments of the number, frequency, variety, and time spent on specific websites, such as news outlets (Wojcieszak et al., 2023), politicians' social media profiles (Stier et al., 2018), or other content types, such as pornography (von Andrian-Werburg et al., 2023). Typically, this approach requires a predefined list of URLs relevant to the research focus (e.g., news sites), which is then compared to the tracked URLs to extract relevant data. Alternatively, tracked URLs can be coded manually or automatically to identify content relevant to the study. In either case, comparing tracked URLs poses practical challenges, particularly due to the so-called long-tail issue, where a large number of URLs are visited only infrequently.

If URLs are fully tracked, this exposure data can be enriched by scraping the content associated with the recorded URLs and using the URL as a unique identifier to link content to individual users (Clemm von Hohenberg et al., 2024; Kühnemann, 2021; Munzert & Nyhuis, 2019). This means that, for both desktop and mobile browsing, it is already possible to identify exposure and assign the corresponding content to the respective user using the URL as a unique identifier. However, this approach faces limitations in today's increasingly mobile-centric media landscape, as in-app content (e.g., Instagram posts) cannot be easily tracked, restricting data collection to browser-accessed sites. Accordingly, research encounters two main challenges when capturing in-app content: first, correctly identifying relevant exposure, and second, collecting the content that has been accessed and linking it to the recipient's exposure. Assigning content to in-app exposure is particularly because there is no generalizable identifier, such as a URL, available within apps.

The collection of mobile automated tracking data faces several challenges. Smartphones, equipped with numerous sensors, have become an integral part of daily life (Görland, 2020) and collect a wide variety of information through these sensors (Ferreira et al., 2015). Several commercial and open-source applications allow researchers to access this data for research purposes (Christner et al., 2022). These mobile tracking apps record which applications are used on a smartphone and share this information with researchers (Ferreira et al., 2015; Tong et al., 2022; Toth & Trifonova, 2021; Parry & Toth, 2025). This enables the collection of data on app-specific usage duration, such as time spent on Instagram, as well as general smartphone use, including overall screen time (Fan et al., 2021; Tong et al., 2020). Unlike URL-based tracking, this method captures app usage in a mobile context. However, the content within those apps remains inaccessible due to standard iOS and Android security features, making it difficult to infer either exposure to relevant content or the nature of the content viewed based solely on app-tracking data.

Illuminating this methodological "black box" is key advantage of screen capturing and screen recording (Krieter, 2019, 2020; Krieter et al., 2024). In this approach, a tracking application is installed on the device, typically a smartphone, that captures screenshots or records screen videos at regular intervals (Krieter et al., 2024; Reeves et al., 2021; Yee et al., 2023). Screen recording techniques enable automated analysis by generating log files derived from these recordings (Frisson et al., 2016; Krieter, 2020; Yee et al., 2023). This allows researchers to capture of not only general system events, such as opening an app like Instagram (Böhmer et al., 2011; McMillan et al., 2015), but also specific in-app activities, such as viewing or sharing content. These analyses rely on artificial intelligence, computer vision, and machine learning methods to analyze mobile screen recordings (Krieter, 2020; Reeves et al., 2021; Yee et al., 2023).

Depending on the technical setup, screenshots or videos may be transmitted directly to a server operated by the research team, which provides rich data but raises substantial privacy concerns. Alternatively, screen recordings can be 'processed locally on the participant's device, with only derived results uploaded to the 'server. While this latter approach yields less detailed information, it substantially enhances privacy protection (Krieter, 2020; Krieter et al., 2024). Accordingly, screen recording techniques can help identify relevant exposure within apps, although capturing and storing the viewed content itself remains challenging due to privacy-preserving restrictions'.

To summarize the current state of mobile tracking technologies, existing methods primarily enable the recording of web browsing on smartphones by logging visited URLs. When URLs are fully captured, exposure data can be enriched by scraping the corresponding webpages and matching their content to user exposure, using the URL as a unique identifier. However, this procedure is not feasible for in-app environments, such as social media or messaging applications. Consequently, significant methodological gaps remain, both in the identification of relevant in-app content and in the assignment of that content to recorded exposure. This paper addresses these gaps by posing two research questions: the identification of relevant in-app exposure (RQ1) and the linkage of in-app content to exposure (RQ2). Despite the limitations discussed above, automated tracking enables the unobtrusive capture of user behavior and, in some cases, content across multiple platforms, making it particularly well suited to studying the complex contemporary media landscape.

3. Towards a holistic approach to mobile tracking

Most people carry their smartphones with them throughout their everyday lives, and these devices continuously generate data related to daily events, either directly or indirectly. Mobile tracking data exhibit several distinctive characteristics, including high granularity, sequentiality, large volume, and the interweaving of device and software affordances with human actions. These features underscore the need for a comprehensive approach to mobile tracking in research design that extends beyond the technical implementation of data collection. Accordingly, this paper addresses issues of data protection and privacy, socio-technical considerations, and extensions to the operationalization of exposure.

3.1 Data protection and privacy

Due to the fine-grained and sequential nature of data collection on mobile app usage and screen-recorded content, tracking data may contain a wide range of highly sensitive and private information (Krieter, 2020). Such data may include information about health, dating, or finance applications, which can allow inferences about characteristics such gender, sexual orientation, or socioeconomic status. Accordingly, these data require particularly careful, responsible, and conscientious handling by researchers.

In a tracking study, it is essential to inform participants comprehensively and transparently about the functionality of the tracking system, the scope of the data collected, and how these data will be used, and to obtain their informed consent within the framework of the study. Providing a clear and comprehensible explananation of how app tracking and screen recording operate, as well as how participants' data are handled and stored, is an important way to ensure that participants are adequately informed.

All raw data should be pseudonymized as early as possible in the research process. Any identifiers that could directly or indirectly reveal the identity of participants must be removed or replaced with randomly generated codes. Where feasible, additional anonymization procedures should be applied to ensure that participants cannot be re-identified. Particular attention must be paid to the risk of de-pseudonymization through the combination of study data with third-party information.

Studies have shown that the central storage of data with third parties (e.g. commercial providers of cloud services) conflicts with users' privacy requirements (Hong et al., 2003; Spiekermann & Cranor, 2009). Therefore, it appears more advantageous to use a secure storage location on servers operated by 'one's own institution. Access to stored data must be restricted to authorized members and governed by strict authentication procedures.

3.2 Socio-technical considerations

Mobile automated tracking data constitute a type of digital trace that is often regarded as a more precise—and obstensibly superior—representation of social reality. However, such data primarily provide a partial and surface-level view of human interaction with smartphones. Relying exclusively on these data to reconstruct complex social and societal realities risks oversimplification and misrepresentation of the social world (Breiter & Hepp, 2018; Jungherr & Theocharis, 2017). There is a tendency to privilege what can be quantified through tracking over the inherent complexity of human behavior, with these observations being mistakenly treated as comprehensive and accurate (Jungherr, 2015). To mitigate this risk, a robust strategy that applies the concept of measurement error to digital traces, analogous to other data generating processes in the social sciences is thus called for (Sen et al., 2021).

At the same time, the assumption that mobile automated tracking data possess comprehensive validity often coincides with the belief in their neutrality. Yet, like all digital trace data, they are not neutral; they are produced through technical procedures implemented by powerful institutions (such as corporations and governments) with specific interests (Breiter & Hepp, 2018; Gitelman, 2013). Consequently, analyses of mobile automated tracking data engage with information that is institutionally controlled, generated for particular purposes, and inherently biased. Against this backdrop, it is essential to critically assess digital traces as indicators of human behavior and social reality (Breiter & Hepp, 2018).

3.3 Operationalization of media exposure

Although tracking data almost always include a timestamp for each data point, enabling conclusions about duration and sequence of use, the majority of studies in automated tracking research rely on a relatively superficial operationalization of media exposure. Often, exposure is aggregated across individual media outlets or broad types (e.g. legacy media). Automated tracking data, however, allow media exposure to be operationalized along multiple dimensions, including duration, frequency, order, and the content of visited websites or applications (Christner et al., 2022). Despite this potential, most studies focus primarily on frequency and duration, while neglecting order, sequence, and specific content. As a result, they overlook the complexity and sequential nature of actual media use behavior.

Therefore, usage behavior before, during, and after exposure is captured only inadequately, if at all. Yet, qualitative studies demonstrate that individuals exhibit a wide range of behaviors when reading, watching, or listening to content, which in turn affects reception (Merten, 2021; Wieland, 2023). For many of these behavioral patterns, such as scanning, which refers to the quick and superficial viewing of content (Meijer & Kormelink, 2015), both the temporal sequence and the duration of engagement are crucial for characterizing the behavior.

Incorporating temporal information into the measurement of exposure can provide a more accurate reflection of reality in the data and enhance the explanatory power of exposure in statistical models, as demonstrated by Richter & Stier (2022). Several notable studies have operationalized and analyzed sequential tracking data using information-theoretical approaches (Kulshrestha et al., 2021; Möller et al., 2020). The works of Möller et al. (2020), Jürgens and Stark (2022), and Kulshrestha et al. (2021) showcase the level of depth and detail that can be achieved with automated tracking data when it is properly prepared and operationalized. Nevertheless, this level of granularity is underutilized in most studies, which often resort to simple aggregation of usage times. While aggregation may suffice for some research questions, it risks overlooking valuable information that has already been collected.

Against this backdrop, mobile tracking data—characterized by its granularity, temporal sequence, and close connection to 'users' everyday experiences—presents specific challenges for research. These challenges include ensuring trustworthy data handling, accurately interpreting individual data points, and and effectively operationalizing the information collected. All of these factors must be carefully considered when designing a study using mobile tracking.

The characteristics of mobile tracking data and the implications for its handling are further illustrated in the first case study.

4. Case study I - Identification of relevant exposure

The first case study addresses the first research question (RQ1), focusing on the identification of research-relevant exposure on mobile devices. It examines how politically active young people engage with political and climate-related information on their smartphones. Within this context, we developed and conceptually assessed the use of screen recordings as a method for measuring media use.

4.1 Data collection and methods

To obtain a detailed picture of information behavior, a total of 25 participants were recruited. Their smartphone use was recorded, with consent, between 14 February 2022 and 6 May 2022. General smartphone usage and the detection of relevant events were measured using a combination of two mobile tracking applications: the app tracking application AWARE and the Keywordlogger developed by Krieter (2020; 2024). This approach resulted in a comprehensive dataset comprising over 9 million data points.

4.2 Mobile automated app tracking

In this case study, we used the AWARE framework as an open-source solution for collecting hardware-, software-, and human-based data from smartphones (Ferreira et al., 2015). The AWARE client was implemented to monitor app usage and overall smartphone activity. Each time a user switched to a different application, a new timestamp was recorded. Additionally, the screen sensor captured the screen's status, such as on/off and lock/unlock events. This setup enabled the collection of information on both the frequency of interactions with the smartphone and their duration.

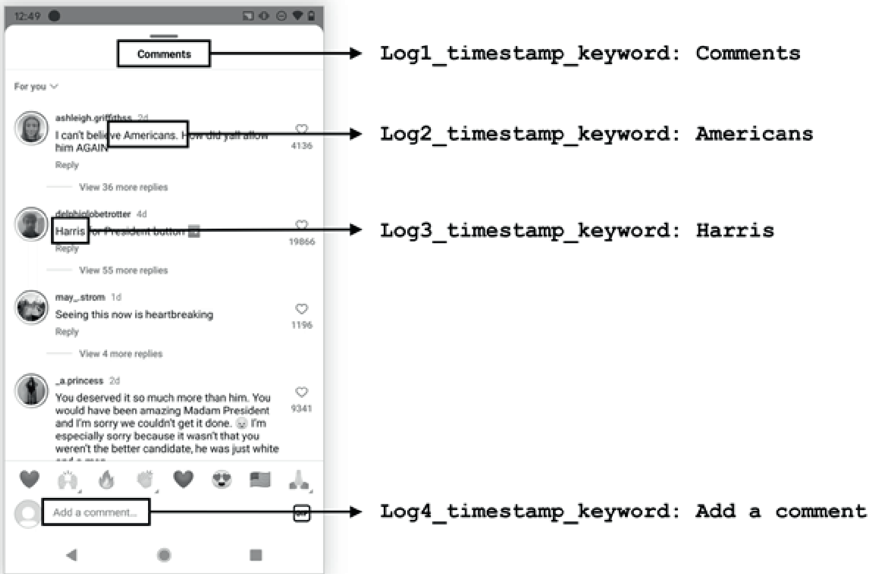
4.3 Screen recording

In addition to monitoring app usage, identifying relevant content within apps presents a central challenge. To address this, we applied screen recording methods adapted for communication science research as part of this case study.

Depending on the research focus, there are two main tracking approaches: (1) monitoring keywords related to smartphone or specific in-app functions (e.g., sharing, forwarding, etc.); (2) using content-related keywords that target specific vocabulary or references pertinent to particular topics (for instance, names of political parties, politicians, or terms describing specific media content such as Instagram accounts, newspapers, and TV channels). It is important that the use of keywords does not compromise privacy protections.

The Keywordlogger analyzes the screen recordings based on these previously defined keywords and records each visible keyword, with the timestamp, for each frame of the video examined. In this way, log files are generated which provide information about the occurrence, duration, and sequence of the defined keywords during the smartphone's usage time.

Figure 1: Example of how the Keywordlogger identifies keywords in the Instagram comments section.



The Keywordlogger essentially takes two approaches to strengthen the privacy of participants. First, the screen videos, which can contain very sensitive data, are only stored temporarily and do not leave the device, as they are only evaluated locally on the smartphone. The evaluation of the videos follows a fixed catalog of predefined events that are to be recognized in the videos. This means that only what is relevant for the research purpose is recognized and saved. The log files created locally on the device are transmitted to the research team in a pseudonymized form via a secure connection.

Second, the participants have the option of stopping the tracking throughout the entire study period by revoking the authorizations for the Keywordlogger & AWARE app on their Android smartphones. In this way, control over the data remains with the participants (Ferreira et al., 2015).

4.4 Operationalization of media exposure

The case study combines different data sources to create a comprehensive dataset that reflects the complex reality of user behavior. Enriching an initial dataset with additional information provided by other data collection methods, such as screen recording or other smartphone sensors, involves several steps.

Table 1: Data structure combined

Participant[1]	**Screen- and app-tracking by AWARE**						**Screen recordings by Keyword-logger**
	Smartphone display on and off	*Mobile session duration in seconds*	*Mobile application*	*Start of the mobile application*	*End of the mobile application*	*Usage duration in seconds*	*Keywords*
1000	08:33:56–08:35:14	78.05	Clock	08:34:03.988	08:34:06.757	2.77	Eis, Spiegel
1000	08:33:56–08:35:14	78.05	WhatsApp	08:34:06.757	08:34:07.436	0.68	[no keywords recorded]
1000	08:33:56–08:35:14	78.05	Mobile phone keypad[2]	08:34:07.436	08:34:15.012	7.58	Eis, Spiegel, Stadt, Erde
1000	08:33:56–08:35:14	78.05	WhatsApp	08:34:15.012	08:34:53.791	38.78	Wald, Erde, Luft, Wetter, Eis, Jahr, Wald, Erde, Abfall, Luft, Wetter, USA
1000	08:33:56–08:35:14	78.05	Mobile phone keypad[2]	08:34:53.791	08:35:04.034	10.24	Erde, Wetter, Handel, Eis, Erde, Luft, Erde, Luft, Wetter, Handel, Eis
1000	08:33:56–08:35:14	78.05	WhatsApp	08:35:04.034	08:35:12.552	8.52	Stadt, Jahr, Stadt
1000	08:33:56–08:35:14	78.05	WhatsApp	08:35:12.552	08:35:14.605	2.05	[no keywords recorded]

[1] Participant 1000 is the test device.
[2] AWARE records the app that is active in the foreground of the smartphone at that moment. The temporarily prior and subsequent app indicates within which app the keypad was used.

(1) *Data collection* using two different app tracking applications provides an initial raw dataset. We used app tracking and screen recording to capture information about app launches, screen status, and timestamped keywords. This information about smartphone use is constituted in several initial data sets, which must be processed and linked in the following.

(2) The *session concept* is a core conceptual component that enables different data sets to be linked together in a meaningful way. The session concept is based on the work of Peng and Zhu (2020), who understand uninterrupted usage behavior as a session. In this context, thinking in temporal order is a relevant part of the concept. In our understanding, the concept can be applied to different levels, resulting in a matryoshka-like data structure that

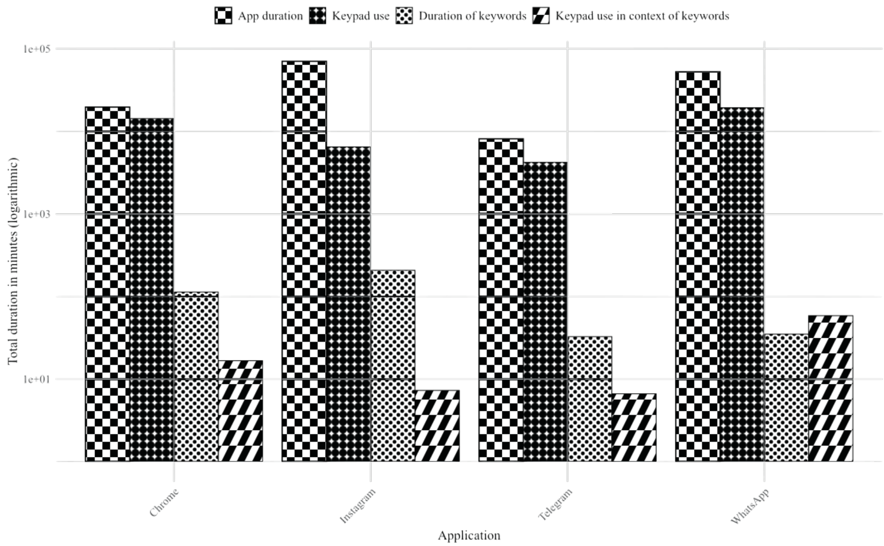
enables flexible handling of multi-level data. In this case study, several session levels were applied, first, *keyword sessions*, which comprise the time span of one or more visible terms; *app sessions*, which represent the uninterrupted usage sequence of an app; and *mobile sessions*, which comprise the aggregated usage behavior between unlocking and locking the smartphone. Among other things, this structure makes it possible to assign terms recorded by the Keywordlogger to the corresponding app session, e.g. Instagram or Whatsapp.

(3) *Contextual enrichment* describes precisely this step of correctly assigning nested session levels. For a better understanding, let us describe the initial situation: the first dataset contains the start and end times of the apps used on the smartphone, while the second dataset is a list of timestamps with the respective recorded keywords. The goal is to correctly assign the terms to the app that matches the time. Therefore, for each keyword it must be checked whether its timestamp falls within the time span of the app usage. The result is a data set that can be aggregated to different session levels without losing contextual information. Furthermore, aggregations at other levels, such as general app or screen usage time, are still possible.

4.5 Results and discussion

The first research question (RQ1) aims to identify research-relevant exposure within apps on smartphones. The combination of app tracking and screen recording enables us, in this case study focusing on the information usage of supporters of climate protests, to make statements about exposure to climate-relevant terms and actors on smartphones and within specific apps.

Figure 2: The visibility of climate-protest terms in apps (N = 25)



Note: An active keypad is labeled as "keypad use" in the context of keywords if at least one keyword is identified during the active and possibly interrupted period of keypad use. This means that the duration of keypad use may exceed the time span in keywords are actually visible.

The dataset created in this way allows for the operationalization of media exposure at different levels (across participants and within participants) and at different points in time (month, week, day, hour, etc.), while capturing multiple dimensions such as binary (non-)exposure, frequency, duration, order, or sequence. For example, the visibility of climate protest-related terms can be analyzed across the entire sample for each app (see Figure 2).

An even more granular analysis consists of calculating the duration of active keypad use with simultaneously visible climate protest terms per app (see Figure 2). Further breakdowns over time or comparisons between individuals can also be conducted.

In this way, the combination of app tracking and screen recording enables the identification of research-relevant constructs, allowing statements about the presence or absence of participants' exposure to specific actors or media-outlets. This approach makes it possible to examine questions regarding the range of information sources within social media apps or messengers, the frequency of exposure, and patterns of use throughout the day to be examined.

5. Case study II – Matching in-app exposure and content

Our second case study demonstrates how the limitations discussed above can be addressed by integrating content-level information with in-app exposure data. Specifically, this case study focuses on answering the second research question (RQ2), which concerns the recording of in-app content. We illustrate this by examining how participants' in-app exposure on TikTok is distributed across various public actors. Additonally, we investigate the share of politically related content on the platform. The case study measures political exposure on TikTok over a one-week period, beginning 1 September 2025.

5.1 Data collection and methods

Using accessibility-based mobile tracking, we recorded the in-app activity of 452 participants', capturing fragments of post descriptions and associated metadata whenever predefined potential political accounts (Zerrer, et al., 2024) or keywords appeared on screen. To reconstruct incomplete textual information, these fragments were matched with a parallel scraping database of TikTok posts from the observed accounts, based on string similarity alignment. The resulting dataset linked participants' exposure events to complete post-level content and corresponding engagement metrics, enabling a distinction between political and non-political content within each actor category. This approach allows for a deeper examination of the content participants received and interacted with, through a combination of automated data extraction, content analysis, and contextual interpretation (Ohme et al., 2024). Moreover, this step proposes a framework for integrating tracking data with context-rich metadata and insights from further content analysis procedures (Freelon et al., 2024; Reeves et al., 2021).

5.2 Capturing In-App Communication: Methodological Pathways

Analyzing in-app communication requires access to specific types of information, most importantly account names and post descriptions. Obtaining such information, however, is not straightforward, as different methodological approaches can vary considerably in coverage, granularity, feasibility, and privacy implications. One option is to adapt the approach outlined in the first case study, where app tracking was combined with keyword logging to approximate exposure to relevant content. This strategy allows for broad coverage across applications, as it does not require a priori restrictions to specific platforms. However, while this method provides valuable insights into exposure patterns, it remains limited in capturing complete content and its contextual richness. As a result, researchers often gain only partial visibility into what users actually encountered.

A second option is the so-called Screenomics framework (Reeves et al., 2021). Screenomics systematically records a continuous stream of screenshots or screen videos, enabling a highly detailed reconstruction of all visible content, including post descriptions and user interface elements. This method offers a comprehensive perspective on digital communication practices and supports rich contextual analyses. At the same time, it produces massive volumes of data and raises substantial privacy concerns, as continuous recording of all on-screen activities may capture highly sensitive personal information.

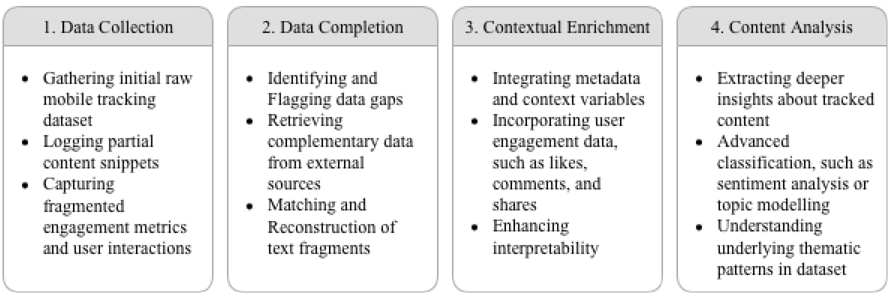
A third approach is to rely on a smartphone's Accessibility Service as a data collection instrument. Originally designed to assist users with disabilities—by enabling enhanced interaction options, such as reading screen content aloud, voice commands or automated actions—the system infrastructure also provides researchers with a powerful tool to study digital behavior on smartphones and obtain a more detailed and continuous record of activity (Andone et al., 2016). By leveraging the Accessibility Service, researchers can track event-based interactions, app usage patterns, displayed content, and user engagement across different platforms, offering more detailed information than screen recordings.

In practice, the Accessibility Service can be configured to capture a wide range of behavioral data directly at the system level. It records which applications are active, the duration and frequency of usage, and interactions within apps, such as clicks and scrolling patterns. Additionally, it can extract text displayed on screen, including advertisements, search queries, and other publicly available content from social media platforms like TikTok, YouTube, and Instagram. If text is not natively accessible through accessibility features, this process can be complemented by established data processing techniques, such as OCR (Optical Character Recognition) algorithms, to extract relevant information. Beyond app interactions, the Accessibility Service allows for the monitoring of visited websites, providing insights into browsing behavior by tracking domain visits and search activities.

5.3 Operationalization of media exposure and in-app content

Building on the methodological approach of the first case study, the second case study applies a structured, multi-stage approach to operationalizing and linking media exposure with in-app content. The main stages of this process are illustrated in Figure 3.

Figure 3: Data enrichment pipeline



Building on the general framework illustrated in Figure 3, the following subsections detail the procedures used to collect, match, and enrich mobile tracking data to reconstruct participants' exposure to political content.

(1.) *Data Collection* captures raw tracking data, using mobile tracking approaches, providing a foundational dataset. However, raw data may be incomplete and fragmented. In this case study, we used the Accessibility Service-based tracking approach to capture posts from predefined accounts of interests and keywords. The raw dataset includes basic smartphone interaction metrics, such as the timestamp a post appeared and the partial description text visible on participants' screens (Table 2, *Raw Tracking Data*).

Table 2: The data enrichment pipeline

1. Raw tracking data			**2. Data matching**	**3. contextual enrichment**		**4. content analysis**
User ID	Timestamp	Raw Description	Full-Text Description	Account	Account Type	Political content?
1000	2025-09-01 11:39:03	Bayer Leverkusen hat sich von seinem Trainer Erik ten Hag getrennt. Das bestätigt der Club via X. Der Niederländer wurde erst zu dieser Saison verpflichtet. Nach nur zwei Spieltagen ist der 55-Jährige fr…	Bayer Leverkusen hat sich von seinem Trainer Erik ten Hag getrennt. Das bestätigt der Club via X. Der Niederländer wurde erst zu dieser Saison verpflichtet. Nach nur zwei Spieltagen ist der 55-Jährige freigestellt worden. Es ist der erste Trainer-Rauswurf in der laufenden Bundesliga-Saison.	RTLWest	Media	No
1000	2025-09-01 21:30:14	Die Fakenews der Woche: Trump ist nicht tot, Es gibt keinen Killerrob...	Die Fakenews der Woche: Trump ist nicht tot, Es gibt keinen Killerroboter namens Noisy, es gibt kein Kopftuchverbot und keine Tankbeschränkung. #1minutejura #nachrichten #lernenmittiktok	Herr Anwalt	Influencer	Yes

Note: Data in the table have been anonymized to protect participants' privacy.

Further, we collected data to create a content library (cache) based on two predefined input sources: an *account list* and a *keyword list*. The account list contained official TikTok profiles of political actors, parties, and institutions, all manually verified as *profiles of interest*. The keyword list included political terms, enabling the detection of additional relevant exposures beyond the predefined accounts.

All profiles from the account list were then systematically scraped to create a content library. While API-based data collection is preferred for its compliance with platform Terms of Service and its ability to systematically gather public metadata, it has limitations such as rate restrictions and unpredictable policy changes (Freelon, 2018). When APIs are insufficient, web scraping can serve as an alternative (Perriam et al., 2020), though it raises ethical and legal concerns related to platform policies and data privacy laws (Trezza, 2023). The content library contains full post-level metadata, including descriptions, publication dates, and engagement metrics such as likes, views, and comments, and serves as a structured reference database for the subsequent enrichment step. The data enrichment process links the raw mobile tracking data with the specific TikTok content that participants were exposed to. The goal is to reconstruct and identify the exact posts that appeared on screen during the observation period, enabling the analysis of content and actor characteristics.

(2.) The *Data Matching* step is required to process the data collected in the initial stage, filling in missing content fragments and refining the dataset's integrity. A key challenge is the fragmentation of textual content captured by methods like OCR. Mobile tracking often produces incomplete text snippets rather than full descriptions or captions. This issue is particularly common on dynamic mobile platforms (e.g., YouTube, TikTok), where text frequently appears in scrolling interfaces or overlays, resulting in partial captures of the content viewed by users. For exampe, a user might see only a portion of a video description on screen, causing mobile tracking apps to record incomplete segments.

We performed exact and fuzzy string alignment, utilizing the Token-Sort-Ratio for its robustness in handling fragmented texts, to match the recorded text fragments with post descriptions from the scraped dataset. Through this matching process, 1,392 unique exposures were linked to corresponding posts in the content library. The resulting enriched dataset combined participants' individual exposure logs with additional post-level information.

(3.) *Contextual enrichment* integrates metadata and contextual variables, adding depth to the dataset and improving interpretability. In the second case study, this step incorporated additional metadata to expand and clarify the information retrieved through the matching process. After linking each exposure event to its corresponding TikTok post, we systematically added post-level context variables from the parallel scraping dataset. These included quantitative engagement indicators (likes, comments, shares, and view counts) as well as descriptive metadata such as publication date, video duration, and uploader identity. These additions transformed the enriched dataset from a purely behavioral log into a content-rich analytical resource. They allowed us to examine not only which political posts participants encountered, but also how visible and popular those posts were within the wider TikTok environment. For instance, by combining exposure frequency with view and like counts, we could distinguish between highly visible, viral posts and those with limited audience reach. Through this contextual enrichment, the dataset captured both the micro-level experiences of individual users and the broader communicative environment in which political content circulates (see Table 2).

(4.) The final *in-depth content analysis* applies advanced analytical techniques to extract further thematic and semantic insights. This step is designed to extract thematic and semantic insights from the TikTok posts linked to participants' exposure data. While the contextual enrichment stage added engagement metrics and structural information to the raw mobile tracking records, this stage goes further by using LLMs to classify previously matched TikTok video descriptions as either political or non-political, utilizing LLama 70B.

We note that, in this case study, we have by no means exhausted the possibilities of manual or automated methods of analyzing social media content. Further opportunities to expand the dataset include conducting more in-depth content textual analysis, or placing greater focus on image and video data.

Text-Level. The in-depth content analysis step systematically examines these complex aspects by integrating computational text analysis techniques, such as *automated sentiment and stance detection*. *Sentiment and stance detection* are key methods for understanding the emotional and ideological dimensions of digital content. Sentiment analysis applies natural language processing (NLP) techniques to classify text as positive, negative, or neutral, providing insights into the affective dimensions of media consumption (Liu, 2012). While sentiment analysis captures the emotional tone embedded in content, stance detection identifies ideological biases and patterns of selective exposure (Bestvater & Monroe, 2023; AlDayel & Magdy, 2021).

*Image and Video-Analysi*s. While the present case study focused primarily on textual content, future work could extend the data enrichment pipeline by incorporating additional research tools, such as visual and video content analysis, further broadening the analytical scope. Multimodal content analysis integrates computer vision, speech-to-text processing, and video analysis to assess the thematic and stylistic characteristics of images, videos, and multimedia posts. Techniques such as image classification and object recognition help categorize visual elements, distinguishing between different content types, including news graphics, political advertisements, and entertainment media. Speech recognition and automated transcription enable textual analysis of spoken content in video-based platforms like YouTube and TikTok, allowing researchers to apply sentiment and stance detection to audiovisual narratives. By combining textual, visual, and auditory cues, multimodal content evaluation provides a more comprehensive understanding of digital media exposure.

5.4 Results and discussion

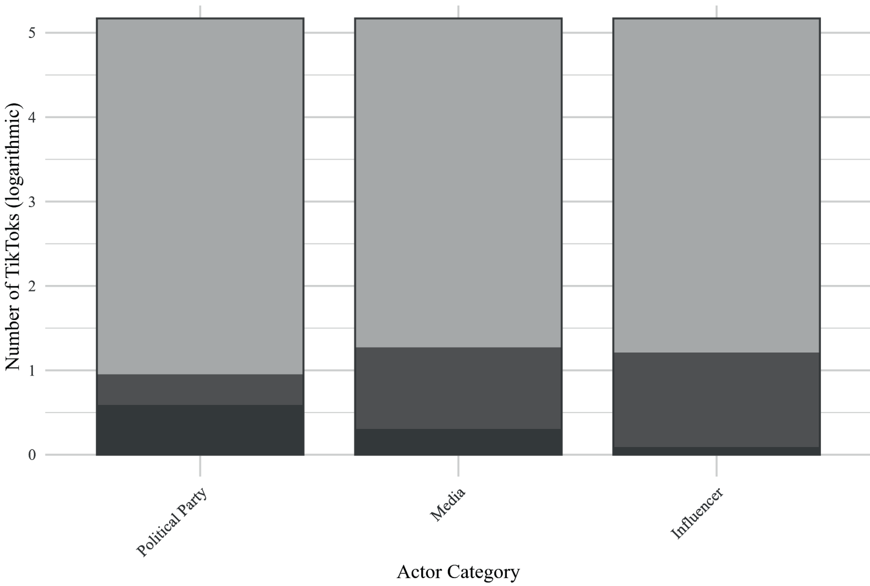
The second research question (RQ2) aims to capture research-relevant exposure within smartphone apps. Enriching the exposure data with a predefined list of relevant TikTok accounts, including media outlets, politicians, and influencers, enables analysis of participants' exposure to posts from these actors.

Figure 3 shows the daily average number of impressions per participant by actor category, with the number of TikToks consumed varying across participants and days.

Both political party and media content are present in the media mix. However, daily exposure to political posts is relatively low, suggesting that these actors adopt a more entertainment-oriented approach on TikTok. The comparatively low proportion of influencer content is also noteworthy. This may be due to limitations in identifying relevant accounts from the available account list, meaning that not all relevant posts were captured.

Nevertheless, this analysis demonstrates the potential of linking exposure with in-app content to gain insights into the individual composition of TikTok feeds within our sample.

Figure 3: Composition on TikTok Exposure per Actor Category (N=452)



Note: light gray = total daily TikToks per person; medium gray = category share; dark grey = share of political posts

In summary, the second case study adopts an integrated approach, transforming fragmented, surface-level tracking data into a structured dataset suitable for in-depth analysis. Each phase addresses specific limitations while building on the output of the previous step. While raw data alone often lacks coherence and context, targeted reconstruction and contextual layering make it possible to interpret not only *what* media content was encountered, but also *how*, *why*, and *with what potential impact*. The strength of this multi-process lies in its cumulative logic: each step enhances the interpretability and analytical depth of the data, enabling researchers to trace connections between exposure patterns, content characteristics, and user engagement. This highly customizable, layered structure opens new possibilities for research questions that cut across different steps, such as linking specific advertising themes to patterns of engagement, or examining content framing varies across sessions and platforms. In doing so, the pipeline not only improves data quality, but also expands the scope of what can be empirically observed using mobile tracking methods.

6. Conceptual Key Principles

When working with (mobile) automated tracking data, a distinctive feature is that data preparation constitutes a significant part of the project, as it can be complex and directly influence subsequent results.

One reason for the need for thorough processing of mobile automated tracking data is that such data was not originally created for scientific purposes, but primarily serves technical or administrative functions (Breiter & Hepp, 2018; Riebling, 2019). Consequently, it must be transformed into a format that is meaningful for research. It is important to recognize that, like all types of digital trace data, automated tracking data is not a completely unaltered raw material (Breiter & Hepp, 2018). In fact, this type of data is shaped by social institutions and technological products or devices they produce (Breiter & Hepp, 2018; Freelon, 2014). Given this context, it can be inferred that social, political, and societal assumptions may be embedded in the data (Breiter & Hepp, 2018).

We align with the perspective of Breiter & Hepp (2018) that tracking data, like other digital traces, reveals its full significance when viewed within a broader context or connected to real-world scenarios (Breiter & Hepp, 2018). Building on this understanding, we propose that the processing of tracking data should focus on five specific dimensions.

(1) The first dimension emphasizes that the privacy of the study participants must be protected, a principle reinforced by the concept of data minimalism. Accordingly, all information that is not relevant to the research interest is removed from the data to be analyzed. During pre-processing, additional measures were implemented to safeguard participant privacy. These included the categorizing apps to reduce the risk that third parties could comprimise the pseudonymization of the participants. Furthermore, sensitive and non-relevant apps—such as those related to health, dating, shopping, games, and finance—were replaced with placeholders.

(2) The second dimension involves the meaningful and targeted processing of automated tracking data into interpretable units of information. In our study, the goal was to obtain as comprehensive and context-rich a picture of information use as possible. To achieve this, we aggregated the data into sessions, defined as sequential, uninterrupted sequences of observed usage behavior (Peng & Zhu, 2020; Zhu et al., 2018). This session-based approach allows additional information about mobile usage to be captured and analyzed, while maintaining a manageable unit of analysis. It also enables the inclusion of previously or subsequently used apps or app categories facilitating the identification of broader usage patterns (Tong et al., 2022).

(3) The third dimension emphasizes the incorporation of a socio-technical perspective, including reflection on structural biases. Tracking data, as a form of digital trace data generated by technical systems for specific purposes, inherently exhibits biases (Breiter & Hepp, 2018). Researchers should take care to ensure that these biases are minimized, or ideally absent, in the research results. To achieve this, we recommend employing mixed-methods designs that incorporate user perspectives through quantitative or qualitative (survey-)methods. Combined with critical reflection on platform logics, this approach provides a foundation for examining methodological choices and analysis strategies.

(4) The fourth dimension involves data enrichment through the accurate linking of information from different sources to create a more comprehensive dataset that supports the analysis of smartphone usage behavior. By integrating recorded terms or content with an application that logs the start and end times of apps usage (Ferreira et al., 2015), it becomes possible to connect tracked events to specific applications. Data collected in this manner provides insights into the content participants accessed (as shown in Table 1 & 2), the channels through which they received the information (e.g., Telegram), and how they subsequently engaged with that content (e.g., continued communication).

(5) The fifth dimension emphasizes replicability by third parties. In this context, precise and comprehensible documentation of the research approach is crucial, as is the avoidance of unnecessarily complex procedures—adhering to the principle of keeping the process as simple as possible.

7. When to use which approach?

While the first approach focuses on identifying relevant exposure, the second goes further by capturing the associated content and linking it to that exposure. Both approaches also allow for the recording of device context—that is, upstream and downstream events on the mobile phone, including previously or subsequently opened apps, as well as exposure to content from different entities. They additionally capture the general duration of smartphone use, both overall and within specific applications. Accordingly, both approaches enable the mapping of a relatively accurate picture of smartphone usage practices. Despite these similarities, the two approaches differ in terms of (1) the type of content data, (2) the scope of data collection, and (3) the accessibility of the required data collection tools.

The identification of exposure enables the recording of research-relevant terms displayed on the smartphone across all apps. Unlike the second approach, there is no need to preselect specific social media platforms, allowing the first approach to adopt a more open and exploratory strategy. However, this openness comes at the cost of depth, as the actual content viewed by participants is not recorded.

The ability to capture content and link it to the corresponding exposure is the primary advantage of the second approach. Although it is less open due to the preselection of one or more social media platforms, it provides both the content viewed as and the associated social media metrics. Recording the actual content also enables further content analysis, which can integrate social media metrics with the individually tracked usage behavior of participants. Accordingly, this approach allows for the creation of a comprehensive picture, although it requires a certain degree of prior knowledge is required, such as lists of relevant accounts and keywords.

The granularity of the data is reflected in the scope of data collection. Collecting data to identify relevant exposure requires relatively fewer steps: first, the mobile tracking sensors (e.g., screen and app usage) are selected; next, the keywords are chosen; a database is set up (e.g., PostgreSQL or MySQL); and the appropriately configured apps are installed on participants' devices. Capturing content data and linking it to exposure, however, involves additional steps. This includes obtaining a unique identifier for each piece of content, setting up and configuring the scraper, and subsequently scraping the content. Consequently, the collection of content entails a more complex and multi-step data collection process.

Furthermore, the identification of exposure focuses on capturing a comprehensive record of user behavior on the entire smartphone. Due to its high granularity and the ability to aggregate data at different levels, this approach offers considerable flexibility in analytical options. For example, aggregated usage times per app across all participants can be examined alongside the visibility of climate-protest terms, such as *Fridays for Future*, within these apps (see Figure 2). Other possibilities include identifying instances when the smartphone keyboard is used simultaneously with the appearance of protest-relevant terms, which may indicate interpersonal or one-to-many communication in the context of climate protests.

In contrast, collecting content and linking it to exposure allows for more extensive analytical possibilities, including the examination of content characteristics, but it also requires additional prerequisites. Crucially, the content must be accessible, either through scraping or API access, which typically necessitates knowledge of an account name or post ID. Moreover, this approach is mostly limited to social media content and is only partially transferable to other mobile applications, such as messaging platforms. Nevertheless, it supports different levels of data aggregation, enabling further analyses of users' direct interactions with various types of content, particularly through recorded content features and engagement metrics, such as topics, stances, and likes.

8. Limitations of mobile tracking

Mobile automated tracking data presents several challenges and limitations. While it is often regarded as a reliable method for capturing digital information and media usage (Parry et al., 2021), it has notable drawbacks. This type of data provides a high level of detail and numerous measurement points, resulting in a dense and complex dataset.

The tracking app primarily logs apps brought to the foreground, meaning that only visible apps are recorded. As a result, background apps—particularly audio and music applications that continue playing without user interaction—may be underrepresented, potentially leading to an underestimation of audio usage duration.

Furthermore, most tracking apps—such as the AWARE client, Keywordlogger, and the software used in the second case study—are Android-only. This limitation may introduce sample bias, as iOS users are not represented in the study.

Most tracking apps, such as the AWARE client, also capture system apps and other background processes that are irrelevant to the research questions. These are categorized as system apps and excluded from analysis. However, this labeling process can introduce errors, as the sample includes a variety of smartphones with different Android versions, each of which may name system apps and background processes differently.

Additionally, the enrichment of tracking data depends on data quality. Data retrieved from screen recordings is often noisy and fragmented due to inconsistencies, incomplete logging, and platform variations. Similarly, API- and scraping-based data may also be incomplete, as public data is not always consistent, and official APIs impose restrictions or provide only sample datasets, which can make it difficult to accurately assign content to tracked exposure.

Another important consideration is how accurately automated tracking data reflects genuine human behavior. Two key aspects are particularly relevant: distinguishing between data points generated by human behavior and those resulting from technical artifacts. In mobile automated tracking data, it can be challenging to differentiate between actions stemming from human interaction, such as checking the time, and data produced by technical processes, such as a notification triggering the screen to turn on. Both types of data are recorded in the tracking logs, but only the former represents actual human engagement. Furthermore, tracking data cannot provide definitive information about 'users' attention to the content displayed on their devices. Accordingly, we advocate for a realistic understanding of the capabilities of mobile automated tracking and recommend employing mixed-methods approaches to more effectively capture the intricacies of human behavior.

9. Conclusion

Our paper illustrates the methodological possibilities of in-app tracking using two case studies. The first case study focuses on identifying relevant exposure within smartphone apps, while the second demonstrates how content can be captured, enriched, and linked to exposure. From these studies, we derive five conceptual core principles for working with in-app tracking data: privacy and data minimalism, meaningful preprocessing, incorporation of a socio-technical perspective, data enrichment, and replicability. Additionally, we highlight the strengths and limitations of both approaches.

This work contributes to the development of in-app measurement methods, tools, and standards, as well as to conceptual thinking about data handling and analysis. We hope that these methodological advances will inspire further research to illuminate in-app user behavior.

References

Adam, S., Makhortykh, M., Maier, M., Aigenseer, V., Urman, A., Gil Lopez, T., Christner, C., de León, E., & Ulloa, R. (2024). Improving the quality of individual-level web tracking: Challenges of existing approaches and introduction of a new content and long-tail sensitive academic solution. https://doi.org/10.1177/08944393241287793

AlDayel, A., & Magdy, W. (2021). Stance detection on Social Media: State of the art and trends. Information Processing & Management, 58(4), 102597. https://doi.org/10.1016/j.ipm.2021.102597

Andone, I., Błaszkiewicz, K., Eibes, M., Trendafilov, B., Montag, C., & Markowetz, A. (2016). How age and gender affect smartphone usage. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, 9–12. https://doi.org/10.1145/2968219.2971451

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. Communication Methods and Measures, 11(3), 173–190.

Behre, J., Hölig, S., & Möller, J. (2024). Reuters Institute Digital News Report 2024: Ergebnisse für Deutschland. Verlag Hans-Bredow-Institut. https://www.ssoar.info/ssoar/handle/document/94461

Bestvater, S. E., & Monroe, B. L. (2023). Sentiment is not stance: Target-aware opinion classification for political text analysis. Political Analysis, 31(2), 235–256. https://doi.org/10.1017/pan.2022.10

Böhmer, M., Hecht, B., Schöning, J., Krüger, A., & Bauer, G. (2011). Falling asleep with Angry Birds, Facebook and Kindle: A large scale study on mobile application usage. Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, 47–56.

Bosch, O. J., Sturgis, P., Kuha, J., & Revilla, M. (2024). Uncovering digital trace data biases: Tracking undercoverage in web tracking data. Communication Methods and Measures, 1–21. https://doi.org/10.1080/19312458.2024.2393165

Breiter, A., & Hepp, A. (2018). The complexity of datafication: Putting digital traces in context. Communicative figurations: Transforming communications in times of deep mediatization, 387–405.

Christner, C., Urman, A., Adam, S., & Maier, M. (2022). Automated tracking approaches for studying online media use: A critical review and recommendations. Communication Methods and Measures, 16(2), 79–95. https://doi.org/10.1080/19312458.2021.1907841

Clemm von Hohenberg, B., Stier, S., Cardenal, A. S., Guess, A. M., Menchen-Trevino, E., & Wojcieszak, M. (2024). Analysis of web browsing data: A guide. Social Science Computer Review, 42(6), 1479–1504. https://doi.org/10.1177/08944393241227868

Fan, Y., Tu, Z., Li, T., Cao, H., Xia, T., Li, Y., Chen, X., & Zhang, L. (2021). Understanding the long-term dynamics of mobile app usage context via graph embedding. IEEE Transactions on Knowledge and Data Engineering, 35(3). https://doi.org/10.1109/TKDE.2021.3110141

Ferreira, D., Kostakos, V., & Dey, A. K. (2015). AWARE: Mobile context instrumentation framework. Frontiers in ICT, 2. https://doi.org/10.3389/fict.2015.00006

Freelon, D. (2014). On the interpretation of digital trace data in communication and social computing research. Journal of Broadcasting & Electronic Media, 58(1), 59–75. https://doi.org/10.1080/08838151.2013.875018

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., Pruden, M. L., Malmer, D., Wu, Q., Xia, Y., Johnson, D., Chen, E., & Crist, A. (2024). What's in your PIE? Understanding the contents of personalized information environments with PIEGraph. Journal of the Association for Information Science and Technology, 75(10), 1119–1133. https://doi.org/10.1002/asi.24869

Frisson, C., Malacria, S., Bailly, G., & Dutoit, T. (2016). Inspectorwidget: A system to analyze users' behaviors in their applications. 1548–1554. https://doi.org/10.1145/2851581.2892388

Gitelman, L. (Ed.). (2013). "Raw data" Is an oxymoron. MIT Press.

Görland, S. (2020). Medien, Zeit und Beschleunigung: Mobile Mediennutzung in Interimszeiten. Springer.

Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. Nature Human Behaviour, 4(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

Hasebrink, U., & Popp, J. (2006). Media repertoires as a result of selective media use. A conceptual approach to the analysis of patterns of exposure. Communications, 31(3), 369–387. https://doi.org/10.1515/COMMUN.2006.023

Hong, J. I., Boriello, G., Landay, J. A., McDonald, D. W., Schilit, B. N., & Tygar, J. D. (2003). Privacy and security in the location-enhanced world wide web. Proceedings of Fifth International Conference on Ubiquitous Computing: Ubicomp, https://scholar.google.com/citations?view_op=view_citation&hl=en&user=MoFbcc0AAAAJ&cstart=300&pagesize=100&sortby=pubdate&citation_for_view=MoFbcc0AAAAJ:4TOpqqG69KYC.

Jungherr, A. (2015). Analyzing political communication with digital trace data: The role of twitter messages in social science research. Springer.

Jungherr, A., & Theocharis, Y. (2017). The empiricist's challenge: Asking meaningful questions in political science in the age of big data. Journal of Information Technology & Politics, 14(2), 97–109. https://doi.org/10.1080/19331681.2017.1312187

Jürgens, P., & Stark, B. (2022). Mapping exposure diversity: The divergent effects of algorithmic curation on news consumption. Journal of Communication, 72(3), 322–344. https://doi.org/10.1093/joc/jqac009

Krieter, P. (2019). Can I record your screen? Mobile screen recordings as a long-term data source for user studies. 1–10. https://doi.org/10.1145/3365610.3365618

Krieter, P. (2020). Looking inside—mobile screen recordings as a privacy friendly long-term data source to analyze user behavior. Universität Bremen. https://doi.org/10.26092/elib/103

Krieter, P., Zerrer, P., Puschmann, C., & Geise, S. (2024). Following topics across all apps and media formats: Mobile keyword tracking as a privacy-friendly data source in mobile media research. Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization, 126–131.

Kühnemann, H. (2021). Anwendungen des Web Scraping in der amtlichen Statistik. AStA Wirtschafts- Und Sozialstatistisches Archiv, 15(1), 5–25. https://doi.org/10.1007/s11943-021-00280-5

Kulshrestha, J., Oliveira, M., Karaçalık, O., Bonnay, D., & Wagner, C. (2021). Web routineness and limits of predictability: Investigating demographic and behavioral differences using web tracking data. 15, 327–338. https://doi.org/10.48550/arXiv.2012.15112

Liu, B. (2012). Sentiment analysis and opinion mining. Springer International Publishing. https://doi.org/10.1007/978-3-031-02145-9

Maier, M., Adam, S., Gil Lopez, T., Makhortykh, M., Bromme, L., Christner, C., de León, E., & Urman, A. (2025). Populist radical-right attitudes, political involvement and selective information consumption: Who tunes out and who prefers attitude-consonant information. Mass Communication and Society, 28(1), 101–129. https://doi.org/10.1080/15205436.2024.2310156

McMillan, D., McGregor, M., & Brown, B. (2015). From in the wild to in vivo: Video Analysis of Mobile Device Use. 494–503. https://doi.org/10.1145/2785830.2785883

Meijer, C. I., & Kormelink, G. T. (2015). Checking, sharing, clicking and linking: Changing patterns of news use between 2004 and 2014. Digital Journalism, 3(5), 664–679. https://doi.org/10.1080/21670811.2014.937149

Menchen-Trevino, E., & and Karr, C. (2012). Researching real-world web use with roxy: Collecting observational web data with informed consent. Journal of Information Technology & Politics, 9(3), 254–268. https://doi.org/10.1080/19331681.2012.664966

Merten, L. (2021). Block, hide or follow—Personal news curation practices on social media. Digital Journalism, 9(8), 1018–1039. https://doi.org/10.1080/21670811.2020.1829978

Möller, J., van de Velde, R. N., Merten, L., & Puschmann, C. (2020). Explaining online news engagement based on browsing behavior: Creatures of habit? Social Science Computer Review, 38(5), 616–632. https://doi.org/10.1177/0894439319828012

Muise, D., Markowitz, D., Reeves, B., Ram, N., & Robinson, T. (2024). (Mis) measurement of political content exposure within the smartphone ecosystem: Investigating common assumptions. Journal of Quantitative Description: Digital Media, 4.

Munzert, S., & Nyhuis, D. (2019). Die Nutzung von Webdaten in den Sozialwissenschaften. In C. Wagemann, A. Goerres, & M. Siewert (Eds.), Handbuch Methoden der Politikwissenschaft (pp. 1–25). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-16937-4_22-1

Naab, T. K., Karnowski, V., & Schlütz, D. (2019). Reporting mobile social media use: How survey and experience sampling measures differ. Communication Methods and Measures, 13(2), 126–147. https://doi.org/10.1080/19312458.2018.1555799

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

Parry, D., Davidson, B., Sewall, C., Fisher, J., Mieczkowski, H., & Quintana, D. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. Nature Human Behaviour, 5, 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Parry, D., & Toth, R. (2025). Extracting meaningful measures of smartphone usage from Android event log data: A methodological primer. Computational Communication Research, 7(1), 1.

Peng, T.-Q., & Zhu, J. J. H. (2020). Mobile phone use as sequential processes: From discrete behaviors to sessions of behaviors and trajectories of sessions. J. Comput. Mediat. Commun., 25(2), 129–146. https://doi.org/10.1093/jcmc/zmz029

Perriam, J., Birkbak ,Andreas, & and Freeman, A. (2020). Digital methods in a post-API environment. International Journal of Social Research Methodology, 23(3), 277–290. https://doi.org/10.1080/13645579.2019.1682840

Reeves, B., Ram, N., Robinson, T. N., Cummings, J. J., Giles, C. L., Pan, J., Chiatti, A., Cho, M., Roehrick, K., & Yang, X. (2021). Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them. Human–Computer Interaction, 36(2), 150–201. https://doi.org/10.1080/07370024.2019.1578652

Richter, S., & Stier, S. (2022). Learning about the unknown Spitzenkandidaten: The role of media exposure during the 2019 European Parliament elections. European Union Politics, 23(2), 309–329. https://doi.org/10.1177/14651165211051171

Riebling, J. (2019). The medium data problem in social science. In C. M. Stützer, M. Welker, & M. Egger (Eds.), Computational Social Science in the Age of Big Data. Concepts, Methodologies, Tools, and Applications (pp. 77–103). Herbert von Halem Verlag.

Scharkow, M. (2016). The accuracy of self-reported internet use—A validation study using client log data. Communication Methods and Measures, 10(1), 13–27. https://doi.org/10.1080/19312458.2015.1118446

Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. Public Opinion Quarterly, 85(S1), 399–422. https://doi.org/10.1093/poq/nfab018

Spiekermann, S., & Cranor, L. F. (2009). Engineering privacy. IEEE transactions on software engineering, 35(1), 67–82. https://doi.org/10.1109/TSE.2008.88

Stier, S., Bleier, A., Bonart, M., Mörsheim, F., Bohlouli, M., Nizhegorodov, M., Posch, L., Maier, J., Rothmund, T., & Staab, S. (2018). Systematically monitoring social media: The case of the German federal election 2017 (GESIS Papers). GESIS - Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.21241/ssoar.56149

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. Social Science Computer Review, 38(5), 503–516. https://doi.org/10.1177/0894439319843669

Stier, S., Kirkizh, N., Froio, C., & Schroeder, R. (2020). Populist attitudes and selective exposure to online news: A cross-country analysis combining web tracking and surveys. The International Journal of Press/Politics, 25(3), 194016122090701. https://doi.org/10.1177/1940161220907018

Stier, S., Mangold, F., Scharkow, M., & Breuer, J. (2021). Post post-broadcast democracy? News exposure in the age of online intermediaries. American Political Science Review, 1–7. https://doi.org/10.1017/s0003055421001222

Tong, L., Mingyang, Z., Hancheng, C., Yong, L., Sasu, T., & Pan, H. (2020). What apps did you use? : Understanding the long-term evolution of mobile app usage. 66–76. https://doi.org/10.1145/3366423.3380095

Tong, L., Tong, X., Huandong, W., Zhen, T., Sasu, T., Zhu, H., & Pan, H. (2022). Smartphone app usage analysis: Datasets, methods, and applications. IEEE communications surveys and tutorials, 24(2), 937–966. https://doi.org/10.1109/comst.2022.3163176

Toth, R., & Trifonova, T. (2021). Somebody's watching me: Smartphone use ttracking and reactivity. Computers in Human Behavior Reports, 4, 100142. https://doi.org/10.1016/j.chbr.2021.100142

Trezza, D. (2023). To scrape or not to scrape, this is dilemma. The post-API scenario and implications on digital research. Frontiers in Sociology, 8. https://doi.org/10.3389/fsoc.2023.1145038

von Andrian-Werburg, M. T., Siegers, P., & Breuer, J. (2023). A Re-evaluation of Online Pornography Use in Germany: A Combination of Web Tracking and Survey Data Analysis. Archives of Sexual Behavior, 1–13. https://doi.org/10.1007/s10508-023-02666-8

Wieland, M. (2023). Informiert oder (doch nur) abgelenkt? Potenziale und Herausforderungen automatisierter Nachrichtenkontakte in mobilen sozialen Medien. Herbert von Halem Verlag.

Wojcieszak, M., de Leeuw, S., Menchen-Trevino, E., Lee, S., Huang-Isherwood, K. M., & Weeks, B. (2023). No polarization from partisan news: Over-time evidence from trace data. The International Journal of Press/Politics, 28(3), 601–626. https://doi.org/10.1177/19401612211047194

Yee, A. Z. H., Yu, R., Lim, S. S., Lim, K. H., Dinh, T. T. A., Loh, L., Hadianto, A., & Quizon, M. (2023). ScreenLife Capture: An open-source and user-friendly framework for collecting screenomes from Android smartphones. Behavior Research Methods, 55(8), 4068–4085. https://doi.org/10.3758/s13428-022-02006-z

Zerrer, P. T. (2024). Political action and news use of the Fridays for Future movement in Germany. https://doi.org/10.26092/elib/3604

Zerrer, P., Puschmann, C., & Pressmann, P. (2025, January 28). German Politics Online: A List of social media accounts. https://doi.org/10.17605/OSF.IO/A7SJD

Zhu, J. J. H., Chen, H., Peng, T.-Q., Liu, X. F., & Dai, H. (2018). How to measure sessions of mobile phone use? Quantification, evaluation, and applications. Mobile Media & Communication, 6(2), 215–232. https://doi.org/10.1177/2050157917748351