

Digital Phenotyping via Passive Network Traffic Monitoring: Feasibility and Acceptability in University Students

Rameen Mahmood, Donghan Hu, Annabelle David, Zachary Beattie, Jeffrey Kaye, Nabil Alshurafa, Lou Haux, Josiah Hester, Andrew Kiselica, Shinan Liu, Chenxi Qiu, Chao-Yi Wu, Danny Yuxing Huang

Submitted to: Journal of Medical Internet Research
on: September 24, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	40
Figures	41
Figure 1.....	42
Figure 2.....	43
Figure 3.....	44
Figure 4.....	45
Figure 5.....	46
Please Select A Types.....	47
Please Select A Type 6.....	48
Figures	49
Figure 7.....	50
Figure 8.....	51
Figure 9.....	52
Figure 10.....	53
Figure 11.....	54
Figure 12.....	55

Digital Phenotyping via Passive Network Traffic Monitoring: Feasibility and Acceptability in University Students

Rameen Mahmood¹ BSc; Donghan Hu¹ BS, MS, PhD; Annabelle David¹ BA; Zachary Beattie² BS, PhD; Jeffrey Kaye³ BA, MD; Nabil Alshurafa⁴ BS, MS, PhD; Lou Haux⁵ BS, MS, PhD; Josiah Hester⁶ BSc, PhD; Andrew Kiselica⁷ BA, MA, PhD; Shinan Liu⁸ BEng, MS, PhD; Chenxi Qiu⁹ BS, PhD; Chao-Yi Wu⁹ BS, MS, PhD; Danny Yuxing Huang¹ BA, PhD

¹Department of Electrical and Computer Engineering Tandon School of Engineering New York University Brooklyn US

² Oregon Health & Science University Portland US

³ Oregon Alzheimer's Disease Research Center & Center for Aging & Technology (ORCATECH) Oregon Health & Science University Portland US

⁴ Northwestern University Chicago US

⁵ Max Planck Institute for Human Development Berlin DE

⁶ Georgia Institute of Technology Atlanta US

⁷ University of Georgia Athens US

⁸ University of Hong Kong Hong Kong HK

⁹ Neurology Harvard Medical School Harvard University Boston US

Corresponding Author:

Rameen Mahmood BSc

Department of Electrical and Computer Engineering

Tandon School of Engineering

New York University

370 Jay St

Brooklyn

US

Abstract

Background: Digital behaviors such as sleep, social interaction, and productivity reflect how individuals structure daily life. Among university students, online activity patterns mirror academic schedules, social rhythms, and lifestyle habits, with disruptions linked to sleep, stress, and well-being. Existing approaches—including wearables, apps, and surveys—yield useful insights but depend on self-report or active participation, limiting adherence in real-world use. Passive sensing of network traffic provides a scalable and less burdensome alternative, enabling unobtrusive capture of smartphone usage patterns while preserving privacy.

Objective: This study evaluated whether encrypted smartphone network traffic, collected via a standard virtual private network (VPN), can be used to capture patterns of digital behavior. We assessed feasibility (sustained data capture) and acceptability (usability, burden, and privacy perceptions), and examined whether traffic-derived features reveal aspects of digital behavior—including timing, intensity, and regularity—relevant to health and daily functioning.

Methods: We conducted a two-week prospective observational study at New York University. Thirty-eight students enrolled; 29 provided valid network data, 27 remained active for more than five days, and 25 completed the exit interview. Participants installed the WireGuard VPN client on personal smartphones, which enabled passive capture of encrypted network traffic. Feasibility was assessed across two domains: user retention and data coverage. Acceptability was evaluated using the System Usability Scale (SUS), NASA Task Load Index (NASA-TLX), and semi-structured exit interviews. Beyond evaluating feasibility and acceptability, we conducted exploratory analyses that visualized traffic-derived features in relation to digital activity patterns.

Results: Of the 29 participants who contributed valid data, 27 (93%) remained active for more than five days. Mean data coverage was 74.1% (median 77.1%). Participants contributed an average of 311.6 hours of monitored traffic (~13 days, SD 3.5), with totals ranging from 121 to 496 hours. Usability ratings were high (mean SUS score = 78) and perceived workload low (NASA-TLX scores minimal). Participants described the system as easy to install, unobtrusive, and generally trustworthy,

though some reported temporarily disabling the VPN during activities they considered private. No inferential statistical tests were conducted; analyses were descriptive. Exploratory analyses indicated that traffic-derived features reflected daily digital activity rhythms and revealed distinctive lifestyle patterns, including gaming and irregular late-night food delivery use.

Conclusions: VPN-based monitoring of encrypted smartphone traffic was feasible and acceptable, enabling sustained passive data collection with minimal burden. The findings demonstrate the potential of this approach as a scalable and device-agnostic method for digital phenotyping—capable of capturing fine-grained behavioral rhythms while preserving privacy. With broader validation and deployment, the technique could expand the toolkit for studying health, well-being, and cognitive function in everyday life. Clinical Trial: Not applicable. This study was not registered as a clinical trial because it did not involve randomization.

(JMIR Preprints 24/09/2025:84618)

DOI: <https://doi.org/10.2196/preprints.84618>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in a JMIR journal, my accepted manuscript PDF will be visible to all users.

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in a JMIR journal, my accepted manuscript PDF will be visible to all users.

Original Manuscript

Digital Phenotyping via Passive Network Traffic Monitoring: Feasibility and Acceptability in University Students

Abstract

Background

Digital behaviors such as sleep, social interaction, and productivity reflect how individuals structure daily life. Among university students, online activity patterns mirror academic schedules, social rhythms, and lifestyle habits, with disruptions linked to sleep, stress, and well-being. Existing approaches—including wearables, apps, and surveys—yield useful insights but depend on self-report or active participation, limiting adherence in real-world use. Passive sensing of network traffic provides a scalable and less burdensome alternative, enabling unobtrusive capture of smartphone usage patterns while preserving privacy.

Objective

This study evaluated whether encrypted smartphone network traffic, collected via a standard virtual private network (VPN), can be used to capture patterns of digital behavior. We assessed feasibility (sustained data capture) and acceptability (usability, burden, and privacy perceptions), and examined whether traffic-derived features reveal aspects of digital behavior—including timing, intensity, and regularity—relevant to health and daily functioning.

Methods

We conducted a two-week prospective observational study at New York University. Thirty-eight students enrolled; 29 provided valid network data, 27 remained active for more than five days, and 25 completed the exit interview. Participants installed the WireGuard VPN client on personal smartphones, which enabled passive capture of encrypted network traffic. Feasibility was assessed across two domains: user retention and data coverage. Acceptability was evaluated using the System Usability Scale (SUS), NASA Task Load Index (NASA-TLX), and semi-structured exit interviews. Beyond evaluating feasibility and acceptability, we conducted exploratory analyses that visualized traffic-derived features in relation to digital activity patterns.

Results

Of the 29 participants who contributed valid data, 27 (93%) remained active for more than five days. Mean data coverage was 74.1% (median 77.1%). Participants contributed an average of 311.6 hours of monitored traffic (~13 days, SD 3.5), with totals ranging from 121 to 496 hours. Usability ratings were high (mean SUS score = 78) and perceived workload low (NASA-TLX scores minimal). Participants described the system as easy to install, unobtrusive, and generally trustworthy, though some reported temporarily disabling the VPN during activities they considered private. No inferential statistical tests were conducted; analyses were descriptive. Exploratory analyses indicated that traffic-derived features reflected daily digital activity rhythms and revealed distinctive lifestyle patterns, including gaming and irregular late-night food delivery use.

Conclusions

VPN-based monitoring of encrypted smartphone traffic was feasible and acceptable, enabling sustained passive data collection with minimal burden. The findings demonstrate the potential of this approach as a scalable and device-agnostic method for digital phenotyping—capable of capturing fine-grained behavioral rhythms while preserving privacy. With broader validation and deployment,

the technique could expand the toolkit for studying health, well-being, and cognitive function in everyday life.

Keywords

Feasibility Studies; Passive Monitoring; Digital Phenotyping; Virtual Private Network (VPN); Network Traffic

Trial Registration

Not applicable. This study was not registered as a clinical trial because it did not involve randomization.

Introduction

Background

Digital behaviors—defined by how people interact with devices and digital platforms—offer insight into routines, social relationships, and overall health. Among university students, digital behaviors mirror academic, social, and lifestyle demands, and irregularities have been tied to stress and poorer health [84–87]. Beyond short-term effects, disruptions in behavioral rhythms, such as sleep disturbance, are increasingly recognized as early indicators of psychiatric and neurocognitive disorders, including depression [23,32–34,66], bipolar disorder [31, 35–37], schizophrenia [24,25], Alzheimer’s disease [3–12,60,61], and Parkinson’s disease [16,17]. Notably, these changes often emerge years before clinical diagnosis [134–137], highlighting the potential of passively collected digital behaviors as low-burden, noninvasive markers for earlier detection of health risks.

Yet, capturing such behaviors in real-world contexts remains constrained by the limitations of current measurement approaches. For example, polysomnography (PSG) is the gold standard for quantifying sleep architecture and fragmentation, but it is expensive, invasive, and typically limited to single-night laboratory assessments [67, 68]. Wearable and sensor-based systems offer the possibility of extending monitoring into daily life [38–41], though their adoption is often limited by the psychological burden of continuous device use, adherence challenges, battery consumption, restricted coverage, and behavioral changes they impose, [45,46,52,55–60,68,89]. These limitations have motivated researchers to seek more naturalistic, low-burden sources of behavioral data, most notably smartphones [101–103].

Deeply embedded in the daily routines of young adults, smartphones generate continuous digital traces that reflect a wide range of everyday rhythms, including social interaction, online engagement, mobility, and rest–activity cycles [101,104]. Prior work has shown that messaging activity patterns can reveal diurnal rhythms [92, 93], while periods without device use can approximate disengagement, such as sleep or downtime [90,91]. While such signals cannot capture fine-grained quality measures, they provide scalable proxies for broader temporal patterns of behavior. In our study, we leverage passive digital traces—captured through encrypted network traffic—from everyday smartphone use to estimate key behavioral regularities, such as activity onset and offset, usage intensity, and daily rhythm stability, without relying on self-reports, surveys, or other active input from users.

Prior Work

Translating digital traces into meaningful behavioral indicators requires methods that are scalable, minimally burdensome, and privacy-preserving. Yet achieving this balance has proven difficult under current mobile operating system constraints. iOS and Android impose strict limits on continuous background monitoring by third-party sensing apps. Built-in analytics tools like Apple’s Screen Time and Android’s Digital Wellbeing provide only coarse-grained metrics—such as hourly, daily, or

weekly app totals—and lack contextual detail, including timing, sequencing, and situational context [105,106,114,129,130]. More intrusive approaches—such as rooting, jailbreaking, or continuous screen capture—are burdensome to participants, disruptive to daily device use, and raise privacy concerns [107,108]. As a result, current tools either overburden participants or yield only fragmented signals, leaving a critical gap in methods that can link everyday digital behaviors to health and well-being at scale [90].

Several efforts have attempted to bridge this gap by using smartphone interaction data as a proxy for digital behavioral rhythms. Studies in young adults show that late-night device activity can signal irregular rest patterns [69,71], while excessive nighttime use has also been linked to depressive symptoms [70,71]. Beyond sleep, other forms of digital engagement—such as messaging frequency, app usage patterns, and periods of disengagement—have been studied as potential indicators of mood, social well-being, and cognitive function [91,109–113]. Collectively, these associations highlight the promise of digital behaviors as scalable markers of overall well-being, even though current methods for capturing them remain technically and ethically constrained.

Current approaches reveal a clear trade-off between burden and precision. Self-report methods, while simple to administer, are limited by recall errors, social desirability bias, and participant fatigue when repeated frequently [94,138,139]. By comparison, high-burden methods, such as custom tracking apps, can deliver greater accuracy but demand intrusive permissions and elevated privileges [52,54,58]. Beyond these approaches, mobile sensing platforms such as EARS and Beiwe have been widely deployed in behavioral health research with adolescents and young adults [96–98]. While these apps enable collection of location (GPS) traces, accelerometer data, and app usage logs, they also illustrate common limitations of dedicated phone monitoring: dependence on OS-level APIs (e.g., iOS Screen Time), participant burden from whitelisting or daily enabling of data collection, and concerns about battery consumption and usability. These constraints can substantially limit data yield. For example, in a large-scale deployment of the EARS app involving 4,754 adolescents, fewer than half (1,463) contributed analyzable data, underscoring the friction introduced by participant-dependent steps and technical restrictions [99]. Our approach addresses these constraints by collecting network traffic data, thereby operating at the network level rather than the operating system level. This design makes our solution agnostic to both device type (e.g., mobile, computer) and underlying operating system (e.g., iOS, Android, Windows, Mac), enabling capture of digital engagement across devices and contexts—including, but not limited to, smartphone interactions—while minimizing participant burden.

While our approach emphasizes network-level traces as a scalable and low-burden data source, other paradigms aim for maximal detail at the content level. Screenomics (the record of individual experiences represented as a sequence of screens that people view and interact with over time [140]), for example, is notable for its fine-grained, content-level capture: by recording screenshots every five seconds, it provides unparalleled visibility into moment-to-moment digital activity [95]. Yet these very strengths also create significant drawbacks. Continuous screenshot collection is highly intrusive, raises substantial privacy concerns, and scales poorly for everyday, long-term use [99,149,150]. Even with safeguards, participants remain aware that their screen contents are logged, which may alter natural behavior. As a result, while screenomics demonstrates what is technically possible at the content level, it is not feasible as a scalable, low-burden approach to behavioral monitoring.

These limitations highlight the need for methods that balance resolution with scalability and privacy. Our study pursues this balance by capturing encrypted network traffic via a VPN tunnel (see Methods for details). By design, this approach does not capture message content or typed input (e.g., the text of a message), thereby preserving privacy while still providing temporal markers of engagement. In practice, the metadata we collected in this study provided sufficient signal to infer a range of behavioral rhythms—such as social media use, streaming sessions, or gaming—while requiring minimal participant effort. In this sense, our system offers a privacy-preserving approach to

behavioral monitoring: it forgoes content-level detail but achieves scalability, minimizes participant burden, and enables longitudinal analysis of everyday digital life in naturalistic settings.

Goal of This Study

The goal of this study was to evaluate the feasibility and acceptability of this VPN-based approach and to explore whether network traffic-derived features can capture daily behavioral patterns relevant to mental health and well-being. Unlike prior tools, it requires no app whitelisting or custom keyboards, imposes negligible battery overhead, and avoids intrusive permissions, making it suitable for longitudinal deployments. The network traffic data we collected mirrors what Internet Service Providers or mobile carriers routinely observe, but with stronger safeguards: all traffic is anonymized and never shared with third parties [115]. Such capabilities are critical for advancing studies that aim to clarify links between screen use, sleep, behavioral and psychiatric outcomes in young adults. Building on this rationale, we structured the study around the following aims:

- **Aim 1 (Feasibility):** Assess whether encrypted network traffic can be continuously and passively captured on participants' devices and leveraged to accurately infer app usage.
- **Aim 2 (Acceptability):** Evaluate participants' perceptions of the monitoring method, with emphasis on usability, privacy, and burden.
- **Exploratory Aim:** Explore whether network traffic-derived features can capture daily behavioral patterns related to health-related activities and behaviors, including sleep–wake cycles, usage regularity, and activity rhythms.

Together, these aims position our approach as a scalable, privacy-preserving foundation for longitudinal research on digital behaviors with applications to improving understanding of mental health and cognitive conditions.

Methods

Study Tools

We conducted a two-week observational study with university students at New York University (NYU), collecting VPN-based mobile network traffic in real time to evaluate the feasibility and acceptability of continuous digital phenotyping. Passive data collection was carried out using the WireGuard VPN app [116] installed on participants' personal smartphones, complemented by self-reported measures, including the System Usability Scale (SUS) [77], NASA Task Load Index (NASA-TLX) [76], and semi-structured exit interviews. This design allowed us not only to assess the technical feasibility of continuous capture and participants' perceptions of usability, privacy, and burden, but also to explore whether network traffic-derived features can capture broader daily behavioral patterns. Further, by grounding the evaluation in a naturalistic setting, the study directly addresses our overarching goal of developing scalable, privacy-preserving methods for studying everyday digital behaviors relevant to health and well-being.

Participants & Recruitment

After obtaining IRB approval from NYU (IRB-FY2025-9589), participant recruitment was conducted through academic advisors across multiple NYU schools and departments (across STEM, arts, and social sciences departments), since the university does not maintain a centralized student mailing list. This decentralized strategy leveraged existing academic networks to reach a diverse student population relevant to our focus on young adult digital behaviors. Advisors distributed our invitation to their internal lists, describing a study of on screen-time behaviors and linking to a sign-up form. Outreach began in April 2025, generating an initial pool of interested students. In mid-May, these students received a consent packet outlining study procedures, privacy risks, and consent materials. Those who consented were onboarded via a secure web portal, where they installed the WireGuard app and completed setup. This cohort constituted the final study sample. Because this

was a feasibility pilot, we did not set a predetermined target sample size; instead, we adopted a best-effort recruitment approach. As illustrated in the CONSORT diagram (Figure 7), the final cohort reflects the number of students who could be recruited and onboarded within the study timeframe. Detailed demographic characteristics (age range, sex, major/department distribution) and inclusion criteria are reported in the Results – Participant Demographics section.

Study Instrument: WireGuard VPN for Passive Monitoring

A central challenge in mobile behavioral research is capturing objective measures of phone usage without modifying the operating system, draining the battery through continuous sensing, or requesting invasive permissions. However, existing approaches, such as rooting a device or deploying platform-specific monitoring apps, are unstable and challenging to scale. They often break with OS updates, raise privacy concerns, and significantly drain battery life, making them impractical for long-term deployments. To overcome these limitations, we sought a generic, noninvasive method that leverages a universal property of modern applications: most are internet-connected. Monitoring network traffic, therefore, provides a window into user activity that is agnostic to both device type and operating system, provided it can be collected seamlessly.

Virtual Private Networks (VPNs) [116] offer a lightweight mechanism to capture such traffic. By design, a VPN forwards all device-to-internet communication through an intermediary server. Without a VPN, the device exchanges encrypted packets (small units of transmitted data) directly with the destination server. With a VPN, these packets are first sent to the VPN server, which then relays them to the intended destination and returns responses to the device. While packet contents remain encrypted, the VPN operator can observe metadata, including contacted hostnames (the domain names of contacted services), packet counts, and byte volumes, in real time.

Among available VPN protocols (e.g., L2TP, IPsec, OpenVPN), we selected WireGuard for its simplicity, cross-platform support, and robustness in continuous operation. Setup requires only scanning a QR code, eliminating manual configuration and making onboarding nearly frictionless. To further reduce complexity and foster trust, we chose the official WireGuard client rather than building a custom app. The WireGuard app is widely available across iOS, Android, Windows, and macOS, and consistently ranks among the most reliable VPN solutions. In practice, participants downloaded the WireGuard app directly from the app store, scanned a personalized QR code generated by our NYU server, and their devices were automatically configured to route all traffic through our research VPN.

As shown in Figure 1, when active, the WireGuard client forwards all traffic through our Linux-based VPN server, where the open-source tool tshark [117] captures packets and aggregates them into summaries per 10 seconds. For each participant and timestamp, we retain the anonymous participant ID, the contacted hostname, the number of packets sent and received, and the bytes transmitted in each direction. A downstream data processing pipeline then maps hostnames to mobile applications (described later in this section).

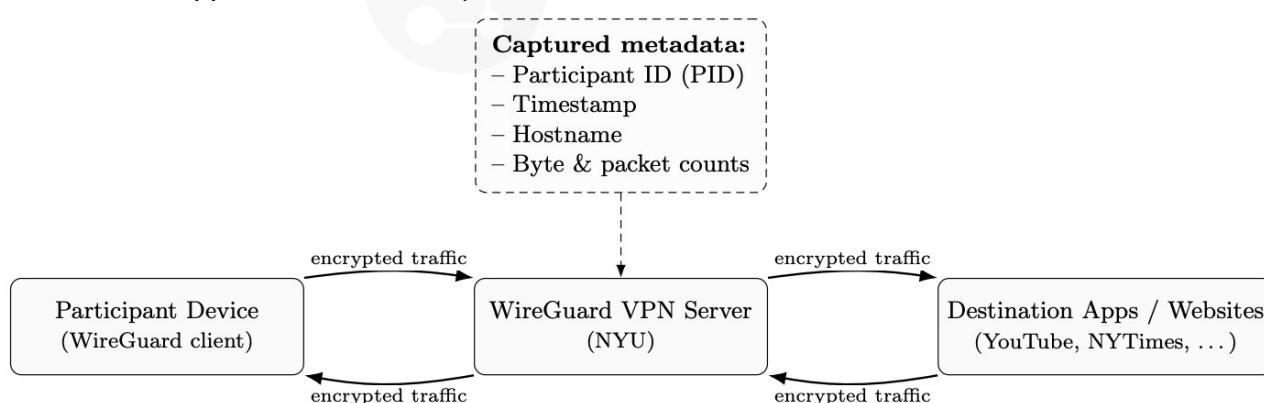


Figure 1. Architecture of the system. All device-to-internet traffic is routed through a WireGuard VPN server

hosted at NYU. While packet contents remain end-to-end encrypted, the server captures metadata (e.g., hostnames and traffic volumes), enabling passive behavioral analysis while preserving message and content privacy.

WireGuard demonstrated strong reliability in handling routine disruptions (e.g., Wi-Fi–cellular handoffs, airplane mode toggles) and minimized data gaps compared to other VPN clients. Energy efficiency was also critical: in 24-hour continuous-use tests on a lab iPhone (iOS 18), WireGuard consumed less than 1% of overall battery life, ensuring participants could keep the VPN active indefinitely without noticeable impact on usability.

Field Study

Onboarding Process

After enrollment, participants were directed to a secure online portal to complete setup. The portal allowed them to generate cryptographically random PIDs and, if desired, regenerate new ones until satisfied. Once a PID was chosen, participants were prompted to store it somewhere else securely, as it was required for re-entry to the portal and could not be recovered if lost. Participants were then instructed to download the WireGuard app (iOS or Android). The portal generated a unique QR code embedding the study-specific VPN configuration, which participants scanned within the app to automatically configure their device to route all internet traffic through our NYU-managed VPN server. The portal dashboard confirmed successful installation and guided participants to the next stage of the study, as shown in Figure 2.

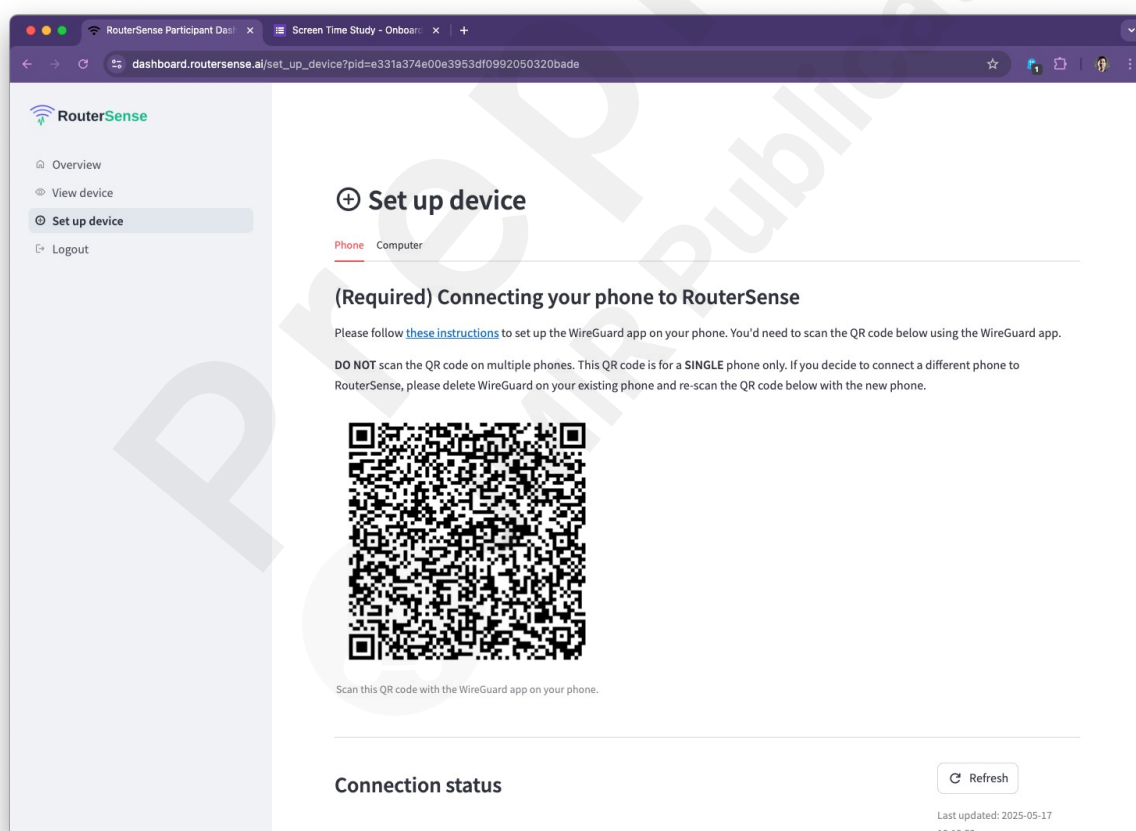


Figure 2. Participant onboarding portal. Each participant scanned a unique QR code to automatically configure the WireGuard VPN client and route traffic through the study server.

Data Collection

Once the WireGuard VPN is activated on a participant's device, it remains active by default and

continues to route all network traffic through the study server unless explicitly disabled or interrupted by a phone reboot. During active periods, the VPN server captures and aggregates metadata in fixed 10-second windows. Each record contains the participant ID (PID), timestamp, contacted hostname, and the number of packets and bytes transmitted in each direction. Packet contents are never collected. For example, when a participant streams YouTube, we can observe only that packets are exchanged with *youtube.com* and the corresponding traffic volume, but not which specific video was viewed, as that information is encrypted. The same applies to news sites, banking portals, or messaging apps: we capture only high-level traffic characteristics, never sensitive user content. This metadata is stored in a secure SQLite database on an NYU-managed server, which serves as the basis for subsequent inference. By limiting collection to traffic metadata (e.g., hostnames, timestamps, and volume measures), we preserve analytic utility for reconstructing temporal patterns of app activity while protecting privacy by excluding any content-level information.

Participants have the ability to manually disable the VPN within the WireGuard application. However, there are scenarios where the VPN may remain off unintentionally—for example, if the device reboots or if the participant forgets to re-enable it. The visibility of the VPN status also varies across devices: older iPhones display a persistent VPN icon in the status bar, newer iPhones with a notch, and most Android devices omit such an indicator. As a result, participants may be unaware when the VPN is inactive.

To mitigate this issue, the study protocol incorporates reminder emails sent every two to three days. These reminders direct all participants to the study portal, which allows them to verify their connection status and ensure that data are actively being collected. The portal provides real-time feedback on cumulative active study days and automatically calculates the corresponding compensation amount, as shown in Figure 3. This continuous feedback loop served both as a technical safeguard against accidental disengagement and as a behavioral nudge to maintain adherence.

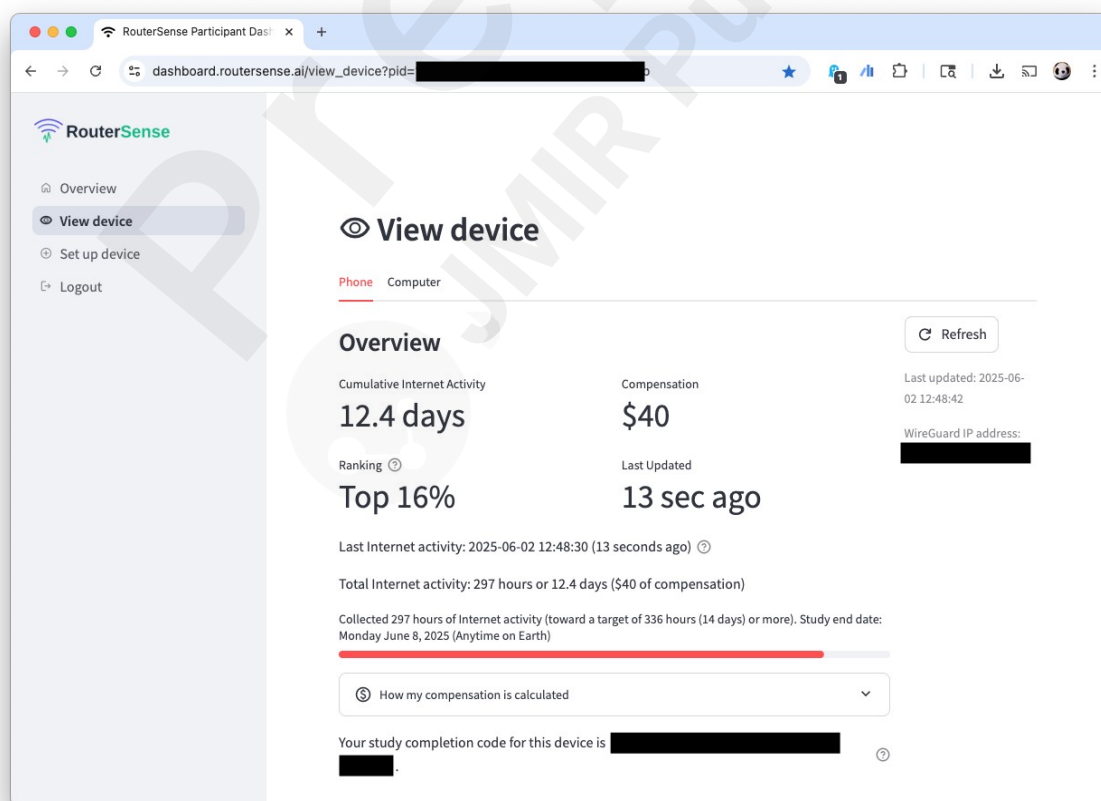


Figure 3. Participant-facing study portal. The dashboard displays cumulative internet activity, compensation amount, and completion code, providing real-time transparency into data collection and payments.

Compensation followed a tiered structure. Participants received \$0 for fewer than five cumulative days of VPN activity, \$38 for 5–7 days, \$39 for 7–10 days, \$40 for 10–14 days, and \$41 for more than 14 days. This structure was deliberately designed to encourage sustained participation beyond the initial five-day warm-up period while minimizing the influence of financial incentives on longer-term engagement. The incremental increases in payment were deliberately small to limit the influence of financial incentives on participation, since the marginal benefit of continued engagement diminished over time. Rolling recruitment meant that the overall study period—from the first participant onboarded to the last participant exited—spanned approximately 20 days.

Beyond payment tracking, the portal also provided participants with a transparency dashboard. This dashboard included two complementary visualizations. The first is a real-time table that displays the hostnames and applications currently observed (Figure 4), providing immediate visibility into the kinds of data being collected. The second is a historical heat map that represents participant activity patterns across hours of the day and days of the week (Figure 5). Together, these views allow participants to contextualize their own digital behaviors while simultaneously reinforcing transparency regarding the scope of data collection. Participants have continuous access to the portal for the duration of the study and may remain enrolled until the final day of data collection, after which payments cease and VPN monitoring is terminated.

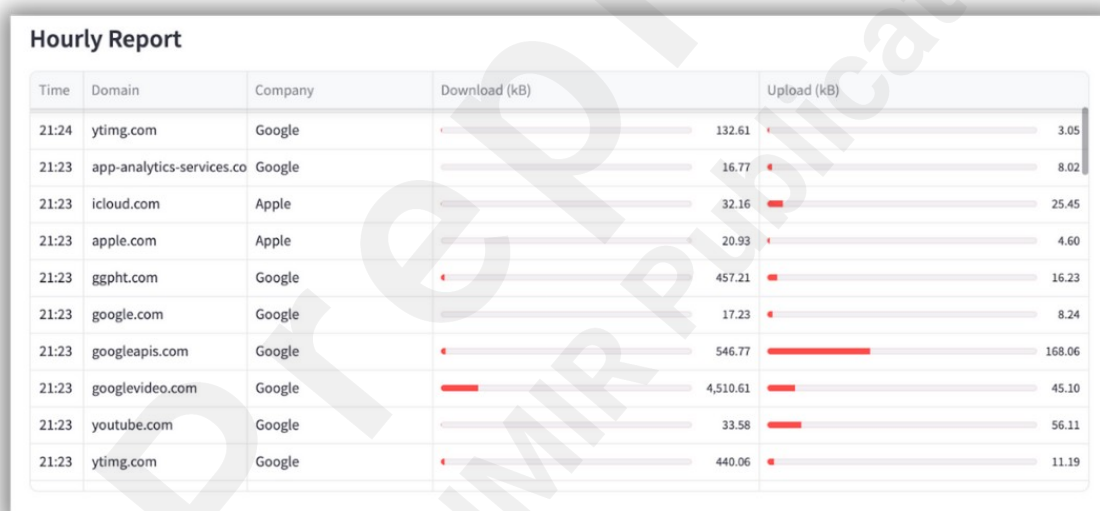


Figure 4. Real-time table on the participant dashboard displaying hostnames contacted during app use (shown here for a YouTube session: youtube.com, googlevideo.com, yting.com).

Internet Activity Heatmap (10-minute intervals) - (Show all companies)

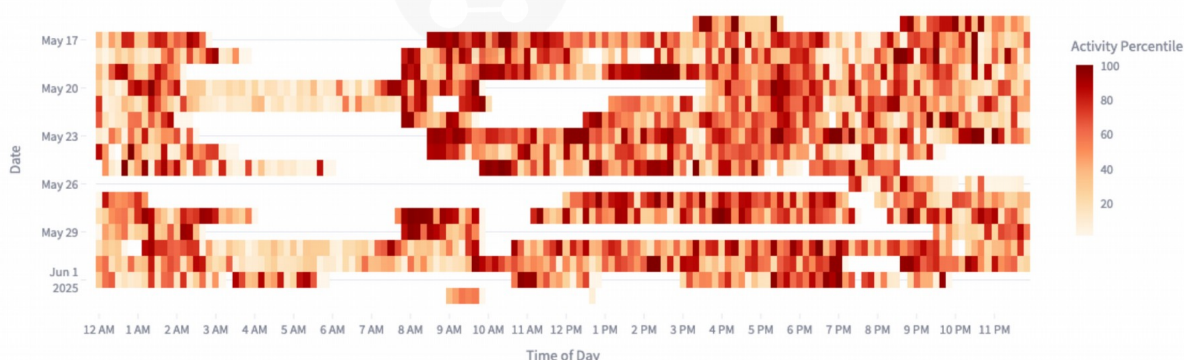


Figure 5. Internet activity heatmap aggregated in 10-minute intervals over 24 hours. Rows represent study days and columns represent time of day; color intensity indicates the percentile of observed activity.

Privacy model

To prevent linkage between digital behaviors and real-world identities, each participant created a random participant ID (PID) during onboarding; no mapping between PIDs and personal information was stored, and participants could regenerate a new PID at any time. The VPN server logs *metadata only*—PID, timestamp, destination hostname, and byte/packet counts; packet contents are never collected. Data is stored on NYU-managed infrastructure with encryption and restricted access to authorized study personnel. Before enrollment, participants received a plain-language briefing (slides, video, live demonstration) describing the privacy model and risks, and provided informed consent.

Exit Surveys and Interviews

At the conclusion of the study, researchers invited all participants via email to complete a short exit survey and schedule a follow-up interview. The survey collects free-text entries for the top ten apps used during the study period, along with standardized usability assessments: the NASA Task Load Index (NASA-TLX) and the System Usability Scale (SUS).

In addition to the survey, participants were invited to a follow-up interview conducted via Zoom. To schedule interviews and issue payments, researchers needed access to participants' real-world identifiers (e.g., names, email addresses). Consistent with the study's privacy model, compensation was decoupled from collected metadata: the study portal generated a cryptographically signed completion code for each participant, encoding the payment amount without linking it to their PID. Participants presented this code during the interview, and researchers validated its authenticity to confirm compensation while preserving study anonymity. Interviews were transcribed using Otter AI [128], and researchers also took contemporaneous notes to capture qualitative observations.

During the semi-structured interview, researchers asked participants open-ended questions about their experience, with prompts designed to elicit reflections on VPN usage, dashboard engagement, and perceived study burden. Questions included whether the VPN was kept active throughout, if and when it was turned off, and how frequently the portal was consulted. Participants are also asked whether they would have been willing to extend participation beyond 14 days, since earlier enrollees had that option. The interview format prioritizes participant-led discussion, allowing concerns or observations to surface organically. A complete set of interview questions is provided in the Appendix.

As an optional step, participants using iPhones could choose to share their App Privacy Report, an Apple-generated summary of app activity over the preceding seven days (further described in the Data Analysis section below). Before sharing, participants were briefed on the scope and limitations of this report: it reveals which apps were active at given times but not the specific content accessed (e.g., the YouTube app used, but not which video was viewed). Sharing the report was voluntary and served only to provide a complementary ground truth for app usage, augmenting the network metadata collected through the VPN.

Data Analysis

To evaluate feasibility, acceptability, and the potential of passive traffic data to everyday activity patterns, we combined qualitative insights from participant interviews with quantitative analysis of network traffic metadata.

Qualitative Analysis: Interviews

Two authors independently reviewed interview transcripts and conducted a thematic analysis [72–74], generating an initial set of 127 codes. Through iterative axial coding, the code set was refined by merging overlapping categories, revising definitions, and removing redundancies. Consensus-based synthesis identified themes relevant to feasibility, acceptability, and participant trust.

Quantitative Analysis: Feasibility

We evaluated feasibility across two domains: user retention and data coverage, following conventions in digital phenotyping and mobile sensing research.

User Retention

User retention was defined as the proportion of participants who contributed at least five days of traffic out of all participants who contributed any traffic. This threshold ensured that estimates were based on individuals with a meaningful amount of monitoring.

Data Coverage

Data coverage was defined as the proportion of hours within each participant's monitoring window (from their first observed active hour to their last) that contained valid VPN traffic. An hour was considered "active" if at least one byte of data was observed. For example, if a participant's monitoring window spanned 120 hours but traffic was observed in only 100 of those hours, their coverage rate was 83%. This definition captures transient lapses (e.g., when the phone was rebooted or the VPN was manually disabled) while aligning with conventions in digital phenotyping, where adherence is measured as the percentage of expected intervals containing data [131–133]. To further contextualize coverage, bins were stratified by (a) time of day (morning: 06:00–11:59; afternoon: 12:00–17:59; evening: 18:00–23:59; night: 00:00–05:59 Eastern Time) and (b) calendar cycle (weekday vs. weekend). Participant-level coverage was summarized using mean and median coverage across individuals, while aggregate feasibility pooled bins across participants to yield overall coverage rates within each stratum (Table 2).

Exploratory Analysis: Network Traffic

Inferring Behavioral Rhythms

To extract behavioral structure from encrypted network flows (i.e., sequences of packets sharing a common source, destination, and time window), we treated traffic volume as a proxy for smartphone activity. Internet communication involves both downloads, which largely reflect content delivery (e.g., high-volume video streams versus smaller text loads) and uploads, which more directly capture user-initiated actions such as taps, clicks, or navigation. Although exceptions exist—such as high upload traffic during calls or background updates—the relative balance between uploads and downloads provides a useful signal of interaction intensity. Because phones are never completely silent, with low-level background exchanges for notifications and maintenance, we interpret higher-volume deviations above this baseline as more indicative of active user engagement.

In particular, we focus on upload traffic as the most reliable proxy for user-initiated interaction, since many common actions (e.g., sending messages, posting, or searching) generate outbound requests. Raw byte counts were aggregated into contiguous, non-overlapping 10-minute bins, a resolution that suppresses transient bursts while preserving diurnal dynamics. For each participant, this yielded a 24-hour activity vector per day. To account for differences in overall smartphone usage—arising from individual habits, app patterns, and device capabilities—we applied a rank-based, within-day percentile normalization. Specifically, for each day d , the traffic volume in bin t was replaced by its empirical percentile relative to the distribution $\{x_{d,1}, \dots, x_{d,T}\}$, where T denotes the number of bins in that day. This transformation is invariant to monotonic scaling and preserves intra-day temporal structure, enabling direct cross-participant comparison of diurnal phase and amplitude, while suppressing between-subject variance in absolute throughput.

We then visualized and analyzed normalized time series in two complementary ways. First, we projected daily activity onto a radial heatmap (Figure 10), where angular position denotes clock time [0–24h), concentric rings denote successive days, and color denotes normalized traffic intensity (red = high percentile, blue = low percentile). This polar representation highlights rest–activity cycles,

phase shifts, and fragmentation of daily usage patterns. Nighttime intervals (20:00–08:00) were shaded, and weekends were highlighted in green to facilitate alignment with social schedules.

To isolate context-specific rhythms, each network flow was enriched by mapping hostnames to high-level application categories (e.g., Social, Entertainment, Games, Business, Education, Food & Drink). Mappings were derived from Apple's App Privacy Report (APR; see subsequent section), supplemented by manual verification and aggregation to category-level taxonomies. For each participant, we generated a set of paired radial visualizations: one depicting aggregate device activity and additional plots for each activity category of interest (Figure 11). This design enabled structured comparison between global digital behavioral rhythms and the distinct temporal patterns associated with multiple behavioral categories.

The result is a device-agnostic framework that combines upload-sensitive traffic features, percentile normalization, radial time encoding, and app-level categorization to detect time-of-day and behavioral regularities in digital behavior while preserving privacy. While some background traffic may still be misclassified as user activity, and the precise boundary between passive and active phone use remains unresolved, we could still infer broad patterns of usage. More specifically, the approach identified meaningful diurnal and behavioral regularities inferred from encrypted network traffic alone, offering a scalable path for digital phenotyping without OS-level instrumentation.

Inferring Applications from Hostnames

Our network traffic data consisted of full hostnames (e.g., *video.cdn.youtube.com*). For analysis, these were collapsed into base domains (e.g., *youtube.com*), which served as the unit of application mapping. Domains offered a tractable approximation of application usage for category-level analyses; however, raw hostnames rarely map one-to-one onto the specific applications participants were actively using. In some cases, the mapping is transparent—e.g., requests to *youtube.com* almost certainly indicate YouTube activity. In other cases, however, the mapping is ambiguous. For example, content delivery networks (CDNs), such as *cloudflare.com* or *akamai.net*, serve as infrastructure for dozens of unrelated apps, making it impossible to infer app-level usage from hostnames alone. This ambiguity necessitates additional methods for disambiguation.

To address this issue, we leveraged Apple's *App Privacy Report (APR)*, an opt-in iOS feature (described earlier in the Exit Surveys and Interviews section) that logs, at the app level, the set of hostnames contacted over the preceding seven days. The APR is stored locally on the phone as a JSON file and provides a direct mapping between hostnames and the apps that generate them. When a hostname is uniquely associated with a single app, we assign it accordingly. When hostnames are shared across multiple apps, we adopt a conservative approach and exclude them from mapping to avoid misclassification. In our dataset, this affected approximately 11% of observed hostnames, while the remaining 89% exhibited a 1-to-1 mapping to a single app. Despite the seven-day reporting window, the APRs proved informative, particularly when participants' usage patterns were consistent across that period.

To provide higher-level behavioral groupings, apps were mapped to functional categories using Apple's iTunes Search API [148]. Given an app's bundle ID (a unique identifier assigned to each app) obtained from the APR, the API returns its Primary Genre Name (e.g., Social, Entertainment, Productivity). This allowed us to systematically group apps into higher-level behavioral categories, complementing the app-level mappings derived from APR, as shown in Figure 6.

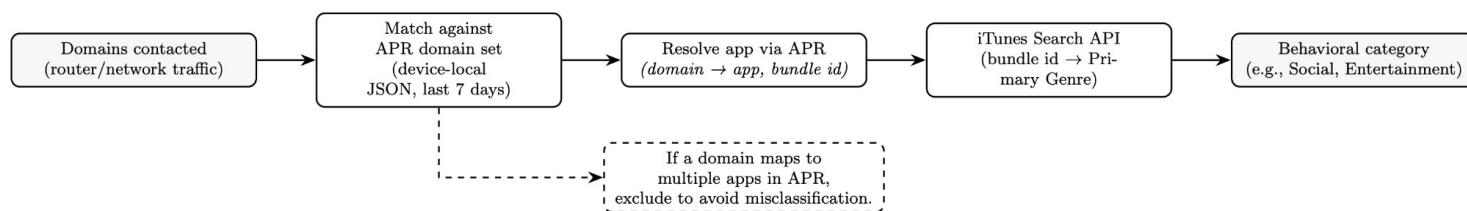


Figure 6. Hostname → App → Category pipeline. Domains observed in network traffic are matched against Apple's App Privacy Report (APR) to recover the generating app (via bundle ID), which is then linked to its primary genre using the iTunes Search API. This process yields higher-level behavioral categories that complement app-level mappings.

For participants who chose not to share their reports, we generated our own mappings using a controlled device: an experimental iPhone connected to the WireGuard VPN. On this device, we manually interacted with a curated set of apps to capture their network traffic. The selection of apps came from two sources:

- (1) Participant-reported apps listed in exit surveys and interviews as most frequently used.
- (2) Top-ranked apps from the Apple App Store across multiple app categories.

Each app was interacted with for approximately one minute, restricted to publicly accessible features (no accounts were created for apps requiring logins). The resulting network traffic traces were cross-referenced with APR logs to isolate hostnames uniquely attributable to each app.

In total, we tested 272 apps from the App Store and 55 apps reported by participants. Together with the participant-contributed APRs, this methodology yielded a well-grounded mapping from hostnames to applications, which we then enriched with category labels retrieved from the iTunes Search API [148].

Results

Outcome 1: Feasibility

User Retention

User retention was quantified as progression along the study pipeline (Figure 7). Specifically, we tracked the proportion of students advancing from initial eligibility screening through consent, onboarding, VPN activation, and study completion. This mirrors conventions in digital health feasibility studies, where recruitment yield and retention are reported alongside data completeness [141,142].

Recruitment outreach reached 100 students, of whom 38 provided informed consent. Of these, 32 completed the onboarding survey, 29 provided valid network data, and 27 remained active for more than five days. At study completion, 26 participants completed the exit survey, and 25 completed the exit interview.

Overall, this corresponds to 93% retention to the analysis stage (27/29) and 66% retention from consent through exit (25/38). Comparisons between students who consented but did not continue and those who completed the study revealed no systematic differences in department, major, or university role, suggesting minimal self-selection bias that might confound usage patterns.

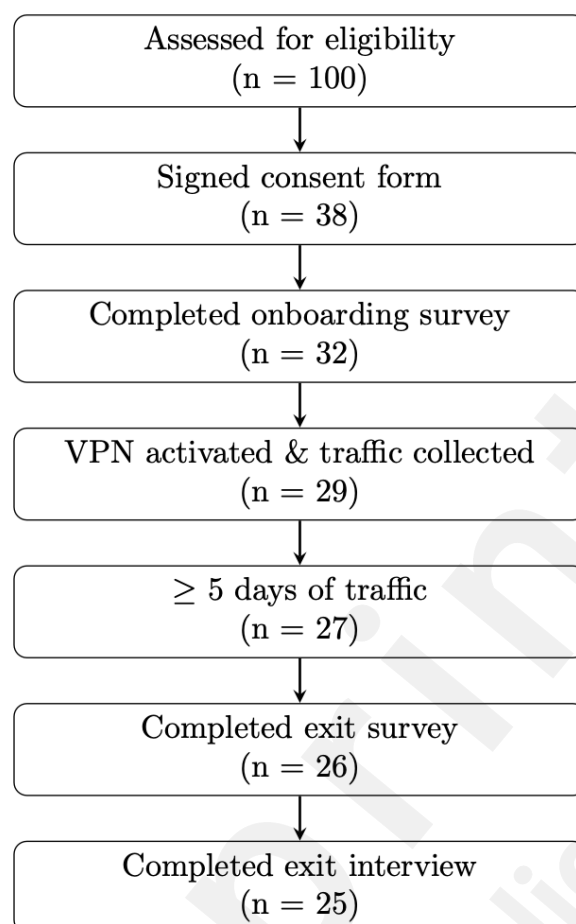


Figure 7. CONSORT-style participant flow diagram, showing progression from initial recruitment through study completion.

This level of retention compares favorably with prior digital health feasibility studies. For example, one large cross-study evaluation found a median retention time of only ~5.5 days across eight remote digital health studies, with participants often disengaging within the first week [143]. A review of 37 studies reported a median completion rate of ~48% (IQR 35–76%), indicating that about half of participants typically finish [144]. More recent cohorts have achieved higher rates under intensive clinical contexts (e.g., ~66% retention at 8 weeks in a mood disorder study [145]; ~73% in a smartphone-based depression biomarker pilot [146]). By comparison, our retention of 93% indicates substantially stronger participant engagement than is commonly observed in remote digital phenotyping studies.

Demographic characteristics of the final sample are summarized in Table 1. Exit interviews were conducted with 25 participants, who reflected on their experiences using our system, reported perceived benefits and burdens, and provided usability feedback. The sample comprised 12 women, 12 men, and one non-binary participant. The mean age was 22 years (SD = 2.9; range 18–29). Educational levels spanned undergraduates (n = 20), one master's student, and four doctoral students, representing a variety of majors. The most common fields were Chemistry (6), Economics (5), Computer Engineering/Science (5), Data Science (3), and Mathematics (3). Racial and ethnic composition included Asian (12), White/European (5), Latin/Hispanic (3), Black/African/Caribbean (1), with four participants choosing not to report.

Demographic Variable	Category/Description	Count	%
Age	Mean (SD)	22.04 (2.95)	
	Range	18-29	

Gender	Male	12	48.0%
	Female	12	48.0%
	Non-binary	1	4.0%
Race/Ethnicity	Asian	13	52.0%
	White/European	6	24.0%
	Latin/Hispanic	2	8.0%
	Black	2	8.0%
	Hispanic	1	4.0%
	Black/African/Caribbean	1	4.0%
Degree Level	Undergraduate	20	80.0%
	Master's	1	4.0%
	Ph.D.	4	16.0%
Major	Chemistry	6	24.0%
	Computer Engineering	3	12.0%
	Economics	3	12.0%
	Biochemistry	2	8.0%
	Data Science	2	8.0%
	Others (9 majors)	9	36.0%

Table 1. Demographic characteristics of interview participants (N = 25), summarizing age, gender, race/ethnicity, degree level, and field of study.

Data Coverage

Feasibility was also evaluated in terms of data coverage, defined as the proportion of expected monitoring intervals in which valid VPN traffic was captured. Participants were onboarded on a rolling basis while the study concluded at a fixed calendar date. As a result, not all individuals had the opportunity to contribute a full 14 days of monitoring. Feasibility was therefore defined relative to each participant's observed monitoring window (first to last active day), rather than by imposing an absolute duration requirement. This definition aligns with conventions in digital phenotyping research, where adherence is typically measured relative to each participant's monitoring opportunity [79, 80].

To ensure comparability of temporal rhythms, we restricted coverage analyses to the 24 participants who remained in a consistent time zone (i.e., not traveling during the study period). Across these participants, mean feasibility was 74.1% (median 77.1%), with coverage ranging from 21% to 100%. This range reflects heterogeneous adherence: some students maintained near-continuous VPN connectivity, while others exhibited frequent gaps, typically due to forgetting to re-enable the VPN. On average, participants contributed 311.6 hours of monitored traffic (~13 days, SD = 3.5), with individual totals spanning 121 to 496 hours. In addition to overall coverage, we examined the temporal distribution of VPN off periods. Across all 24 participants, the median gap in coverage was

7.5 hours (IQR: 2.8–23.8 hours), with most interruptions being relatively short (<24 hours). However, occasional extended lapses occurred, with the maximum gap reaching 304 hours (~12.5 days). This extreme case (Participant P05) showed nearly 13 days of continuous activity followed by ~12.5 days of inactivity, before briefly re-enabling the VPN at the study's end.

Scope	Mean (%)	Median (%)
Overall	74.1	77.1
Morning	71.1	75.0
Afternoon	75.9	79.0
Evening	77.6	82.8
Night	71.9	75.5
Weekday	74.2	79.5
Weekend	74.0	76.4

Table 2. Participant-level data coverage, summarized as mean and median coverage across individuals. Coverage values are reported for overall monitoring, time-of-day strata, and weekday/weekend cycles.

Stratification confirmed that coverage was consistent across daily rhythms and calendar cycles (Table 2). Participant-level mean feasibility clustered tightly between 70–78% across mornings (71.1%), afternoons (75.9%), evenings (77.6%), and nights (71.9%), with weekday (74.2%) and weekend (74.0%) rates nearly identical. Aggregate analysis pooling bins across all participants yielded similar estimates (69.7–76.2% across strata), underscoring that passive VPN sensing did not systematically fail during any particular life-activity period (see Appendix Table 1 for aggregate-level results). On average, participants contributed 16.2 valid days (range 10–23), reflecting staggered onboarding relative to the fixed study endpoint.

Figures 8 and 9 provide complementary perspectives on continuity. The calendar-aligned view (Fig. 8) captures deployment feasibility by showing how staggered onboarding produced overlapping coverage across the fixed study calendar. By contrast, the Day-0-aligned view (Fig. 9) normalizes each participant's trajectory to their own start date, highlighting adherence patterns on equal footing and revealing heterogeneity in continuity and lapses.

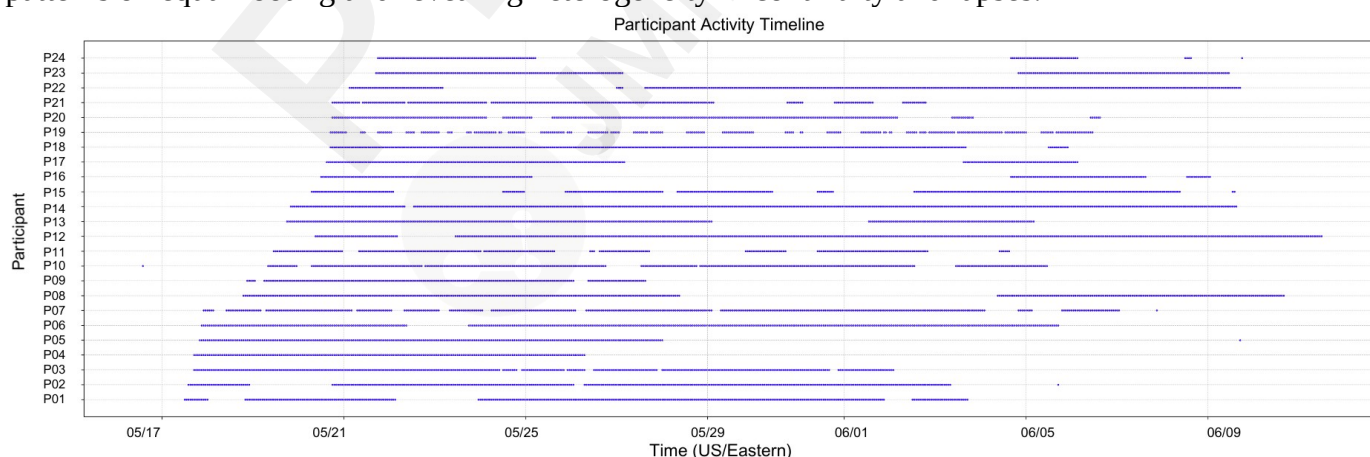


Figure 8. Calendar-aligned VPN continuity timeline. Each dot represents a 60-minute bin with VPN traffic, ordered by enrollment date. This view emphasizes deployment feasibility.

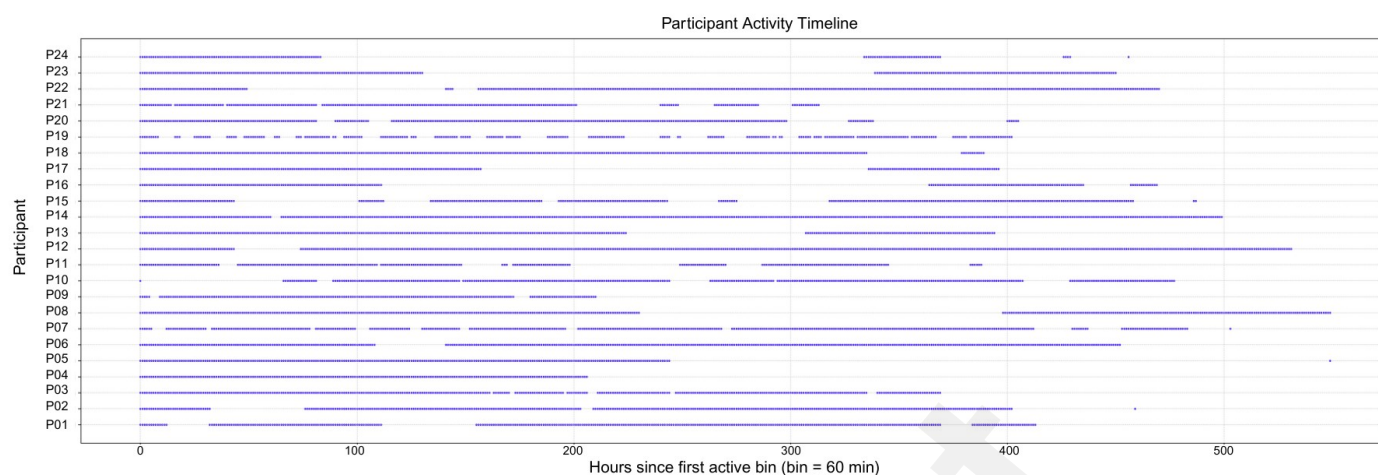


Figure 9. Day-0-aligned VPN continuity timeline. Each dot represents a 60-minute bin with VPN traffic, normalized to participants' first active bin. This view emphasizes individual adherence trajectories.

Taken together, these results indicate that observed feasibility primarily reflects participant adherence rather than limitations of the sensing system. Importantly, the ~74% mean coverage we achieved exceeds benchmarks commonly reported in prior digital sensing studies. For example, early deployments of the Beiwe smartphone app in schizophrenia cohorts reported only ~47–50% mean passive data coverage [81,82], and multi-cohort smartphone studies have found that fewer than half of participants achieved >50% GPS or accelerometer coverage [83]. Similarly, in a large-scale deployment of the EARS app involving 4,754 adolescents, fewer than half (1,463) contributed analyzable data [99], underscoring how adherence challenges can substantially constrain data yield.

Participants' own accounts (obtained via the exit interviews) echoed this impression of feasibility: twenty-one of twenty-five interviewees described the system as easy to set up and effortless to maintain, often noting that onboarding was quicker and smoother than prior experiences with VPN tools. Once activated, the system ran unobtrusively in the background with little or no intervention. Twelve participants produced more than 14 days of valid recordings, and three continued beyond the study window, explaining that they had forgotten it was still running—a strong marker of low-maintenance operation.

In addition, 18 participants emphasized the unobtrusiveness of the system, noting that the VPN ran quietly in the background without disrupting phone use or daily routines. Although the companion app displayed usage and account information every few days, this was not perceived as burdensome. Rather, participants highlighted that the VPN itself imposed no noticeable friction and did not interfere with work, study, or personal life. Six participants even reported that they “forgot it was running,” despite the study being scheduled to end after 14 days. As one participant put it: *“It didn’t interfere at all with anything. Once I set it up, I completely forgot about it.”* [P6]

To better understand the gaps in data coverage, we examined participants' qualitative feedback. The same property that made the VPN unobtrusive also contributed to occasional gaps: when devices powered off or restarted, participants often forgot to manually re-enable it. As P6 recalled, *“Yes, there was one point where my phone died, and it shut off. I think it took about a day for me to realize it wasn’t collecting. But the rest of the time it was consistently on.”* Another participant elaborated that these lapses were not intentional disengagement, but rather a technical limitation: *“I didn’t know that the VPN turned off if my phone died the first time. I think I lost maybe half a day of traffic then, and the second time it died I lost about an hour. It was just forgetting that the VPN doesn’t turn on automatically. Overall it was continuous, except for when my phone died, and those lapses weren’t that long.”* [P7].

A small minority of participants ($n = 4$) reported perceived device performance effects, such as

increased battery drain or occasional network slowdowns during continuous VPN operation. Lab measurements, however, showed that battery consumption remained below 1% over a 24-hour period, suggesting these concerns may be more perceived than actual.

Further themes that emerged from qualitative data analysis are summarized below.

^aParticipant Perspectives on Seamless Operation:

“Just because it’s super simple. It doesn’t bother me. And I’m kind of interested in seeing what the results are.” [P9]

“You know it was all very, very clearly organized so that made it effortless to set up. I think it was good.” [P6]

“I just need to click on one button and turn on the VPN. That’s it. It’s not challenging to set up or use.” [P23]

Participant Perspectives on Unobtrusive Integration:

“Yeah, I would have continued because it was pretty easy. It just ran in the background. I didn’t really have to think about it.” [P16]

“I quite liked it. It was very easy to use—pretty invisible. Most of the time I wasn’t even aware it was happening.” [P22]

^aAll participant quotations were lightly edited for clarity, concision, and privacy, while preserving their original meaning.

Outcome 2: Acceptability

To evaluate the acceptability of our system in naturalistic daily use, we administered adapted versions of the NASA Task Load Index (NASA-TLX)[76] and the System Usability Scale (SUS) [77] during the exit survey. These instruments quantify, respectively, the perceived workload associated with system use and the overall usability of the interface and interaction design.

As shown in Figure 10, NASA-TLX ratings indicated consistently low levels of burden across all six dimensions: mental demand ($M = 2.69$, $SD = 3.27$), physical demand ($M = 1.54$, $SD = 1.79$), temporal demand ($M = 2.92$, $SD = 4.09$), performance ($M = 4.08$, $SD = 4.18$), effort ($M = 3.85$, $SD = 4.31$), and frustration ($M = 2.88$, $SD = 3.78$). The boxplots display the distribution of responses, with medians (horizontal lines) often below the mean values (black dots), reflecting skewness due to a small number of outliers (open circles). Although the NASA-TLX does not prescribe thresholds for “acceptable” workload, the uniformly low scores suggest that participants experienced minimal cognitive, physical, or temporal strain.

For context, prior evaluations of digital health technologies typically report higher NASA-TLX scores. For example, in a study of a COVID-19 exposure notification app, mean mental demand ratings averaged around 29/100, with overall workload ratings across subscales falling in the 25–35/100 range [139]. In contrast, our participants’ mean scores correspond to roughly 10–20/100 equivalents, underscoring that the VPN-based system operated unobtrusively in the background and imposed negligible demands on daily routines.

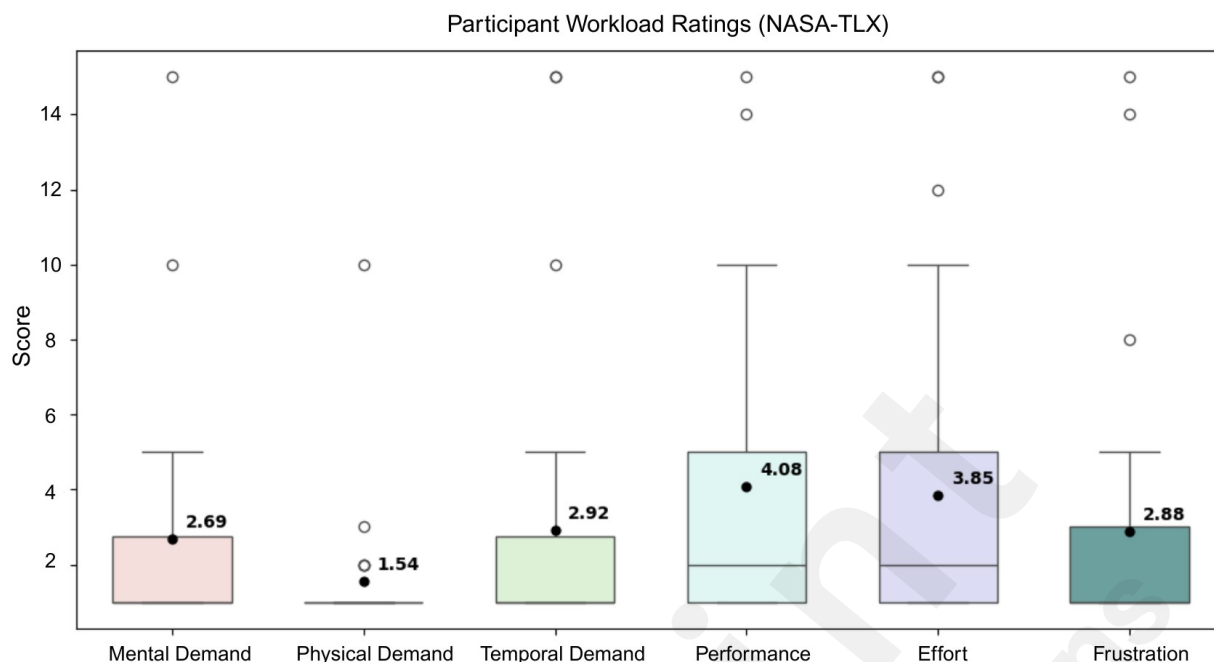


Figure 10. Box plot of NASA-TLX scores from the exit survey (n=26), illustrating participants' perceived workload associated with maintaining the VPN-based monitoring system in daily life. Scores are reported on a 1–21 scale (higher = greater workload).

Additionally, SUS results reinforced this conclusion. The mean SUS score was 78.08 (SD = 14.96), ranging from 52.5 to 100. According to established SUS benchmarks [78], scores above 68 are interpreted as reflecting good usability, positioning our passive monitoring system well above the benchmark for acceptability.

Beyond workload and usability, privacy and trust emerged as critical dimensions of acceptability. Fifteen participants expressed strong confidence in the research team and institutional affiliation, emphasizing that organizational credibility made them comfortable permitting passive data monitoring. At the same time, twelve participants raised concerns about how their data might be used, particularly around anonymity and the risks of being tracked. A small subset of participants (n = 3) continued to experience discomfort during sensitive activities (e.g., banking, video streaming). Seven participants reported occasionally disabling the VPN, citing either privacy concerns or usability issues. Their comments reflected both reassurance and caution: *"You explained it pretty well—what you were doing, how you were capturing the data, how you were making sure it was private and not attached to my name. So I felt like I was going into this with a relative sense of security."* [P14] In contrast, another participant noted: *"Sometimes I turn it off when I'm using my banking apps."* [P20]

Interviews further indicated that acceptability extended well beyond the 14-day study horizon. Nineteen participants expressed willingness to continue using the VPN-based passive monitoring system for longer periods, describing it as unobtrusive, easy to maintain, and, for some, even motivating in that it provided an opportunity to reflect on their digital behaviors. Four participants took the stance of being neutral but permissive: *"no reason to stop using it,"* reflecting that the system imposed little burden. Others stated they would gladly extend participation if specific reliability concerns were resolved. In particular, three participants described instances of unexpected VPN shutdowns (e.g., when the device powered off) that produced gaps in data collection and temporarily interrupted the study. Although relatively infrequent, they directly explain the feasibility gaps observed in our quantitative analysis: participants typically forgot about the VPN once installed, so reactivation after a restart depended on manual action. Together, these accounts indicate that coverage gaps stemmed primarily from benign forgetfulness and limitations of the WireGuard VPN app, rather than participant burden or resistance to monitoring. Participants further suggested that

simple feedback mechanisms could help sustain trust and continuity.

At the same time, participants consistently underscored the system's minimal burden and their willingness to continue. As two participants described:

"Yes, I would continue to use it. Just because it's super simple. It doesn't bother me. And I'm kind of interested in seeing what the results are." [P9]

"It was very little effort. Pretty invisible. And also, I'm happy to participate in research—it feels like a good thing to do if it doesn't cost you anything." [P22]

Some participants also highlighted tangible benefits from using the system, which further supported its acceptability. Twelve participants described becoming more aware of their smartphone usage patterns, with a few ($n = 5$) noting that this awareness reduced their desire to engage with their phones. The dashboard, in particular, was valued as a reflective tool, helping participants visualize sleep–wake cycles, daily routines, and the extent of app or domain usage in a way they had not previously considered. These insights motivated changes in behavior, such as reducing social media use or establishing a regular bedtime routine.

As one participant explained:

"Yeah, I can see when I go to bed and when I wake up. The heat map makes it obvious—those long blank stretches are when I'm sleeping." [P23]

Another emphasized how the data visualization challenged habitual phone use:

"It's really useful. A lot of apps and social media are designed to keep you hooked, and people often don't realize how much time they're spending. This made me notice just how quickly 40 minutes can disappear." [P2]

A third described how our passive VPN monitoring system spurred intentional behavior change:

"Once I saw it was tracking my activity, I actually took a break from Instagram for a week just to see the difference. It made me more conscious of my usage. And since it's also just a VPN, I figured I might as well keep it running." [P13]

Outcome 3 (Exploratory): Capturing Behavioral Rhythms from Network-Derived Features

We examined data from two exemplar participants to identify patterns of daily activity (Figure 10). Participant A exhibited a well-structured and consistent profile across all categories, with internet usage concentrated during daytime hours and near-complete inactivity overnight. The radial plots showed a strong day–night rhythm, and social media activity closely mirrored overall traffic, suggesting that it was an integrated, routine component of daily phone use. Beyond this temporal regularity, category-specific traffic revealed distinct behavioral preferences: consistent engagement with gaming applications, substantial activity in business-related domains (e.g., productivity, professional networking), and minimal interaction with food delivery services. Taken together, these patterns portray an individual with stable routines, a recreational gaming orientation, and professional engagement.

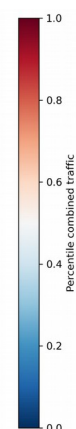
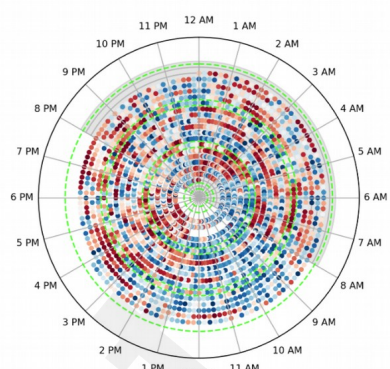
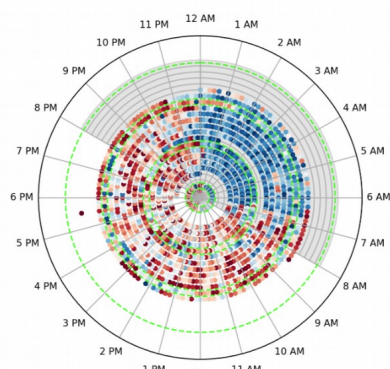
In contrast, Participant B displayed a more fragmented and irregular activity profile. While overall traffic still showed some day–night rhythm, social media engagement was less predictable, with bursts appearing at atypical hours such as late night and early morning. Category-level traces pointed to minimal use of gaming or business-related applications but substantially greater reliance on food-related services, suggesting frequent browsing or ordering through delivery platforms. These patterns together reflect a lifestyle characterized by irregular phone use, lower engagement with recreational or professional apps, and higher dependence on service-oriented applications.

App Category

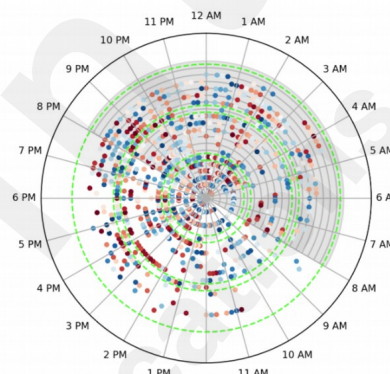
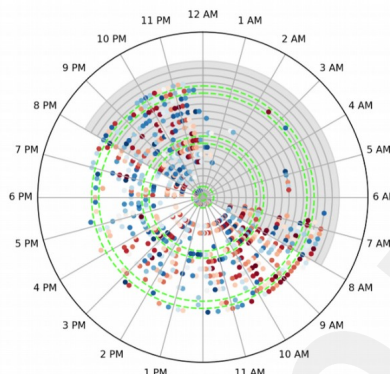
Participant A

Participant B

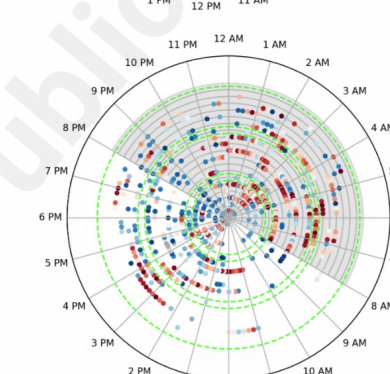
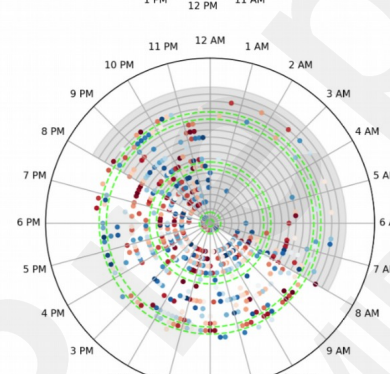
Overall Traffic



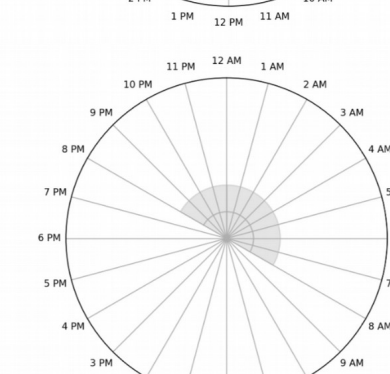
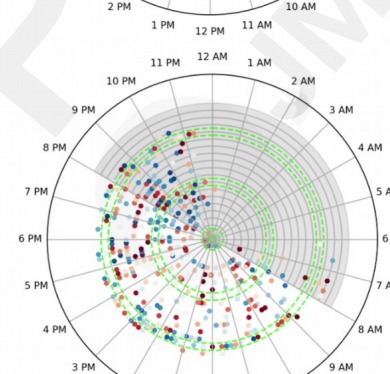
Social



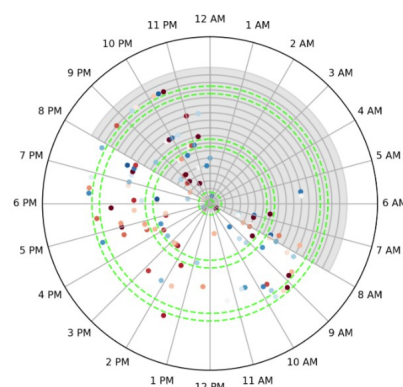
Entertainment



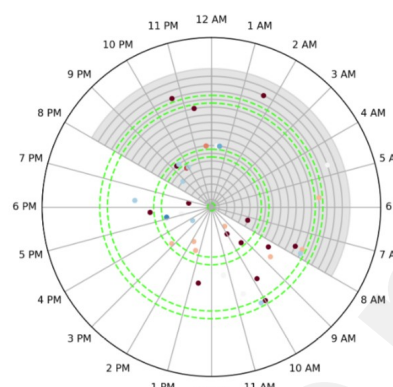
Games



Business



Education



Food

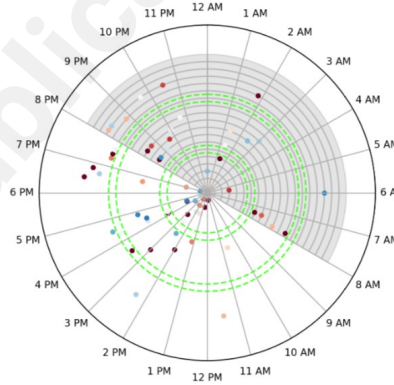
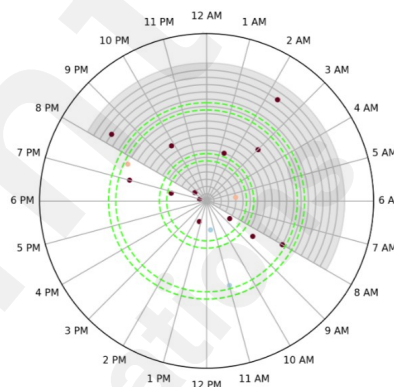
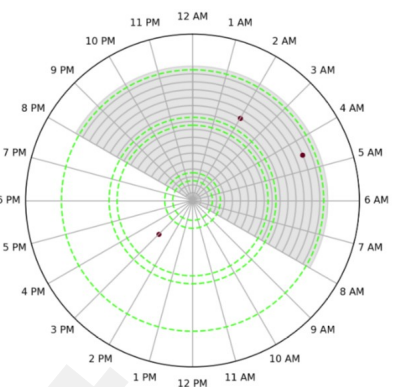
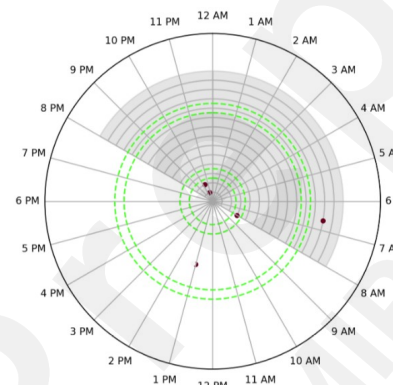


Figure 11. Temporal dynamics of overall and category-specific upload traffic for two representative participants. Each radial heatmap shows 24 hours (angular axis) across successive days (concentric rings). Color indicates normalized traffic intensity, with red denoting higher device usage and blue denoting lower usage. Nighttime hours (20:00–08:00) and weekends (green rings) are marked to facilitate alignment with social schedules.

These contrasts illustrate that encrypted network traffic, when decomposed by app category, can yield two complementary insights: broad daily activity rhythms and finer-grained behavioral signatures. By linking the timing of activity with the intensity of specific app categories, it becomes possible to infer not just *when* someone is online, but also *how* they use different types of applications—for example, whether they rely heavily on social media for communication, gaming for recreation, productivity tools for work, or delivery platforms for on-demand services. Such patterns capture lifestyle markers without requiring access to specific content data.

Although these radial visualizations highlight clear rhythms, it is important to note that the mapping of percentile traffic levels to “active” versus “idle” use is heuristic. High percentile bins likely capture periods of foreground, user-driven activity, while low percentiles often align with background processes or idle states. As these thresholds are derived relative to each participant’s own traffic distribution, the radial plots are best interpreted as exploratory markers of behavioral rhythms

and lifestyle patterns, offering insight into temporal structure rather than absolute measures of device engagement.

Discussion

Principal Results

This study shows that encrypted smartphone network traffic can be transformed into stable, temporally resolved behavioral signatures without inspecting content or requiring intrusive sensing. Our VPN-based monitoring system sustained continuous capture over nearly two weeks on average, with minimal participant effort, confirming feasibility. Participants rated it as lightweight and unobtrusive, with SUS scores well above standard usability benchmarks, and interviews highlighted its “invisible” operation in daily life. Acceptability was shaped by privacy and trust. Most participants expressed confidence in the research team and institutional safeguards, although some temporarily disabled the VPN during sensitive activities, underscoring the importance of this option for long-term adoption. Exploratory analyses further demonstrated that traffic-derived features captured daily activity rhythms and usage regularity, pointing toward applications in behavioral and health research.

Future Work and Limitations

At the same time, several limitations constrain the interpretation and generalizability of these findings. Most notably, our network traffic-based inference cannot definitively establish whether a user is actively engaging with a specific app. Background processes, automatic updates, and notifications may generate network activity even when the device is idle, and although percentile normalization mitigates some noise, it does not fully separate foreground interaction from idle background activity. In practice, however, the traffic profile (background vs. foreground) provides useful cues. Periods of very low, steady upload rates likely correspond to automated background processes—updates, app refreshes, or keep-alive signals—whereas sudden bursts and sustained spikes are characteristic of foreground activity when users are actively engaged. For example, Figure 11 illustrates extended stretches of background traffic during both sleep and work windows, punctuated by sharp peaks that indicate moments of user-driven activity.

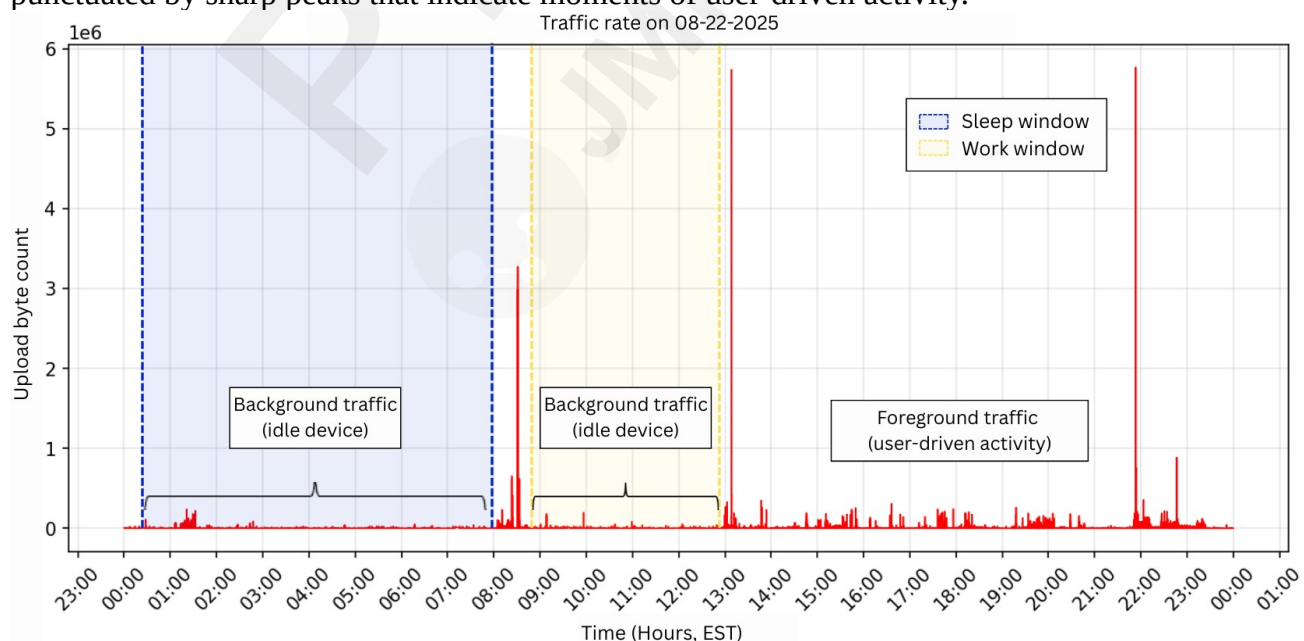


Figure 12. Daily upload traffic from a researcher's device on August 22, 2025. Background periods (e.g., sleep and work windows) are shaded, with bursts of foreground activity visible as spikes in upload volume.

While this separation is not definitive—since some background processes can occasionally mimic bursty activity—it nonetheless provides a reasonable approximation of when the device is idle versus actively used. Coupled with category-specific decomposition, these traffic signatures enrich our ability to reconstruct daily rhythms: not only what apps are used, but also when devices are idle, generating background traffic, or deliberately operated by the user.

It is also important to note that not all behaviors leave a network trace. Actions performed entirely offline—such as composing notes without iCloud sync or using other local-only apps—cannot be observed through our approach. From a research perspective, this granularity gap limits the ability to link outcomes to specific app-level behaviors, such as journaling in an offline notes app. At the same time, however, the inability to pinpoint exact apps can be viewed as a privacy-preserving safeguard, reducing risks of over-identification while still capturing meaningful behavioral rhythms.

A separate, more technical limitation arises from disconnections after phone restarts, which occasionally led to extended lapses in data collection. These gaps highlight the need for stronger reliability safeguards. In future iterations, a key design priority will be automated VPN reactivation and user-facing notifications when the VPN lapses, ensuring continuous coverage without requiring manual intervention.

Beyond technical reliability, a second challenge concerns the Hawthorne effect, where awareness of being monitored may lead participants to temporarily adjust their habits, raising questions about how closely the captured activity reflects “natural” behavior. In our deployment, this effect appeared minimal—several participants reported forgetting about the monitoring entirely—yet others described using the system productively as an opportunity to reflect on their digital habits. Future work should distinguish between these two forms of reactivity, separating distortion from constructive self-awareness.

The demographic profile of our sample represents another important limitation. Participants were university students from a single elite U.S. institution, a group with greater exposure to and familiarity with VPN technology [118]. In contrast, although technology adoption among older adults is increasing, they still lag behind in overall use [119,122]. University students’ relative ease with setup and comfort with privacy trade-offs may not generalize to other populations, where mental models of digital monitoring and tolerance for passive tracking can differ substantially. At the same time, recent studies show that older adults are becoming more familiar with digital tools and increasingly motivated to use them for health-related purposes, such as tracking physical activity, monitoring stress, or measuring vital signs [123–127]. Importantly, once older adults adopt new technologies, they often report positive attitudes toward their continued use [119–121]. Broader recruitment will therefore be critical for assessing both feasibility and acceptability in more diverse and representative cohorts, particularly given the potential relevance of this approach to aging populations, where phone usage patterns have been linked to cognition and health decline [100].

Beyond sample composition, another avenue for future work is cross-device monitoring. While this study focused on smartphones, daily digital activity also unfolds across laptops and tablets. As WireGuard is compatible with all major operating systems, extending the pipeline to multiple devices is technically feasible and would provide a more holistic picture of how individuals transition between contexts and platforms throughout the day.

Equally important is validation against independent ground truth. Our pipeline demonstrated feasibility and acceptability but was not validated against external measures. For example, incorporating accelerometer signals could help disambiguate active interaction from background traffic, while wearables such as smart or fitness wearables could provide measures of sleep timing and fragmentation. Location data, meanwhile, could situate digital behaviors within environmental and social contexts, yielding richer interpretations of behavioral rhythms. In addition, standardized mental health surveys and experience sampling could provide complementary self-reported outcomes, enabling links between digital activity patterns and psychological well-being. In addition, future studies could incorporate self-reported measures of well-being, affect, and mental health

through experience sampling or ecological momentary assessment [147]. These data streams could be temporally aligned with network traffic-derived digital behaviors at the daily level, enabling within-person analyses that connect fluctuations in mood and mental state with patterns of online activity.

Finally, methodological refinements will be necessary to capture the full granularity of app usage. Our preliminary experiments suggest that different functionalities—such as text messaging versus video calls—produce distinct network signatures. By systematically labeling such patterns in controlled environments and pairing them with user-reported logs, classifiers could be trained to infer fine-grained behaviors at scale. Advanced analytic frameworks, including causal modeling, anomaly detection, and rhythm analysis, could then transform temporal network structures into actionable digital biomarkers.

Taken together, these extensions—diverse populations, cross-device integration, multimodal validation, and analytic refinement—would move our system beyond feasibility and acceptability toward a robust, generalizable platform for longitudinal behavioral monitoring, with applications ranging from circadian rhythm research to cognitive health.

Relation to Prior Work

In relation to prior work, our approach occupies a distinct position between high-burden active sensing and opaque aggregate analytics. Self-reports, app usage logs, and wearables have provided important insights into digital behavior but remain limited by recall bias, adherence decay, and participant burden. Prior network-traffic approaches have largely emphasized device identification or coarse activity classification. By contrast, our method emphasizes the timing and structure of encrypted traffic, enabling fine-grained resolution while preserving privacy. Unlike platform-specific tools such as Apple Screen Time or Android Digital Wellbeing, it is device- and operating system-agnostic, offering second-level granularity without requiring invasive permissions or draining battery. This combination of precision, portability, and scalability highlights its potential as a robust and flexible foundation for large-scale behavioral research.

Conclusion

Our study demonstrates the feasibility and acceptability of a passive monitoring system for encrypted network traffic that can be mapped into interpretable activity patterns. By sustaining unobtrusive capture over extended periods, the approach can yield rhythms that align with circadian patterns and digital routines, while minimizing participant burden. Limitations in app-level inference, demographic scope, and device coverage highlight clear avenues for refinement, which hold the promise for improving the method. With broader recruitment, multi-device integration, and validation against ground truth, our passive monitoring system could significantly expand the digital phenotyping toolkit, particularly for investigating sleep, mental health, and cognitive change across the lifespan at scale.

References

1. Paulsen JS, Nance M, Kim JI, et al. A Review of Quality of Life after Predictive Testing for and Earlier Identification of Neurodegenerative Diseases. *Progress in Neurobiology*; 2013;110. doi:10.1016/j.pneurobio.2013.08.003

2. Baquero M. Depressive symptoms in neurodegenerative diseases. *World Journal of Clinical Cases*. 2015;3(8):682. doi:10.12998/wjcc.v3.i8.682
3. Bateman RJ, Xiong C, Benzinger TLS, et al. Clinical and Biomarker Changes in Dominantly Inherited Alzheimer's Disease. *New England Journal of Medicine*. 2012;367(9):795-804. doi:10.1056/nejmoa1202753
4. Price JL, Morris JC. Tangles and Plaques in Nondemented Aging and "preclinical" Alzheimer's Disease. *Annals of Neurology*. 1999;45(3):358-368. doi:10.1002/1531-8249(199903)45:3%3C358::aid-ana12%3E3.0.co;2-x
5. Fortea J, Vilaplana E, Carmona-Iragui M, et al. Clinical and biomarker changes of Alzheimer's disease in adults with Down syndrome: a cross-sectional study. *The Lancet*. 2020;395(10242):1988-1997. doi:10.1016/s0140-6736(20)30689-9
6. Beason-Held LL, Goh JO, An Y, et al. Changes in Brain Function Occur Years before the Onset of Cognitive Impairment. *Journal of Neuroscience*. 2013;33(46):18008-18014. doi:10.1523/jneurosci.1402-13.2013
7. Braak H, Braak E. Frequency of Stages of Alzheimer-Related Lesions in Different Age Categories. *Neurobiology of Aging*. 1997;18(4):351-357. doi:10.1016/s0197-4580(97)00056-0
8. Bäckman L, Small BJ, Fratiglioni L. Stability of the preclinical episodic memory deficit in Alzheimer's disease. *Brain*. 2001;124(1):96-102. doi:10.1093/brain/124.1.96
9. Lim YY, Maruff P, Pietrzak RH, et al. Effect of amyloid on memory and non-memory decline from preclinical to clinical Alzheimer's disease. *Brain*. 2013;137(1):221-231. doi:10.1093/brain/awt286
10. Silveri MC, Reali G, Jenner C, Puopolo M. Attention and memory in the preclinical stage of dementia. *Journal of geriatric psychiatry and neurology*. 2007;20(2):67-75. doi:10.1177/0891988706297469
11. Morris JC, Storandt M, Miller JP, et al. Mild Cognitive Impairment Represents Early-Stage Alzheimer Disease. *Archives of Neurology*. 2001;58(3). doi:10.1001/archneur.58.3.397
12. Small BJ, Mobly JL, Laukka EJ, Jones S, Backman L. Cognitive deficits in preclinical Alzheimer's disease. *Acta Neurologica Scandinavica*. 2003;107(s179):29-33. doi:10.1034/j.1600-0404.107.s179.6.x
13. Snyder PJ, Kahle-Wroblewski K, Brannan S, et al. Assessing cognition and function in Alzheimer's disease clinical trials: Do we have the right tools? *Alzheimer's & Dementia*. 2014;10(6):853-860. doi:10.1016/j.jalz.2014.07.158
14. Fiandaca MS, Mapstone ME, Cheema AK, Federoff HJ. The critical need for defining preclinical biomarkers in Alzheimer's disease. *Alzheimer's & Dementia*. 2014;10:S196-S212. doi:10.1016/j.jalz.2014.04.015
15. Kozauer N, Katz R. Regulatory Innovation and Drug Development for Early-Stage Alzheimer's Disease. *New England Journal of Medicine*. 2013;368(13):1169-1171. doi:10.1056/nejmp1302513
16. Iranzo A, Tolosa E, Gelpi E, et al. Neurodegenerative disease status and post-mortem pathology in idiopathic rapid-eye-movement sleep behaviour disorder: an observational cohort study. *The Lancet Neurology*. 2013;12(5):443-453. doi:10.1016/s1474-4422(13)70056-5
17. dos Santos AB, Kohlmeier KA, Barreto GE. Are Sleep Disturbances Preclinical Markers of Parkinson's Disease? *Neurochemical Research*. 2014;40(3):421-427. doi:10.1007/s11064-014-1488-7
18. Ma L. Depression, Anxiety, and Apathy in Mild Cognitive Impairment: Current Perspectives. *Frontiers in Aging Neuroscience*. 2020;12(9). doi:10.3389/fnagi.2020.00009
19. Gauthier S, Cummings J, Ballard C, et al. Management of behavioral problems in Alzheimer's disease. *International Psychogeriatrics*. 2010;22(3):346-372. doi:10.1017/s1041610209991505
20. Landes AM, Sperry SD, Strauss ME, Geldmacher DS. Apathy in Alzheimer's Disease. *Journal of*

- the American Geriatrics Society. 2001;49(12):1700-1707. doi:10.1046/j.1532-5415.2001.49282.x
21. Mega MS, Cummings JL, Fiorello T, Gornbein J. The spectrum of behavioral changes in Alzheimer's disease. *Neurology*. 1996;46(1):130-135. doi:10.1212/wnl.46.1.130
22. Oken BS, Salinsky MC, Elsas SM. Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*. 2006;117(9):1885-1901. doi:10.1016/j.clinph.2006.01.017
23. Cacioppo JT, Hawkley LC, Berntson GG, et al. Do lonely days invade the nights? Potential social modulation of sleep efficiency. *Psychological Science*. 2002;13(4):384-387. doi:10.1111/1467-9280.00469
24. Birchwood M, Spencer E, McGovern D. Schizophrenia: early warning signs. *Advances in Psychiatric Treatment*. 2000;6(2):93-101. doi:10.1192/apt.6.2.93
25. Wang R, Aung MSH, Abdullah S, et al. CrossCheck. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. Published online September 12, 2016. doi:10.1145/2971648.2971740
26. Kuiper JS, Zuidersma M, Zuidema SU, et al. Social relationships and cognitive decline: a systematic review and meta-analysis of longitudinal cohort studies. *International Journal of Epidemiology*. 2016;45(4):dyw089. doi:10.1093/ije/dyw089
27. Cacioppo JT, Hawkley LC. Perceived social isolation and cognition. *Trends in Cognitive Sciences*. 2009;13(10):447-454. doi:10.1016/j.tics.2009.06.005
28. Friedler B, Crapser J, McCullough L. One is the deadliest number: the detrimental effects of social isolation on cerebrovascular diseases and cognition. *Acta Neuropathologica*. 2014;129(4):493-509. doi:10.1007/s00401-014-1377-9
29. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between Loneliness and Perceived Social Support and Outcomes of Mental Health problems: a Systematic Review. *BMC Psychiatry*. 2018;18(156):1-16. doi:10.1186/s12888-018-1736-5
30. Silbert LC, Dodge HH, Lahna D, et al. Less Daily Computer Use is Related to Smaller Hippocampal Volumes in Cognitively Intact Elderly. *Journal of Alzheimer's Disease*. 2016;52(2):713-717. doi:10.3233/jad-160079
31. Altman S, Haeri S, Cohen LJ, et al. Predictors of relapse in bipolar disorder: a review. *J Psychiatr Pract*; 2006;12(5):269-82. doi:10.1097/00131746-200609000-00002.
32. Zhao J, Liu H, Wu Z, et al. Clinical features of the patients with major depressive disorder co-occurring insomnia and hypersomnia symptoms: a report of NSSD study. *Sleep Medicine*. 2021;81:375-381. doi:10.1016/j.sleep.2021.03.005
33. Bao YP, Han Y, Ma J, et al. Cooccurrence and bidirectional prediction of sleep disturbances and depression in older adults: Meta-analysis and systematic review. *Neuroscience and Biobehavioral Reviews*. 2017;75:257-273. doi:10.1016/j.neubiorev.2017.01.032
34. Lustberg L, Reynolds CF. Depression and insomnia: questions of cause and effect. *Sleep Medicine Reviews*. 2000;4(3):253-262. doi:10.1053/smr.1999.0075
35. Melo MCA, Garcia RF, Linhares Neto VB, et al. Sleep and circadian alterations in people at risk for bipolar disorder: A systematic review. *Journal of Psychiatric Research*. 2016;83:211-219. doi:10.1016/j.jpsychires.2016.09.005
36. Ritter PS, Marx C, Bauer M, Lepold K, Pfennig A. The role of disturbed sleep in the early recognition of bipolar disorder: a systematic review. *Bipolar Disorders*. 2011;13(3):227-237. doi:10.1111/j.1399-5618.2011.00917.x
37. Geoffroy PA, Scott J, Boudebessé C, et al. Sleep in patients with remitted bipolar disorders: a meta-analysis of actigraphy studies. *Acta Psychiatrica Scandinavica*. 2014;131(2):89-99. doi:10.1111/acps.12367

38. Wang Y, Ren X, Liu X, Zhu T. Examining the Correlation Between Depression and Social Behavior on Smartphones Through Usage Metadata: Empirical Study. *JMIR mHealth and uHealth*. 2021;9(1):e19046. doi:10.2196/19046
39. Ahn S, Springer K, Gibson JS. Social withdrawal in Parkinson's disease: A scoping review. *Geriatric Nursing*. 2022;48(48):258-268. doi:10.1016/j.gerinurse.2022.10.010
40. Kaye J, Mattek N, Dodge HH, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. *Alzheimer's & Dementia*. 2014;10(1):10-17. doi:10.1016/j.jalz.2013.01.011
41. Lussier M, Lavoie M, Giroux S, et al. Early Detection of Mild Cognitive Impairment With In-Home Monitoring Sensor Technologies Using Functional Measures: A Systematic Review. *IEEE Journal of Biomedical and Health Informatics*. 2019;23(2):838-847. doi:10.1109/jbhi.2018.2834317
42. Piau A, Wild K, Mattek N, Kaye J. Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care: Systematic Review. *Journal of Medical Internet Research*. 2019;21(8):e12785. doi:10.2196/12785
43. John P.K. Bernstein, Dorociak KE, Mattek N, et al. Passively-Measured Routine Home Computer Activity and Application Use Can Detect Mild Cognitive Impairment and Correlate with Important Cognitive Functions in Older Adulthood. *Journal of Alzheimer's Disease*. 2021;81(3):1053-1064. doi:10.3233/jad-210049
44. Austin J, Hollingshead K, Kaye J. Internet Searches and Their Relationship to Cognitive Function in Older Adults: Cross-Sectional Analysis. *Journal of Medical Internet Research*. 2017;19(9):e307. doi:10.2196/jmir.7671
45. Dorsey E Ray, Papapetropoulos S, Xiong M, Kiebertz K. The First Frontier: Digital Biomarkers for Neurodegenerative Disorders. *Digital Biomarkers*. Published online July 4, 2017. doi:10.1159/000477383
46. Schütz N, Knobel SEJ, Botros A, et al. A systems approach towards remote health-monitoring in older adults: Introducing a zero-interaction digital exhaust. *npj Digital Medicine*. 2022;5(1):1-13. doi:10.1038/s41746-022-00657-y
47. Vasudevan S, Saha A, Tarver ME, Patel B. Digital biomarkers: Convergence of digital health technologies and biomarkers. *npj Digital Medicine*. 2022;5(1). doi:10.1038/s41746-022-00583-z
48. Babrak Lmar M, Menetski J, Rebhan M, et al. Traditional and Digital Biomarkers: Two Worlds Apart? *Digital Biomarkers*. 2019;3(2):92-102. doi:10.1159/000502000
49. Vega J, Jay C, Vigo M, Harper S. Unobtrusive Monitoring of Parkinson's Disease Based on Digital Biomarkers of Human Behaviour. *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*. Published online October 19, 2017:351-352. doi:10.1145/3132525.3134782
50. Mehrotra AK, Mirco Musolesi. Designing Effective Movement Digital Biomarkers for Unobtrusive Emotional State Mobile Monitoring. *UCL Discovery (University College London)*. Published online June 23, 2017. doi:10.1145/3089341.3089342
51. Wacks Y, Weinstein AM. Excessive Smartphone Use Is Associated With Health Problems in Adolescents and Young Adults. *Frontiers in Psychiatry*. 2021;12:669042. doi:10.3389/fpsy.2021.669042
52. Onnela JP, Rauch SL. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. *Neuropsychopharmacology*. 2016;41(7):1691-1696. doi:10.1038/npp.2016.7
53. Nickels S, Edwards MD, Poole SF, et al. Toward a Mobile Platform for Real-world Digital Measurement of Depression: User-Centered Design, Data Quality, and Behavioral and Clinical Modeling. *JMIR Mental Health*. 2021;8(8):e27589. doi:10.2196/27589

54. Onnela JP. Opportunities and challenges in the collection and analysis of digital phenotyping data. *Neuropsychopharmacology*. Published online July 17, 2020. doi:10.1038/s41386-020-0771-3
55. Teo JX, Davila S, Yang C, et al. Digital phenotyping by consumer wearables identifies sleep-associated markers of cardiovascular disease risk and biological aging. *Communications Biology*. 2019;2(1). doi:10.1038/s42003-019-0605-1
56. Alireza Sameh, Rostami M, Mourad Oussalah, Raija Korpelainen, Vahid Farrahi. Digital phenotypes and digital biomarkers for health and diseases: a systematic review of machine learning approaches utilizing passive non-invasive signals collected via wearable devices and smartphones. *Artificial Intelligence Review*. 2024;58(2). doi:10.1007/s10462-024-11009-5
57. Smets E, Rios Velazquez E, Schiavone G, et al. Large-scale wearable data reveal digital phenotypes for daily-life stress detection. *npj Digital Medicine*. 2018;1(1). doi:10.1038/s41746-018-0074-9
58. Juan Carlos Torrado, Husebo BS, Allore HG, et al. Digital phenotyping by wearable-driven artificial intelligence in older adults and people with Parkinson's disease: Protocol of the mixed method, cyclic ActiveAgeing study. *PLOS ONE*. 2022;17(10):e0275747-e0275747. doi:10.1371/journal.pone.0275747
59. Overview of depression and psychosis in Alzheimer's disease. *American Journal of Psychiatry*. 1989;146(5):577-587. doi:10.1176/ajp.146.5.577
60. Spira AP, Gamaldo AA, An Y, et al. Self-Reported Sleep and β -Amyloid Deposition in Community-Dwelling Older Adults. *JAMA neurology*. 2013;70(12):1537-1543. doi:10.1001/jamaneurol.2013.4258
61. Ifrah Zawar, Mattos MK, Manning C, Patrie J, Quigg M. Sleep Disturbances Predict Cognitive Decline in Cognitively Healthy Adults. *Journal of Alzheimer's disease*. 2023;92(4):1427-1438. doi:10.3233/jad-221244
62. Mander BA, Winer JR, Jagust WJ, Walker MP. Sleep: A Novel Mechanistic Pathway, Biomarker, and Treatment Target in the Pathology of Alzheimer's Disease? *Trends in Neurosciences*. 2016;39(8):552-566. doi:10.1016/j.tins.2016.05.002
63. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep Patterns and Predictors of Disturbed Sleep in a Large Population of College Students. *Journal of Adolescent Health*. 2010;46(2):124-132. doi:10.1016/j.jadohealth.2009.06.016
64. Wang F, B  r   E. Determinants of sleep quality in college students: A literature review. *EXPLORE*. 2021;17(2):170-177. doi:10.1016/j.explore.2020.11.003
65. Friedrich A, Schlarb AA. Let's talk about sleep: a systematic review of psychological interventions to improve sleep in college students. *Journal of Sleep Research*. 2018;27(1):4-22. doi:10.1111/jsr.12568
66. Suzuki H, Savitz J, Kent Teague T, et al. Altered populations of natural killer cells, cytotoxic T lymphocytes, and regulatory T cells in major depressive disorder: Association with sleep disturbance. *Brain, Behavior, and Immunity*. 2017;66:193-200. doi:10.1016/j.bbi.2017.06.011
67. Gitanjali B. Establishing a polysomnography laboratory in India: problems and pitfalls. *Sleep*. 1998;21(4):331-332. doi:10.1093/sleep/21.4.331
68. Pan Q, Brulin D, Campo E. Current Status and Future Challenges of Sleep Monitoring Systems: Systematic Review. *JMIR Biomedical Engineering*. 2020;5(1):e20921. doi:10.2196/20921
69. Luqman A, Masood A, Shahzad F, Shahbaz M, Feng Y. Untangling the adverse effects of late-night usage of smartphone-based SNS among University students. *Behaviour & Information Technology*. 2020;40(15):1-17. doi:10.1080/0144929x.2020.1773538
70. Dewi RK, Efendi F, Has EMM, Gunawan J. Adolescents' smartphone use at night, sleep disturbance and depressive symptoms. *International Journal of Adolescent Medicine and Health*.

2018;0(0). doi:10.1515/ijamh-2018-0095

71. Lemola S, Perkinson-Gloor N, Brand S, Dewald-Kaufmann JF, Grob A. Adolescents' Electronic Media Use at Night, Sleep Disturbance, and Depressive Symptoms in the Smartphone Age. *Journal of Youth and Adolescence*. 2015;44(2):405-418. doi:10.1007/s10964-014-0176-x

72. Smith JA. *Qualitative Psychology : A Practical Guide to Research Methods*. 3rd ed. Sage; 2015:222-248.

73. Willig C, Rogers WS. *The SAGE Handbook of Qualitative Research in Psychology*. 2nd ed. Sage Publications; 2017.

74. Guest G, Macqueen K, Namey E. Introduction to Applied Thematic Analysis In: *Applied Thematic Analysis Introduction to Applied Thematic Analysis*. SAGE Research Methods. 2012;3(20). doi:10.4135/9781483384436

75. Cairns P, Cox AL. *Research methods for human-computer interaction*. New York, NY: Cambridge University Press; 2008. ISBN:9780521870122.

76. Alexander JD, Linkersdörfer J, Toda-Thorne K, et al. Passively sensing smartphone use in teens with rates of use by sex and across operating systems. *Scientific Reports*. 2024;14(1):17982. doi:10.1038/s41598-024-68467-8

77. Hart SG. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 2006;50(9):904-908. doi:10.1177/154193120605000909

78. Brooke J. SUS: A "Quick and Dirty" Usability Scale. *Usability Evaluation In Industry*. Published online June 11, 1996:207-212. doi:10.1201/9781498710411-35

79. Sauro J. *A Practical Guide to the System Usability Scale : Background, Benchmarks and Best Practices*. Createspace; 2011.

80. Raugh IM, James SH, Gonzalez CM, et al. Digital phenotyping adherence, feasibility, and tolerability in outpatients with schizophrenia. *Journal of Psychiatric Research*. 2021;138:436-443. doi:10.1016/j.jpsychires.2021.04.022

81. Olaye IM, Belovsky MP, Bataille L, et al. Recommendations for Defining and Reporting Adherence Measured by Biometric Monitoring Technologies: Systematic Review. *Journal of Medical Internet Research*. 2022;24(4):e33537. doi:10.2196/33537

82. Torous J, Kiang MV, Lorme J, Onnela JP. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. *JMIR Mental Health*. 2016;3(2):e16. doi:10.2196/mental.5165

83. Torous J, Staples P, Barnett I, Sandoval LR, Keshavan M, Onnela JP. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. *npj Digital Medicine*. 2018;1(1). doi:10.1038/s41746-018-0022-8

84. Currey D, Torous J. Increasing the value of digital phenotyping through reducing missingness: a retrospective review and analysis of prior studies. 2023;26(1):e300718-e300718. doi:10.1136/bmjment-2023-300718

85. Knowlden AP, Sharma M, Bernard AL. A Theory of Planned Behavior Research Model for Predicting the Sleep Intentions and Behaviors of Undergraduate College Students. *The Journal of Primary Prevention*. 2012;33(1):19-31. doi:10.1007/s10935-012-0263-2

86. Pilcher JJ, Ginter DR, Sadowsky B. Sleep quality versus sleep quantity: Relationships between sleep and measures of health, well-being and sleepiness in college students. *Journal of Psychosomatic Research*. 1997;42(6):583-596. doi:10.1016/s0022-3999(97)00004-4

87. Ford DE, Kamerow DB. Epidemiologic Study of Sleep Disturbances and Psychiatric Disorders. *JAMA*. 1989;262(11):1479. doi:10.1001/jama.1989.03430110069030

88. Capraro V, Globig L, Rausch Z, et al. A consensus statement on potential negative impacts of smartphone and social media use on adolescent mental health. Published online May 15, 2025.

doi:10.31234/osf.io/b94dy_v1

89. De Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC. Wearable Sleep Technology in Clinical and Research Settings. *Medicine & Science in Sports & Exercise*. 2019;51(7):1538-1557. doi:10.1249/mss.0000000000001947

90. Toh SH, Howie EK, Coenen P, Straker LM. "From the moment I wake up I will use it...every day, very hour": a qualitative study on the patterns of adolescents' mobile touch screen device use from adolescent and parent perspectives. *BMC Pediatrics*. 2019;19(1). doi:10.1186/s12887-019-1399-5

91. Borger JN, Huber R, Ghosh A. Capturing sleep-wake cycles by using day-to-day smartphone touchscreen interactions. *npj Digital Medicine*. 2019;2(1). doi:10.1038/s41746-019-0147-4

92. Till Roenneberg. Twitter as a means to study temporal behaviour. *Current Biology*. 2017;27(17):R830-R832. doi:10.1016/j.cub.2017.08.005

93. Leypunskiy E, Kiciman E, Shah M, et al. Geographically Resolved Rhythms in Twitter Use Reveal Social Pressures on Daily Activity Patterns. *Current Biology*. 2018;28(23):3763-3775.e5. doi:10.1016/j.cub.2018.10.016

94. Villanti AC, Johnson AL, Ilakkuvan V, Jacobs MA, Graham AL, Rath JM. Social Media Use and Access to Digital Technology in US Young Adults in 2016. *Journal of Medical Internet Research*. 2017;19(6). doi:10.2196/jmir.7303

95. Ram N, Yang X, Cho MJ, et al. Screenomics: A New Approach for Observing and Studying Individuals' Digital Lives. *Journal of Adolescent Research*. 2019;35(1):16-50. doi:10.1177/0743558419883362

96. Karas M, Huang D, Clement Z, et al. Smartphone Screen Time Characteristics in People With Suicidal Thoughts: Retrospective Observational Data Analysis Study. *JMIR mhealth and uhealth*. 2024;12:e57439-e57439. doi:10.2196/57439

97. Onnela JP, Dixon C, Griffin K, et al. Beiwe: A data collection platform for high-throughput digital phenotyping. *Journal of Open Source Software*. 2021;6(68):3417. doi:10.21105/joss.03417

98. Lind MN, Byrne ML, Wicks G, Smidt AM, Allen NB. The Effortless Assessment of Risk States (EARS) Tool: An Interpersonal Approach to Mobile Sensing. *JMIR Mental Health*. 2018;5(3):e10334. doi:10.2196/10334

99. Alexander JD, Linkersdörfer J, Toda-Thorne K, et al. Passively sensing smartphone use in teens with rates of use by sex and across operating systems. *Scientific Reports*. 2024;14(1):17982. doi:10.1038/s41598-024-68467-8

100. Bengtson JF, Scullin MK. A meta-analysis of technology use and cognitive aging. *Nature Human Behaviour*. Published online April 14, 2025. doi:10.1038/s41562-025-02159-9

101. Beukenhorst AL, Burke KM, Scheier Z, et al. Using Smartphones to Reduce Research Burden in a Neurodegenerative Population and Assessing Participant Adherence: A Randomized Clinical Trial and Two Observational Studies. *JMIR mHealth and uHealth*. 2022;10(2):e31877. doi:10.2196/31877

102. Incel OD, Kose M, Ersoy C. A Review and Taxonomy of Activity Recognition on Mobile Phones. *BioNanoScience*. 2013;3(2):145-171. doi:10.1007/s12668-013-0088-3

103. Amor JD, James CJ. Setting the scene: Mobile and wearable technology for managing healthcare and wellbeing. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Published online August 2015. doi:10.1109/embc.2015.7320189

104. Schlosser, Kurt. "New research finds 95% of teens have access to a smartphone; 45% online 'almost constantly'." *GeekWire* (2018).

105. Google. Background Execution Limits. *Android Developers*. Google; 2017. Available from: <https://developer.android.com/about/versions/oreo/background> [accessed Sep 12, 2025]

106. Apple Inc. Security of Runtime Process in iOS, iPadOS, and visionOS. *Apple Platform Security*.

- Apple Inc.; 2024. Available from: <https://support.apple.com/guide/security/security-of-runtime-process-sec15bfe098e/web> [accessed Sep 12, 2025]
107. Lombardi DB, Ciceri MR. More than defense in daily experience of privacy: The functions of privacy in digital and physical environments. *Europe's Journal of Psychology*. 2016;12(1):115-136. doi:10.5964/ejop.v12i1.948
108. Varenkamp, Peter. iPhone Acquisition Using Jailbreaking Techniques. MS thesis. NTNU, 2019.
109. George MJ, Beron K, Vollet JW, Burnell K, Ehrenreich SE, Underwood MK. Frequency of Text Messaging and Adolescents' Mental Health Symptoms Across 4 Years of High School. *Journal of Adolescent Health*. 2021;68(2):324-330. doi:10.1016/j.jadohealth.2020.06.012
110. Guracho YD, Thomas SJ, Win KT. Smartphone application use patterns for mental health disorders: A systematic literature review and meta-analysis. *International Journal of Medical Informatics*. 2023;179:105217. doi:10.1016/j.ijmedinf.2023.105217
111. Naslund JA, Bondre A, Torous J, Aschbrenner KA. Social Media and Mental Health: Benefits, Risks, and Opportunities for Research and Practice. *Journal of Technology in Behavioral Science*. 2020;5(3):245-257. doi:10.1007/s41347-020-00134-x
112. Bucci S, Schwannauer M, Berry N. The digital revolution and its impact on mental health care. *Psychology and psychotherapy*. 2019;92(2):277-297.
113. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. Feasibility and acceptability of Facebook for health promotion among people with serious mental illness. *DIGITAL HEALTH*. 2016;2:205520761665482. doi:10.1177/2055207616654822
114. Almourad MB, Alrobai A, Skinner T, Hussain M, Ali R. Digital wellbeing tools through users lens. *Technology in Society*. 2021;67:101778. doi:10.1016/j.techsoc.2021.101778
115. Statt N. Verizon's 'supercookie' earns the company a \$1.35 million fine from the FCC. *The Verge*. 2016 Mar 7. Available from: <https://www.theverge.com/2016/3/7/11173010/verizon-supercookie-fine-1-3-million-fcc> [accessed 2025-09-15]
116. Donenfeld JA. WireGuard: Next Generation Kernel Network Tunnel. *Proceedings 2017 Network and Distributed System Security Symposium*. Published online 2017. doi:10.14722/ndss.2017.23160
117. Wireshark. tshark(1) Manual Page. Available from: <https://www.wireshark.org/docs/man-pages/tshark.html> [accessed 2025-09-15]
118. Dutkowska-Zuk A, et al. How and why people use virtual private networks. In: 31st USENIX Security Symposium (USENIX Security 22); 2022 Aug 10-12; Boston, MA. Berkeley, CA: USENIX Association; 2022.
119. Smith A. Older adults and technology use. Pew Research Center; 2014. Available from: <https://www.pewresearch.org/internet/2014/04/03/older-adults-and-technology-use/>
120. Heinz M, Martin P, Margrett JA, et al. Perceptions of Technology among Older Adults. *Journal of Gerontological Nursing*. 2013;39(1):42-51. doi:10.3928/00989134-20121204-04
121. Mitzner TL, Boron JB, Fausset CB, et al. Older adults talk technology: Technology usage and attitudes. *Computers in Human Behavior*. 2010;26(6):1710-1721. doi:10.1016/j.chb.2010.06.020
122. Czaja SJ, Charness N, Fisk AD, et al. Factors Predicting the Use of Technology: Findings From the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging*. 2006;21(2):333-352. doi:10.1037/0882-7974.21.2.333
123. Hossein Matlabi, Parker SG, McKee K. Experiences of Extra Care Housing Residents Aged Fifty-Five and Over with Home-Based Technology. *Social Behavior and Personality An International Journal*. 2012;40(2):293-300. doi:10.2224/sbp.2012.40.2.293
124. Uei SL, Tsai CH, Yang MS. Telecare service use among Taiwanese aged 60 and over: Satisfaction, trust, and continued use intention. *Social Behavior and Personality: an international journal*. 2013;41(8):1309-1318. doi:10.2224/sbp.2013.41.8.1309

125. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew Research Center: Internet, Science & Tech; 2017. Available from: <https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/>
126. Aitken, Murray, Brian Clancy, and Deloitte Nass. "The growing value of digital health: evidence and impact on human health and the healthcare system." IQVIA Institute for Human Data Science 1 (2017).
127. Kononova A, Li L, Kamp K, et al. The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change. *JMIR mHealth and uHealth*. 2019;7(4):e9832. doi:10.2196/mhealth.9832
128. Otter.ai. Otter: AI meeting assistant and real-time voice transcription. Available from: <https://otter.ai/> [accessed 2025-09-15]
129. Kristensen PL, Olesen LG, Egebæk HK, Pedersen J, Rasmussen MG, Grøntved A. Criterion validity of a research-based application for tracking screen time on android and iOS smartphones and tablets. *Computers in Human Behavior Reports*. 2022;5:100164. doi:10.1016/j.chbr.2021.100164
130. Parry DA, le Roux DB, Morton J, Pons R, Pretorius R, Schoeman A. Digital wellbeing applications: Adoption, use and perceived effects. *Computers in Human Behavior*. 2023;139:107542. doi:10.1016/j.chb.2022.107542
131. DeBelle H, Packer E, Beales E, et al. Feasibility and usability of a digital health technology system to monitor mobility and assess medication adherence in mild-to-moderate Parkinson's disease. *Frontiers in Neurology*. 2023;14. doi:10.3389/fneur.2023.1111260
132. Shih HJS, Quinn L, Morgan-Jones P, Long K, Schreier AR, Friel CP. Wearable Activity Monitors to Support Physical Activity Interventions in Neurodegenerative Disease: A Feasibility Study. *Neurodegenerative Disease Management*. 2023;13(3):177-189. doi:10.2217/nmt-2022-0028
133. Huberty J, Ehlers DK, Kurka J, Ainsworth B, Buman M. Feasibility of three wearable sensors for 24 hour monitoring in middle-aged women. *BMC Women's Health*. 2015;15(1). doi:10.1186/s12905-015-0212-3
134. Caselli RJ, Langlais BT, Dueck AC, et al. Neuropsychological decline up to 20 years before incident mild cognitive impairment. *Alzheimer's & Dementia*. 2020;16(3):512-523. doi:10.1016/j.jalz.2019.09.085
135. Bloem BR, Okun MS, Klein C. Parkinson's Disease. *The Lancet*. 2021;397(10291):2284-2303. doi:10.1016/S0140-6736(21)00218-X
136. Trojanowski, John Q., and Harald Hampel. "Neurodegenerative disease biomarkers: guideposts for disease prevention through early diagnosis and intervention." *Progress in neurobiology* 95.4 (2011): 491.
137. Leifer, Bennett P. "Early diagnosis of Alzheimer's disease: clinical and economic benefits." *Journal of the American Geriatrics Society* 51.5s2 (2003): S281-S288.
138. Tarrant, Michael A., et al. "Effects of recall bias and nonresponse bias on self-report estimates of angling participation." *North American Journal of Fisheries Management* 13.2 (1993): 217-222.
139. Zhou, Leming, et al. "The mHealth app usability questionnaire (MAUQ): development and validation study." *JMIR mHealth and uHealth* 7.4 (2019): e11500.
140. Reeves, Byron, et al. "Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them." *Human-Computer Interaction* 36.2 (2021): 150-201.
141. Daniore P, Nittas V, von Wyl V. Enrollment and retention of participants in remote digital health studies: scoping review and framework proposal. *Journal of Medical Internet Research*. 2022;24(9):e39910. doi:10.2196/39910

142. Brady B, Zhou S, Ashworth D, Zheng L, Eramudugolla R, Anstey KJ. *Feasibility, adherence and usability of an observational digital health study built using Apple's ResearchKit among adults aged 18–84 years*. *Frontiers in Digital Health*. 2025; 7. doi: 10.3389/fdgth.2025.1520971
143. Pratap A, Neto EC, Snyder P, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *npj Digital Medicine*. 2020;3(1):1-10. doi:10.1038/s41746-020-0224-8
144. Daniore P, Nittas V, von Wyl V. Enrollment and retention of participants in remote digital health studies: a scoping review and framework proposal (Preprint). *Journal of Medical Internet Research*. 2022;24(9). doi:10.2196/39910
145. Aledavood T, Luong N, Baryshnikov I, et al. Multimodal Digital Phenotyping Study in Patients With Major Depressive Episodes and Healthy Controls (Mobile Monitoring of Mood): Observational Longitudinal Study. *JMIR mental health*. 2025;12:e63622. doi:10.2196/63622
146. Kelkar RS, Currey D, Srilakshmi Nagendra, et al. Utility of Smartphone-Based Digital Phenotyping Biomarkers in Assessing Treatment Response to Transcranial Magnetic Stimulation in Depression: Proof-of-Concept Study. *JMIR Formative Research*. 2023;7:e40197-e40197. doi:10.2196/40197
147. Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. *Annual Review of Clinical Psychology*. 2008;4(1):1-32. doi:10.1146/annurev.clinpsy.3.022806.091415
148. Apple. iTunes Search API. Performance Partners. Available from: <https://performance-partners.apple.com/search-api>
149. Shore A, Prena K. Platform rules as privacy tools: The influence of screenshot accountability and trust on privacy management. *New Media & Society*. Published online July 28, 2023. doi:10.1177/14614448231188929
150. Ingber AS. Understanding screenshot collection and sharing on messaging platforms: a privacy perspective. Vitak J, ed. *Journal of Computer-Mediated Communication*. 2025;30(1). doi:10.1093/jcmc/zmae023

Appendix

Semi-structured

Interview

Questions

Q1: How's your overall experience with the system? What was your favourite aspect of the system? What did you find the most troubling/challenging part of the system?

Q2: [If the tracking time is less than 14 days] I see that you've only used the system for N days. The target is 14. You're still being paid, but any reasons why you decided not to hit the 14-day target?

Q3: [If the tracking time is greater than 14 days] I see that you've used the system for more than 14 days, even though we asked you to use it for 14 days. Any reasons why you decided to go beyond the 14-day target (motivation)?

Q3: Have you noticed any changes (attitudes, actions, etc) from the beginning to the end while using the system?

Q4: If unlimited time, would you have continued using the VPN? Why or why not?

Q5: Do you have any general concerns while using the VPN?

Q6: What reasons make you trust the software to keep running in the background and track your network data?

Q7: Did you continuously stay on the VPN the whole time? Did you turn it off at any point? When and Why?

Q8: How often do you check the dashboard during the study? Any difference between the beginning and end of the study?

Q9: How can we improve the user experience of the system?

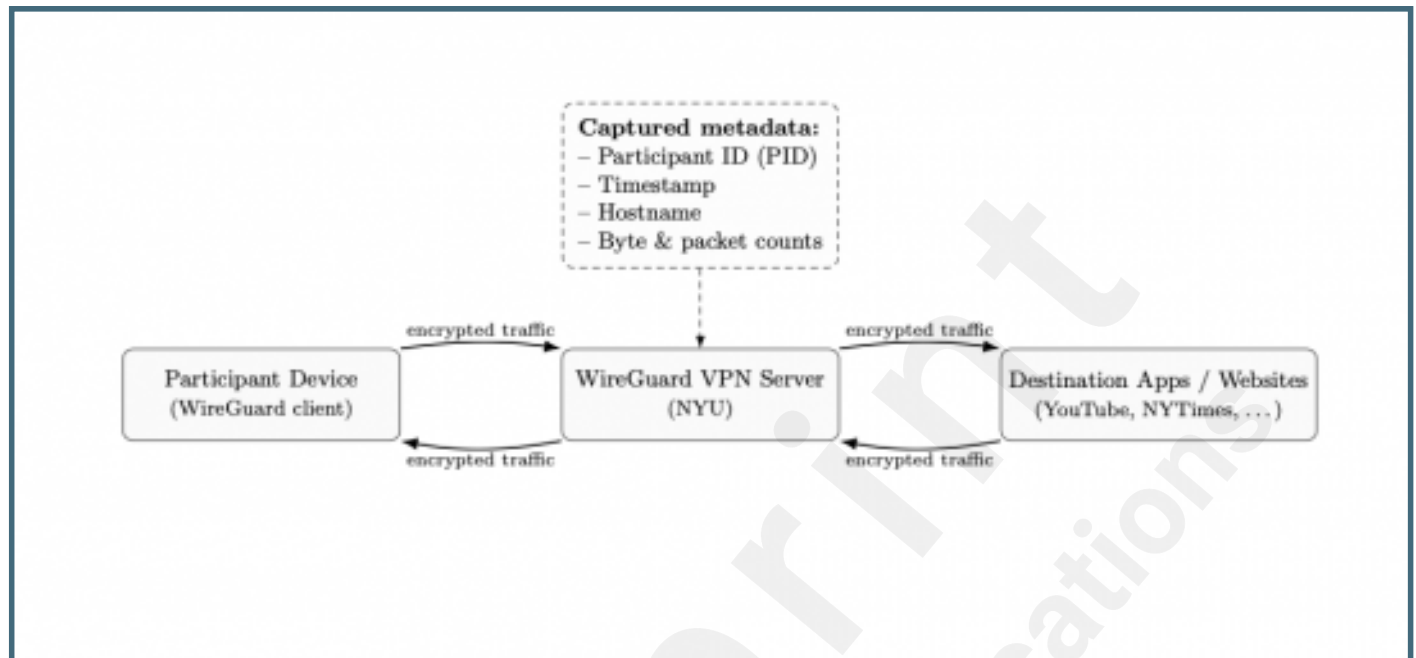
Scope	Feasibility (%)
Overall	72.8
Morning	69.7
Afternoon	74.4
Evening	76.2
Night	70.8
Weekday	73.0
Weekend	72.1

Table 3. Aggregate feasibility, calculated by pooling all monitoring bins across participants. This table reflects raw data yield, expressed as the proportion of expected bins with valid traffic.

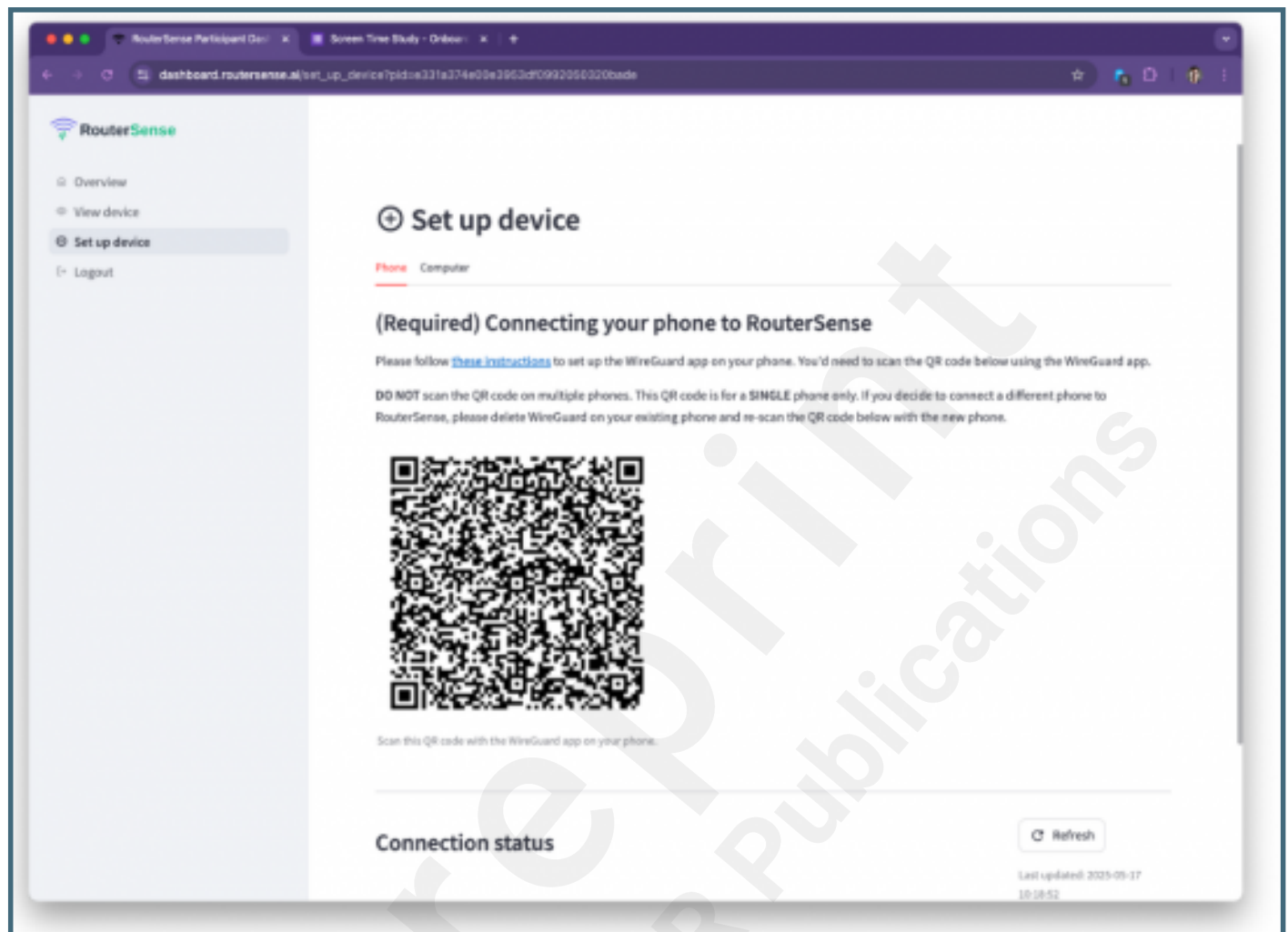
Supplementary Files

Figures

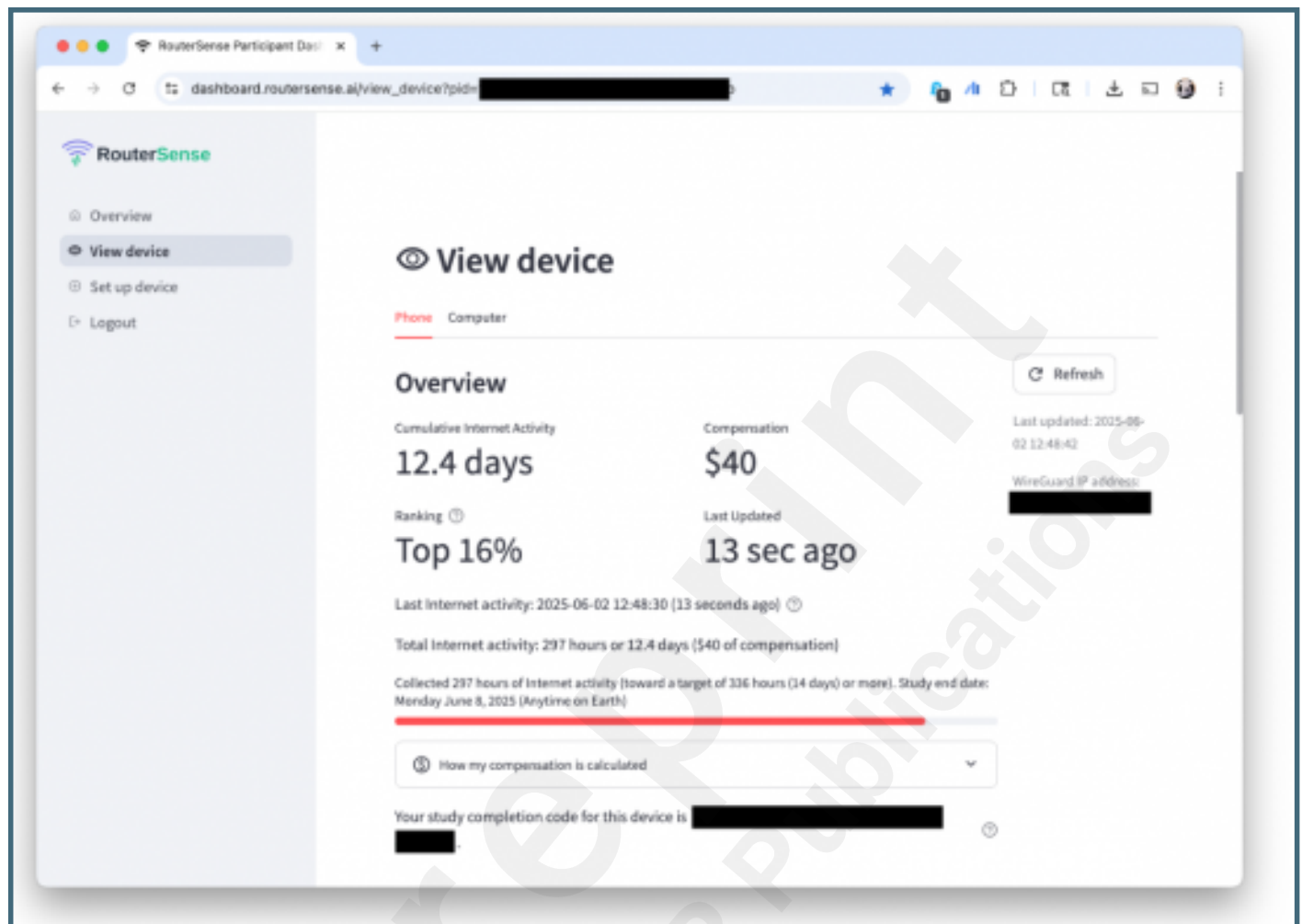
Architecture of the system. All device-to-internet traffic is routed through a WireGuard VPN server hosted at NYU. While packet contents remain end-to-end encrypted, the server captures metadata (e.g., hostnames and traffic volumes), enabling passive behavioral analysis while preserving message and content privacy.



Participant onboarding portal. Each participant scanned a unique QR code to automatically configure the WireGuard VPN client and route traffic through the study server.



Participant-facing study portal. The dashboard displays cumulative internet activity, compensation amount, and completion code, providing real-time transparency into data collection and payments.

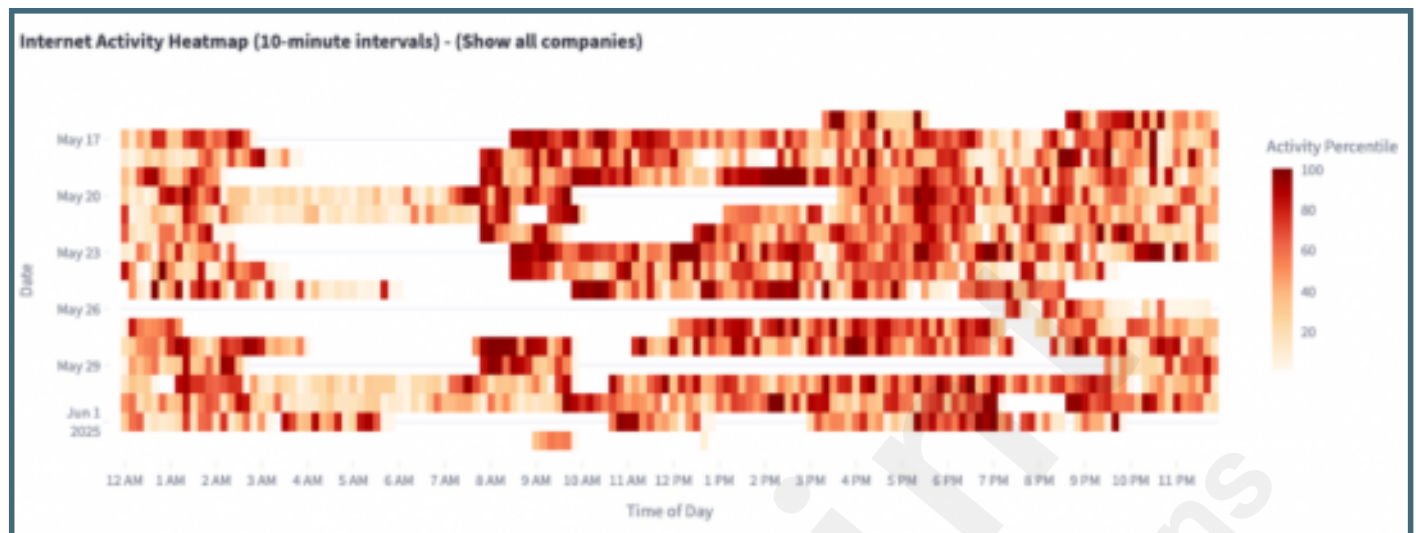


Real-time table on the participant dashboard displaying hostnames contacted during app use (shown here for a YouTube session: youtube.com, googlevideo.com, yting.com).

Hourly Report

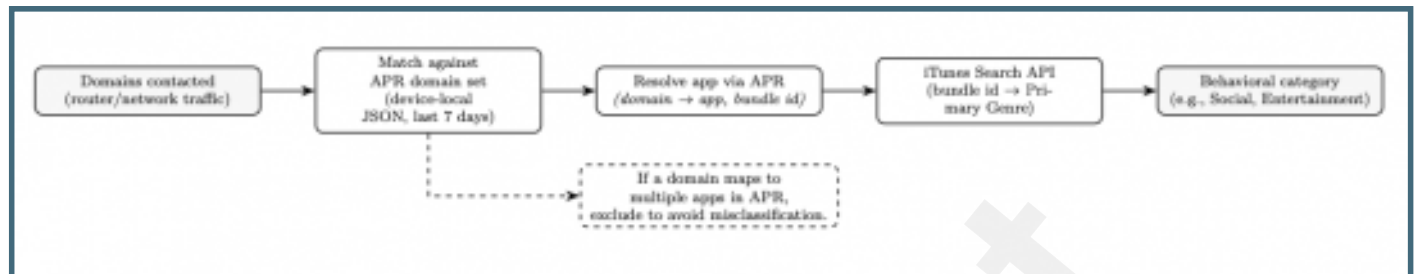
Time	Domain	Company	Download (kB)	Upload (kB)
21:24	yting.com	Google	132.61	3.05
21:23	app-analytics-services.co	Google	16.77	6.02
21:23	icloud.com	Apple	32.36	25.45
21:23	apple.com	Apple	20.93	4.60
21:23	ggght.com	Google	467.21	16.23
21:23	google.com	Google	17.23	8.24
21:23	googleapis.com	Google	546.77	368.06
21:23	googlevideo.com	Google	4,510.61	45.10
21:23	youtube.com	Google	33.58	56.11
21:23	yting.com	Google	440.06	11.19

Internet activity heatmap aggregated in 10-minute intervals over 24 hours. Rows represent study days and columns represent time of day; color intensity indicates the percentile of observed activity.



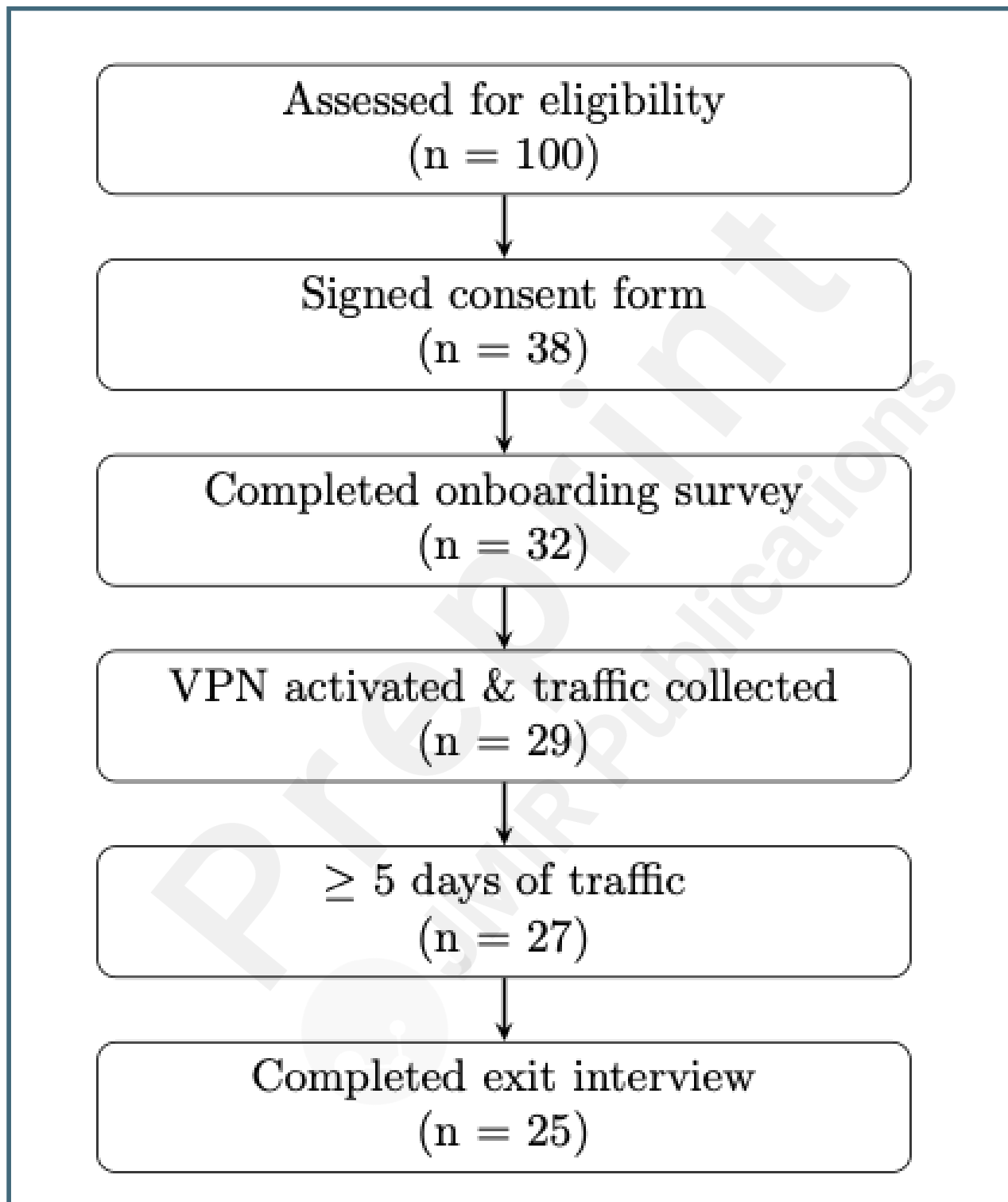
Please Select A Types

Hostname ? App ? Category pipeline. Domains observed in network traffic are matched against Apple's App Privacy Report (APR) to recover the generating app (via bundle ID), which is then linked to its primary genre using the iTunes Search API. This process yields higher-level behavioral categories that complement app-level mappings.

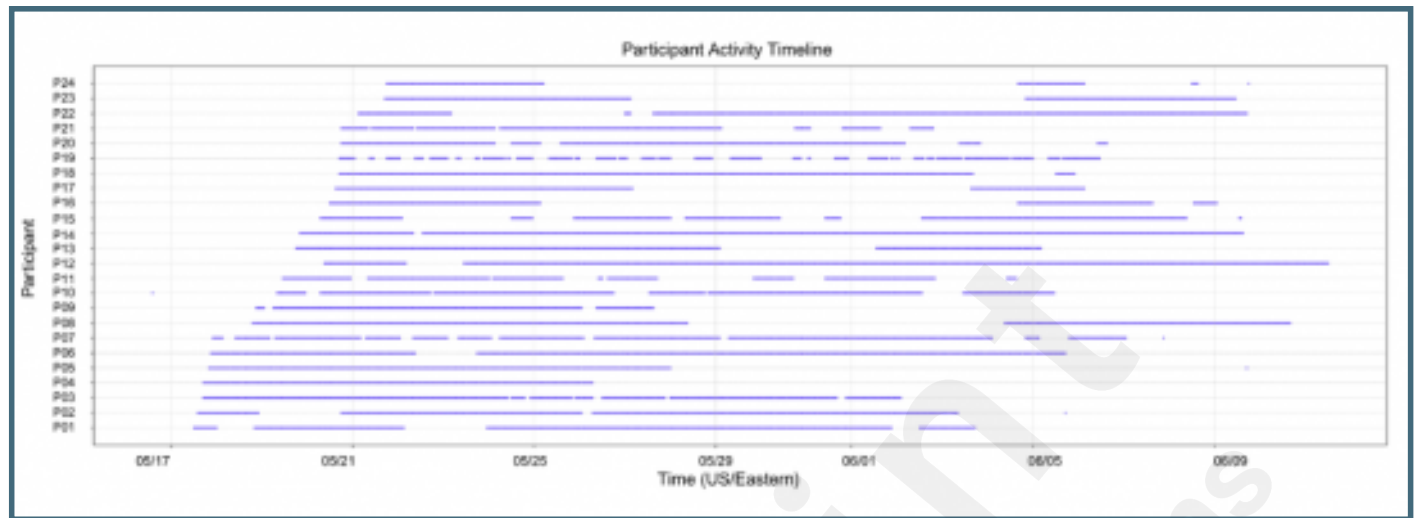


Figures

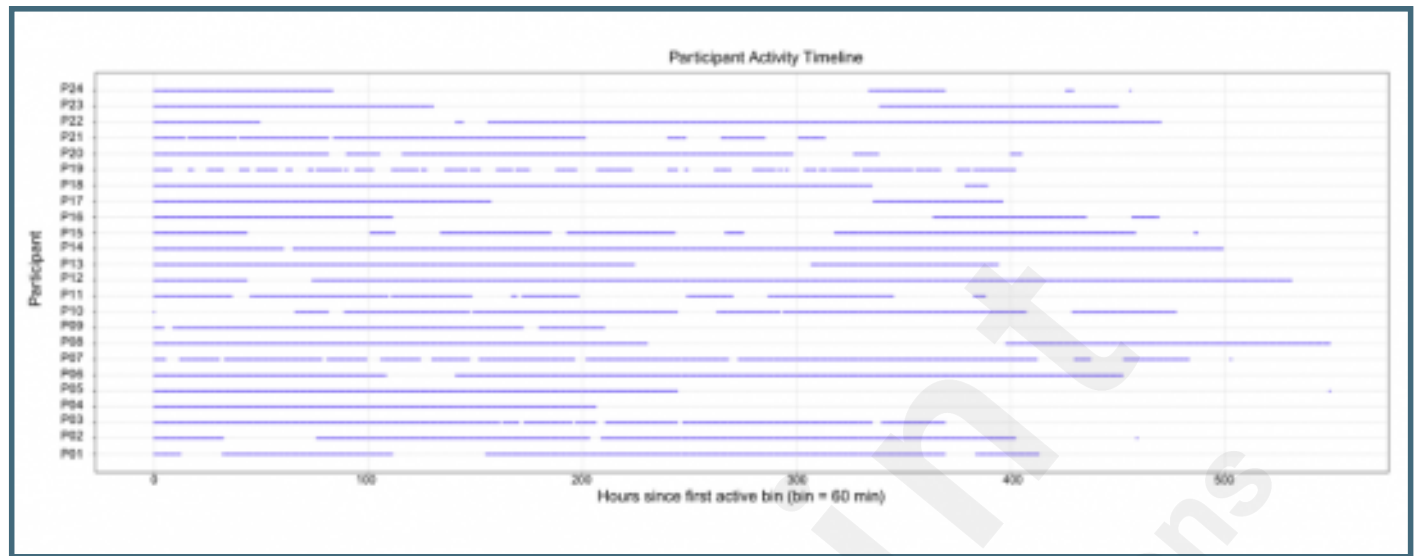
CONSORT-style participant flow diagram, showing progression from initial recruitment through study completion.



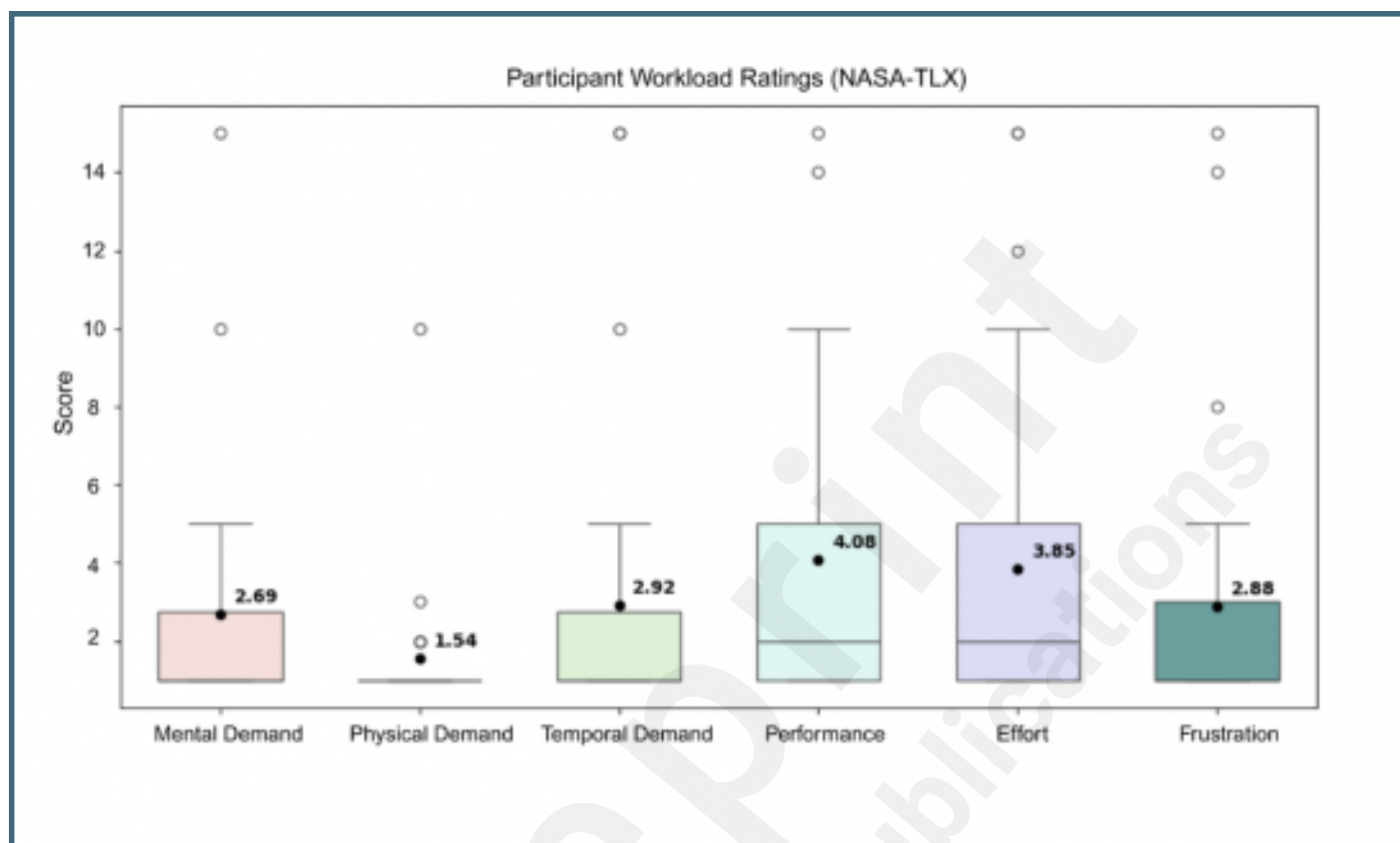
Calendar-aligned VPN continuity timeline. Each dot represents a 60-minute bin with VPN traffic, ordered by enrollment date.



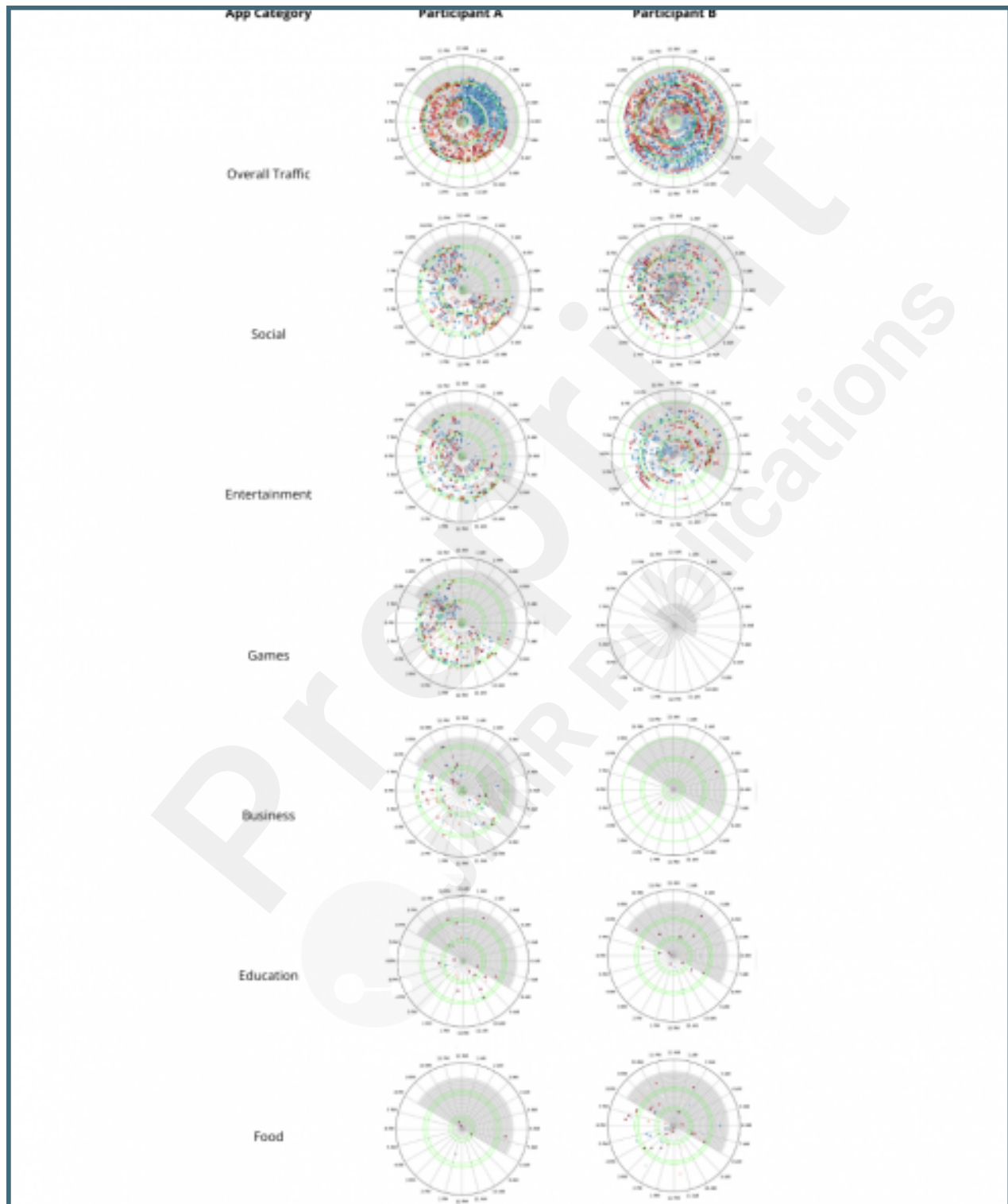
Day-0-aligned VPN continuity timeline. Each dot represents a 60-minute bin with VPN traffic, normalized to participants' first active bin.



Box plot of NASA-TLX scores from the exit survey (n=26), illustrating participants' perceived workload associated with maintaining the VPN-based monitoring system in daily life. Scores are reported on a 1–21 scale (higher = greater workload).



Temporal dynamics of overall and category-specific upload traffic for two representative participants. Each radial heatmap shows 24 hours (angular axis) across successive days (concentric rings). Color indicates normalized traffic intensity, with red denoting higher device usage and blue denoting lower usage. Nighttime hours (20:00–08:00) and weekends (green rings) are marked to facilitate alignment with social schedules.



Daily upload traffic from a researcher's device on August 22, 2025. Background periods (e.g., sleep and work windows) are shaded, with bursts of foreground activity visible as spikes in upload volume.

