

Permission granted?

Overcoming the human and the technical challenges in unlocking digital lifestyles with the Octopus app

Bence Ságvári

Bence Kollányi, Attila Gulyás, Krisztián Kókai, László Sándor

Research → App

App → Research

Functionality

"Reusable app" idea

Core infrastructure with modular design

Mixed method research (sensor/log+quantitative+qualitative)

ESM / EMA / Time and location triggering (geofencing)

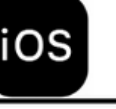
Designed to work offline (if needed)

Real time two-way communication (websocket + webdav)

(Almost) real-time access to data

Branded and/or customizable UI (tailored to specific target groups of participants)



Module	Function	OS
Passive modules		
App usage module	Collects app usage data at the event level (which app, when, and how long it was used)	
Screen time module	Tracks phone usage based on screen activation and inactivation events	
Call log module	Collects anonymized call metadata (caller/called, timestamp, duration)	
Messages module	Collects anonymized SMS metadata (sender/recipient, timestamp)	
Location module	Tracks movement patterns and daily mobility (configurable sampling rate and spatial accuracy)	 
Active modules		
Questionnaire module	Surveys with skip logic and conditional display	 
Camera/image	Capture and upload images from the device	 
Video	Record or upload video files from the device	 
Audio	Record or upload audio notes from the device	 
EMA (Ecological Momentary Assessment) module*	Delivers short questionnaires triggered by conditions or randomly, can include qualitative inputs	 



A TikTok és a YouTube a tinédzserek és a fiatal felnőttek elsőszámú mobil platformjai

Egy nap mintegy négy órát töltenek telefonjuk képernyője előtt a 16-35 évesek, ennek csaknem felét a közösségi médiumok alkalmazásainak használata teszi ki – egyebek mellett ez derült ki egy friss kutatásból,...

onlineplatformok.hu

1

New social stratification models using digital trace data

$n_{\text{app}} = 508$ ($n_{\text{net}} = 1801$)

Sample: 16+ adult population (representative sample from an online panel)

Type of data collected:

- app usage
- location (continuous)
- call ID
- SMS ID

Android only
min. 14 days of data
Ápril - May, 2024

2

Social media use and mental health among adolescence and young people

$n_{\text{app}} = 235$ ($n_{\text{net}} = 1114$)

Sample: 16-35 young adults (representative sample from an online panel)

Type of data collected:

- app usage
- location (continuous)

Android only
avg. ~17 days of data (min: 5, max:41)
June - July, 2025



Research questions

RQ1 - Types of biases:

At what stage, and why do people quit the research?

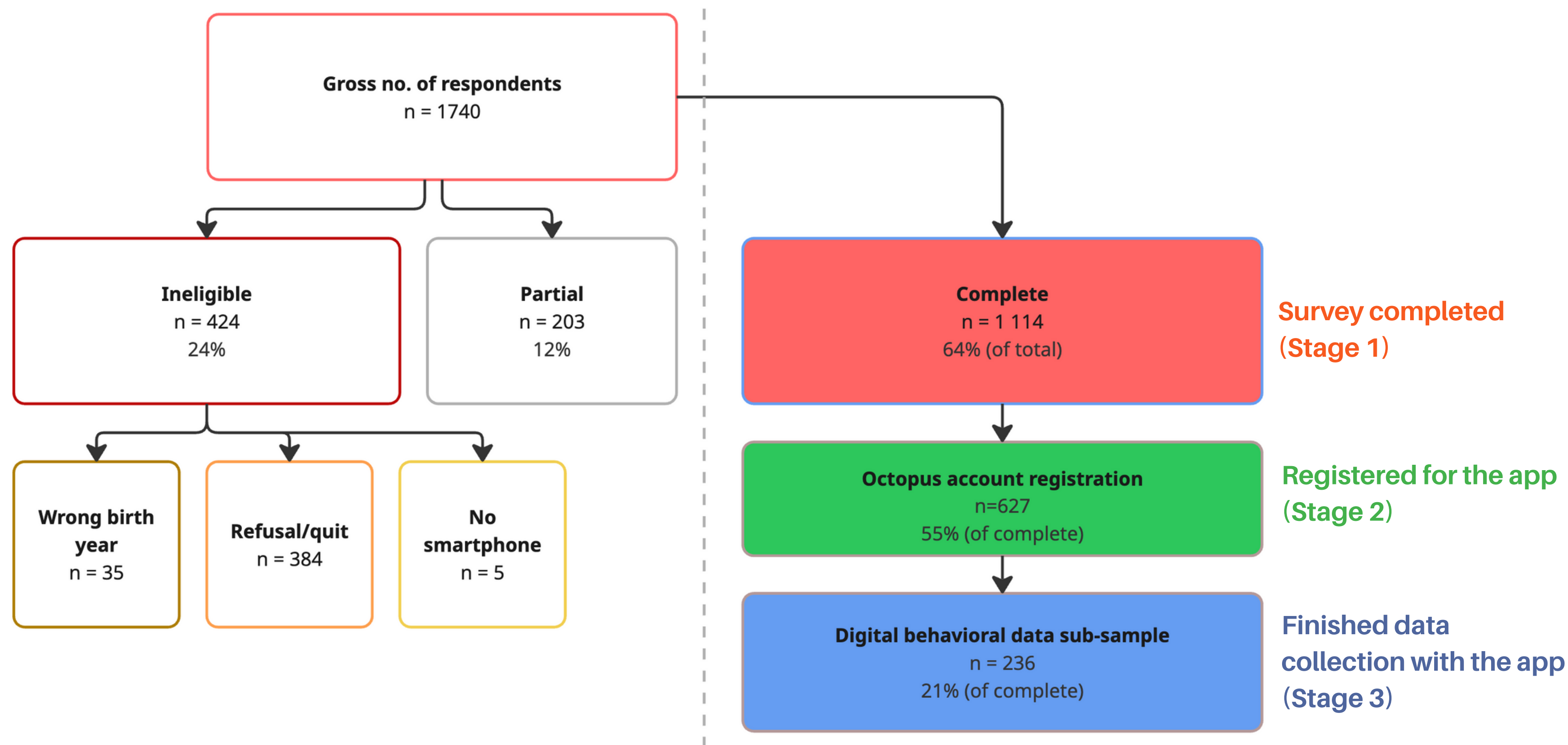
RQ2 - Human factors:

Who are “lost” and how does it change the sample composition?

RQ3 - Technical factors:

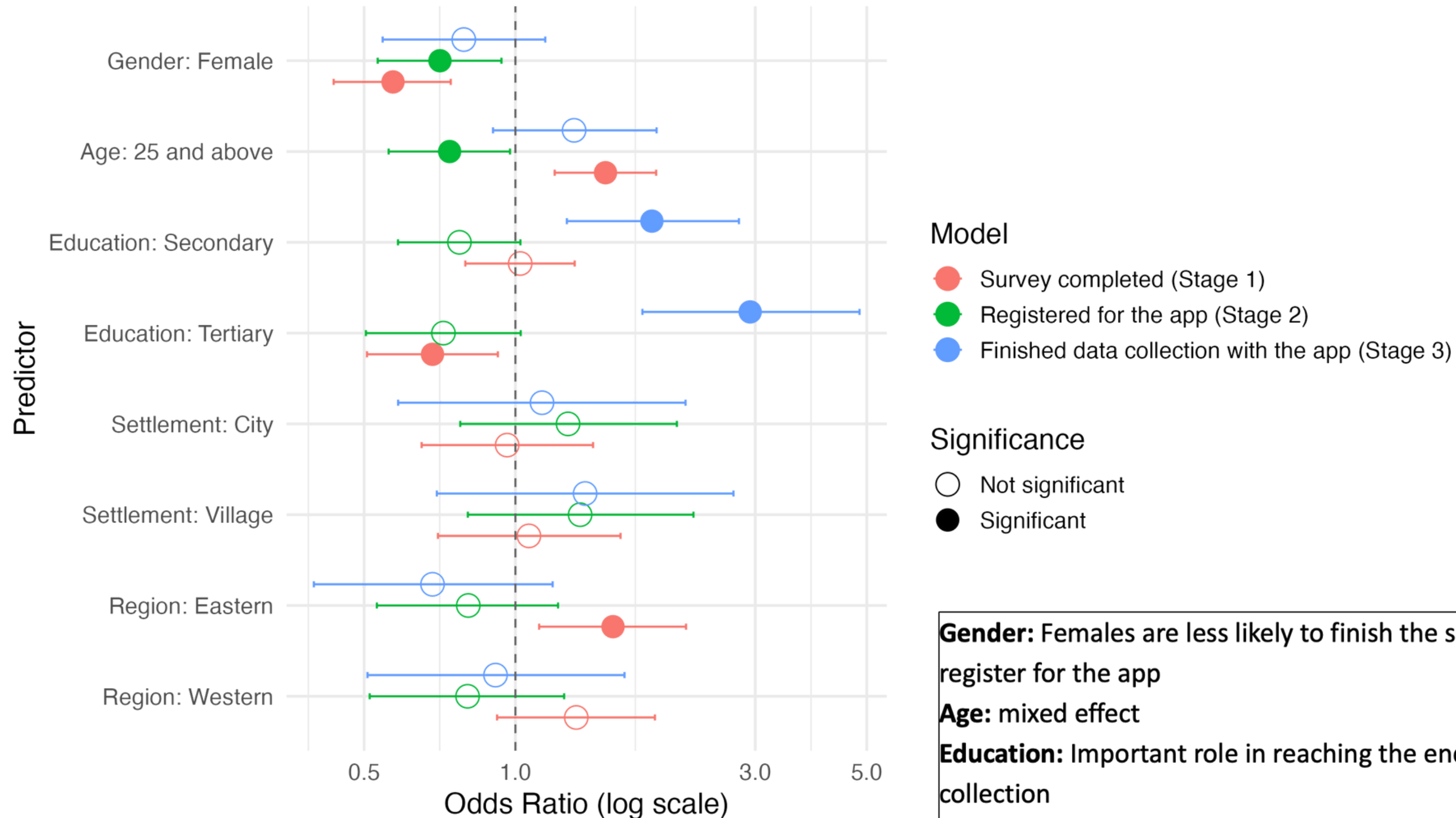
What are the major “under the hood” technical challenges that can have an impact on data quality and sample distribution?

Data collection process (RQ1)



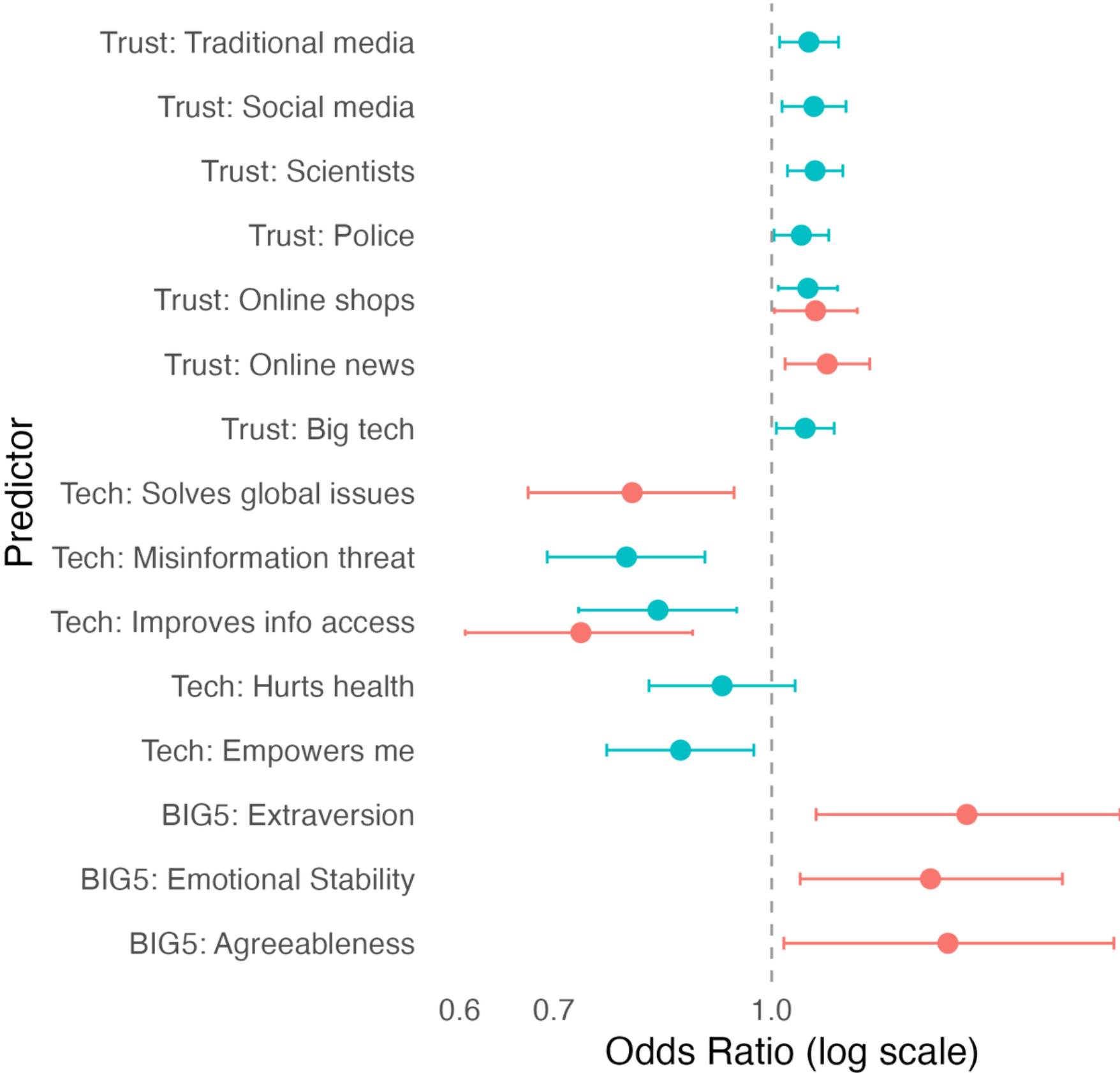
Data collection process - churn analysis: **BASIC DEMOGRAPHY (RQ2)**

Results of logistic regression models for comparing each stage to the previous one



Data collection process - churn analysis: “SOFT” FACTORS (RQ3)

Results of logistic regression models for comparing each stage to the previous one



Analysis of 25 indicators of trust, personality traits and attitudes towards technology. Only significant predictors are shown.

Model

- Finished data collection with the app (Stage 3)
- Registered for the app (Stage 2)

Trust: higher levels of (institutional) trust increase the likelihood of finishing the data collection

Technology attitudes: complex, but one-directional role in decreasing level of participation

Technical challenges and solutions (?) (RQ3)

Role of visible and invisible “device bias”

Background service stopped

(randomly... on specific models... at specific times....)

Permission revoked by the user (accidentally or intentionally)

Permission revoked by the device

(A total of 5 to 8 permission requests needed to be approved, depending on device and OS version.)

Data collection process - churn analysis: BASIC DEMOGRAPHY

Results of logistic regression models for comparing each stage to the previous one

It is far from self-explanatory how we define and calculate screen time data. (Errors in the app usage logs, treatment of system apps, launchers, etc.)

RAW DATA (SAMPLE)

packagename	eventtype	savedtime
com.readdle.spark	Activity_Resume	2024-07-05 08:59:06
com.readdle.spark	Activity_Pause	2024-07-05 08:59:10
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 08:59:10
com.teslacoilsw.launcher	Activity_Pause	2024-07-05 08:59:13
com.facebook.katana	Activity_Resume	2024-07-05 08:59:13
com.teslacoilsw.launcher	Activity_Stop	2024-07-05 08:59:14
com.facebook.katana	Activity_Pause	2024-07-05 09:02:36
com.facebook.katana	Activity_Resume	2024-07-05 09:02:36
com.facebook.katana	Activity_Pause	2024-07-05 09:02:42
com.facebook.katana	Activity_Resume	2024-07-05 09:02:42
com.facebook.katana	Activity_Pause	2024-07-05 09:02:43
com.facebook.katana	Activity_Resume	2024-07-05 09:02:43
com.facebook.katana	Activity_Stop	2024-07-05 09:02:44
com.facebook.katana	Activity_Pause	2024-07-05 09:03:04
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 09:03:04
com.teslacoilsw.launcher	Activity_Pause	2024-07-05 09:03:05
com.brave.browser	Activity_Resume	2024-07-05 09:03:05
com.brave.browser	Activity_Pause	2024-07-05 09:03:54
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 09:03:54
com.teslacoilsw.launcher	Activity_Pause	2024-07-05 09:03:55
com.readdle.spark	Activity_Resume	2024-07-05 09:03:55
com.readdle.spark	Activity_Pause	2024-07-05 09:03:57
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 09:03:57
com.readdle.spark	Activity_Stop	2024-07-05 09:03:58
com.facebook.orca	User_Interaction	2024-07-05 09:04:03
com.teslacoilsw.launcher	Activity_Pause	2024-07-05 09:04:03
com.facebook.orca	Activity_Resume	2024-07-05 09:04:03
com.facebook.orca	Activity_Resume	2024-07-05 09:04:03
com.teslacoilsw.launcher	Activity_Stop	2024-07-05 09:04:04
com.facebook.orca	Activity_Pause	2024-07-05 09:04:09
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 09:04:09
com.teslacoilsw.launcher	Activity_Pause	2024-07-05 09:04:10
com.lastpass.lpandroid	Activity_Resume	2024-07-05 09:04:10
com.lastpass.lpandroid	Activity_Resume	2024-07-05 09:04:10
com.lastpass.lpandroid	Activity_Stop	2024-07-05 09:04:11
com.lastpass.lpandroid	Activity_Pause	2024-07-05 09:04:12
com.teslacoilsw.launcher	Activity_Resume	2024-07-05 09:04:12

PRE-PROCESSED DATA (SAMPLE)

appname	packagename	time_start	time_end	duration_sec	duration_hms
Spark Mail – AI Email Inbox	com.readdle.spark	2024-07-05 08:59:06	2024-07-05 08:59:10	97	00:01:37
Nova Launcher	com.teslacoilsw.launcher	2024-07-05 08:59:10	2024-07-05 08:59:13	3	00:00:03
Facebook	com.facebook.katana	2024-07-05 08:59:13	2024-07-05 09:03:04	231	00:03:51
Brave Private Web Browser, VPN	com.brave.browser	2024-07-05 09:03:05	2024-07-05 09:03:54	49	00:00:49
Spark Mail – AI Email Inbox	com.readdle.spark	2024-07-05 09:03:55	2024-07-05 09:03:57	2	00:00:02
Nova Launcher	com.teslacoilsw.launcher	2024-07-05 09:03:57	2024-07-05 09:04:03	6	00:00:06
Messenger	com.facebook.orca	2024-07-05 09:04:03	2024-07-05 09:04:09	6	00:00:06
LastPass Password Manager	com.lastpass.lpandroid	2024-07-05 09:04:10	2024-07-05 09:04:12	2	00:00:02

CLEANED DATA (SAMPLE)

Launcher and system apps removed

appname	packagename	time_start	time_end	duration_sec	duration_hms
Spark Mail – AI Email Inbox	com.readdle.spark	2024-07-05 08:59:06	2024-07-05 08:59:10	97	00:01:37
Facebook	com.facebook.katana	2024-07-05 08:59:13	2024-07-05 09:03:04	231	00:03:51
Brave Private Web Browser, VPN	com.brave.browser	2024-07-05 09:03:05	2024-07-05 09:03:54	49	00:00:49
Spark Mail – AI Email Inbox	com.readdle.spark	2024-07-05 09:03:55	2024-07-05 09:03:57	2	00:00:02
Messenger	com.facebook.orca	2024-07-05 09:04:03	2024-07-05 09:04:09	6	00:00:06
LastPass Password Manager	com.lastpass.lpandroid	2024-07-05 09:04:10	2024-07-05 09:04:12	2	00:00:02

Detailed log analysis of data collection errors

Device specific patterns

has.error			has.error			has.error			login_class			
Brand	False	True	Android	False	True	Model	False	True	Grand T..	Brand	one	many
Samsung	61.30%	38.70%	14	64.36%	35.64%	Galaxy A52	66.67%	33.33%	100.00%	Honor	15.79%	84.21%
Redmi	62.50%	37.50%	13	67.55%	32.45%	Galaxy A54 5G	63.64%	36.36%	100.00%		3	16
Xiaomi	65.38%	34.62%	12	59.65%	40.35%	Galaxy A53 5G	55.56%	44.44%	100.00%		Huawei	18.18%
Huawei	31.82%	68.18%	11	50.79%	49.21%	Galaxy A13	50.00%	50.00%	100.00%	4		18
Honor	73.68%	26.32%	10	56.10%	43.90%	Galaxy A52s 5G	68.75%	31.25%	100.00%	Sony		25.00%
POCO	66.67%	33.33%	9	25.00%	75.00%	Galaxy A12	56.25%	43.75%	100.00%	Motorola	25.00%	75.00%
Motorola	62.50%	37.50%	7.1.1	33.33%	66.67%	Redmi Note 12	66.67%	33.33%	100.00%		2	6
Deutsche Telekom	57.14%	42.86%	8.0.0	33.33%	66.67%	Redmi Note 10 Pro	50.00%	50.00%	100.00%		POCO	33.33%
Sony	25.00%	75.00%	8.1.0	33.33%	66.67%	Galaxy S23 Ultra	70.00%	30.00%	100.00%	3		6
LGE	75.00%	25.00%				Galaxy A34 5G	90.00%	10.00%	100.00%	Redmi		34.38%
Lenovo	50.00%	50.00%				Galaxy A23 5G	60.00%	40.00%	100.00%		33	63
Oppo	66.67%	33.33%				Redmi Note 11 Pro 5G	100.00%		100.00%		Xiaomi	38.46%
OnePlus	66.67%	33.33%				Galaxy A33 5G	88.89%	11.11%	100.00%	10		16
ZTE	50.00%	50.00%				Redmi Note 12 Pro 5G	75.00%	25.00%	100.00%	Deutsche Telekom		42.86%
Vivo	100.00%					Redmi Note 11	75.00%	25.00%	100.00%		3	4
Ulefone	100.00%					Galaxy S21 FE 5G	75.00%	25.00%	100.00%		Samsung	63.47%
realme		100.00%				Galaxy A51	62.50%	37.50%	100.00%	205		118
Nothing	100.00%					Galaxy A14	50.00%	50.00%	100.00%	Oppo		66.67%
Nokia	100.00%					Redmi 9A	57.14%	42.86%	100.00%		2	1
Google	100.00%					HONOR 90	71.43%	28.57%	100.00%		OnePlus	66.67%
Gigaset	100.00%					Galaxy A20e		100.00%	100.00%	2		1
Blaupunkt		100.00%				P30 Pro	50.00%	50.00%	100.00%	Lenovo		75.00%
Blackview		100.00%				HUAWEI P30 lite	33.33%	66.67%	100.00%		3	1
						Galaxy S20 FE	50.00%	50.00%	100.00%			
						Galaxy A40	33.33%	66.67%	100.00%			
						Galaxy A32 5G	66.67%	33.33%	100.00%			
						Xiaomi 11 Lite 5G NE	20.00%	80.00%	100.00%			
						Redmi Note 8T		100.00%	100.00%			
						Redmi 9T	60.00%	40.00%	100.00%			
						Galaxy S22	60.00%	40.00%	100.00%			
						Galaxy S10e	60.00%	40.00%	100.00%			
						Galaxy A71	60.00%	40.00%	100.00%			
						Galaxy A52 5G	80.00%	20.00%	100.00%			
						Galaxy A32	40.00%	60.00%	100.00%			
						Xiaomi 13T	75.00%	25.00%	100.00%			

Conclusions

Patterns in willingness to participate, actual participation levels, and churn stages show similarity across different studies.

Notable attrition at each stage from survey to full digital participation.

- Demographic factors such as gender and education significantly shaped participation across stages, with dropout patterns varying by subgroup.
 - High level of institutional trust consistently supported continued participation, while certain technology attitudes played a more ambivalent, often limiting role.
-

Less visible technology divide

- iOS vs Android (we all know about)
- Device (model) specific problems à biases
- (Some) differences in the raw log data (app usage)

Solution: thorough testing, brand-specific tutorials and manuals, constant logging, automated emails to respondents to ask for action, benchmarking raw data with built-in and commercial tools (i.e. screen time apps)



Thank You!

Website:

www.octopus-research.hu

Website:

www.tk.hu

Email Address:

sagvari.bence@tk.hu