SoftwareX 34 (2026) 102706

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

# *EasyTrack*: A scalable and general purpose platform for reliable data collection in mHealth studies

Alfred Malengo Kondoro [a], Kobiljon Toshnazarov [b], Muhammad Salman [c], Donggeun Oh [d], Yugyeong Jung [e], Uichin Lee [e], Youngtae Noh [a,*]

[a] Department of Data Science, Hanyang University, Seoul, Republic of Korea
[b] School of Computing, Department of Computer Science, New Uzbekistan University, Tashkent, Republic of Uzbekistan
[c] Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Khyber Pakhtunkhwa, Pakistan
[d] Department of Artificial Intelligence, Hanyang University, Seoul, Republic of Korea
[e] School of Computing, Korea Advanced Institute of Science & Technology (KAIST), Daejeon, Republic of Korea

ARTICLE INFO

*Keywords:*
mHealth data collection
Scalable research platforms
Data quality monitoring
Longitudinal studies

ABSTRACT

Mobile health (mHealth) studies increasingly rely on continuous, in-the-wild data from smartphones and wearable sensors, yet existing platforms often fall short in scalability, data quality assurance, and interoperability, limiting their ability to support reliable longitudinal research under real-world conditions. *EasyTrack* addresses these challenges by providing a cloud-based, general-purpose platform that supports heterogeneous sensing configurations, automated data quality monitoring, and integration with external data collection systems. The platform is implemented through a modular, layered architecture that enables efficient data ingestion, configurable study management, and real-time data quality inspection. *EasyTrack* has been deployed in multiple real-world mHealth studies involving diverse participant cohorts and sensor modalities, and cloud-based scalability experiments further demonstrate its capacity to support large-scale deployments. Together, these empirical evaluations show that *EasyTrack* provides a robust and extensible infrastructure for conducting high-quality longitudinal mHealth research.

## Code metadata

Code metadata (mandatory).

| Nr. | Code metadata description | Metadata |
|---|---|---|
| C1 | Current code version | v1.0 |
| C2 | Permanent link to code/repository used for this code version | https://github.com/imc-hanyang/et-scalability-test |
| C3 | Permanent link to Reproducible Capsule | https://hub.docker.com/r/roganoh/et-dashboard https://hub.docker.com/r/roganoh/et-grpc-server |
| C4 | Legal Code License | MIT License |
| C5 | Code versioning system used | GIT |
| C6 | Software code languages, tools, and services used | Python, Django |
| C7 | Compilation requirements, operating environments & dependencies | (included in repository) |
| C8 | If available Link to developer documentation/manual | https://github.com/imc-hanyang/eva-dashboard/blob/main/README.md |
| C9 | Support email for questions | qobiljon.toshnazarov@gmail.com |

* Corresponding author.
*Email addresses:* alfr3do@hanyang.ac.kr (A.M. Kondoro), q.toshnazarov@newuu.uz (K. Toshnazarov), msalman@ceme.nust.edu.pk (M. Salman), dgoh1620@hanyang.ac.kr (D. Oh), yugyeong.jung@kaist.ac.kr (Y. Jung), uclee@kaist.ac.kr (U. Lee), youngtaenoh@hanyang.ac.kr (Y. Noh).

https://doi.org/10.1016/j.softx.2026.102706
Received 27 December 2025; Received in revised form 27 April 2026; Accepted 4 May 2026
Available online 19 May 2026

## 1. Motivation and significance

Mobile health (mHealth) research increasingly relies on smartphones and wearable devices to collect continuous, in-the-wild data about human behavior, context, and physiology [1]. Such data underpin a wide range of scientific investigations, including behavioral monitoring, mental health assessment, and digital phenotyping [2–4]. To support such investigations, researchers conduct studies that vary widely in scale and duration, ranging from short-term experiments [5] to longitudinal deployments that combine passive sensing with self-reported data [3,6,7]. Despite this progress, a fundamental scientific problem remains unresolved: *ensuring reliable and high-quality longitudinal mHealth data collection under real-world conditions.* As studies scale beyond controlled environments, data streams become increasingly prone to degradation caused by participant non-compliance [8,9], sensor and operating-system limitations [10,11], software faults [9,12], and constrained device resources such as battery life [11,13]. These issues often emerge gradually and, without systematic monitoring, may remain unnoticed, resulting in incomplete, irregular, or anomalous data that undermine the validity, reproducibility, and interpretability of scientific findings. Addressing this problem requires software infrastructure that extends beyond basic sensing and storage to enable scalable, flexible, and continuously monitored longitudinal data collection. We therefore developed *EasyTrack*, a general-purpose platform for reliable real-world mHealth research.

*EasyTrack* has already supported multiple peer-reviewed scientific studies, demonstrating its contribution to longitudinal mHealth research. The platform has enabled passive-sensing investigations of user engagement and dropout by supporting the collection of behavioral and contextual data used to model participation dynamics [14]. It has also been applied in large-scale digital phenotyping studies that leverage smartphone sensing data to predict real-world behaviors, such as departure-time estimation for extending electric vehicle battery lifetimes [7,15]. In the mental health domain, *EasyTrack* has facilitated large-scale studies on short-term depression detection [16,17] and in-the-wild stress sensing using commodity smartwatches [18,19].

Researchers use *EasyTrack* through a web-based interface to configure studies, participants, and data sources, while participant devices collect and transmit passive sensing and self-reported data via lightweight client applications. Together, these studies show how *EasyTrack* enables sustained real-world deployments and high-quality longitudinal data collection.

Prior work on mHealth data collection spans platform development, scalable system architectures, and approaches for managing longitudinal data quality. Several platforms provide end-to-end sensing and monitoring capabilities using smartphones and wearable devices, including RADAR-Base [20], BEIWE [21], the AWARE framework [22], AwarNS [23], CARP Mobile Sensing (CAMS) [23], and mCerebrum [24]. These systems demonstrate important advances in ecological momentary assessment (EMA), passive sensing, and digital phenotyping using commodity devices, but they differ in their support for scalable deployment, extensible study configuration, interoperability with external systems, and integrated monitoring of longitudinal data quality. Other infrastructures provide vendor-independent monitoring for digital health studies [25] and clinical digital phenotyping platforms such as mindLAMP [26]. More recent efforts emphasize monitoring and evaluation through dashboards and big-data infrastructures [27], as well as modular integration of mHealth, wearable, and IoT data streams [28]. In parallel, research on scalable architectures has explored cloud- and serverless-based pipelines, real-time analytics, and large-scale data ingestion for in-the-wild applications [29,30], with general-purpose frameworks such as Sensus demonstrating scalable support for human-subject studies [31]. Methodological studies further highlight the importance of data quality monitoring in longitudinal and digital phenotyping research [32–34].

Interoperability standards such as HL7 FHIR increasingly guide how patient-generated health data are exchanged, but most existing mHealth platforms provide only limited support for mapping heterogeneous sensing and EMA data to such formats. Despite these advances, existing solutions typically address these challenges in isolation. As a result, this work contributes *EasyTrack*, an integrated, research-oriented platform that unifies scalable system architecture, flexible study configuration, and systematic data quality management for longitudinal mHealth studies. A detailed comparative analysis of related work is provided in Section S1 of the supplementary material.

In summary, this paper makes the following contributions:

- A general-purpose research infrastructure that operationalizes reliable longitudinal mHealth data collection under real-world conditions, addressing systematic data degradation caused by participant behavior, device constraints, and software failures.
- A unified architectural model that integrates scalable data ingestion, flexible study configuration, and continuous data quality monitoring within a single platform, bridging capabilities that prior mHealth systems address only in isolation.
- An empirically validated system design demonstrated across multiple real-world deployments and large-scale scalability experiments, providing evidence that high-frequency, multimodal sensing can be reliably supported at scale.
- A reusable and extensible foundation for future mHealth research, enabling reproducible study management, cross-study interoperability, and systematic evaluation of data quality throughout longitudinal deployments.

## 2. Software description

In this section, we explain the software architecture and functionality of *EasyTrack*, which are detailed as follows:

### 2.1. Software architecture

The software architecture of *EasyTrack* follows a four-layer design that reflects the lifecycle of longitudinal mHealth studies, from study configuration to data collection and quality monitoring. As illustrated in Fig. 1, the management layer provides researcher-facing interfaces for study and participant configuration; the data collection layer acquires and standardizes EMA and sensor data on participant devices; and the core layer exposes back-end services for scalable data ingestion, storage, and retrieval. On top of these components, the data quality (DQ) monitoring layer continuously analyzes incoming data streams to identify missingness, irregular sampling, and anomalous patterns, enabling real-time feedback through dashboards and automated notifications. Together, these layers form an integrated pipeline that supports reliable, end-to-end longitudinal data collection under real-world conditions. Each layer is described in detail in the following subsections.

In addition, *EasyTrack* supports GDPR-aligned data protection practices by implementing client-side pseudonymization, encrypted data transport, and role-based researcher access essential safeguards when EMA responses contain sensitive health-related information.

#### 2.1.1. Management layer

This layer provides the web-based interface through which researchers configure and manage data collection studies. Through this portal, researchers can create studies, define data sources, manage participants, and collaborate with other researchers. Authentication is handled through Google OAuth 2.0, enabling seamless login and automatic registration on first use. Once authenticated, users can initiate new studies or modify existing ones.

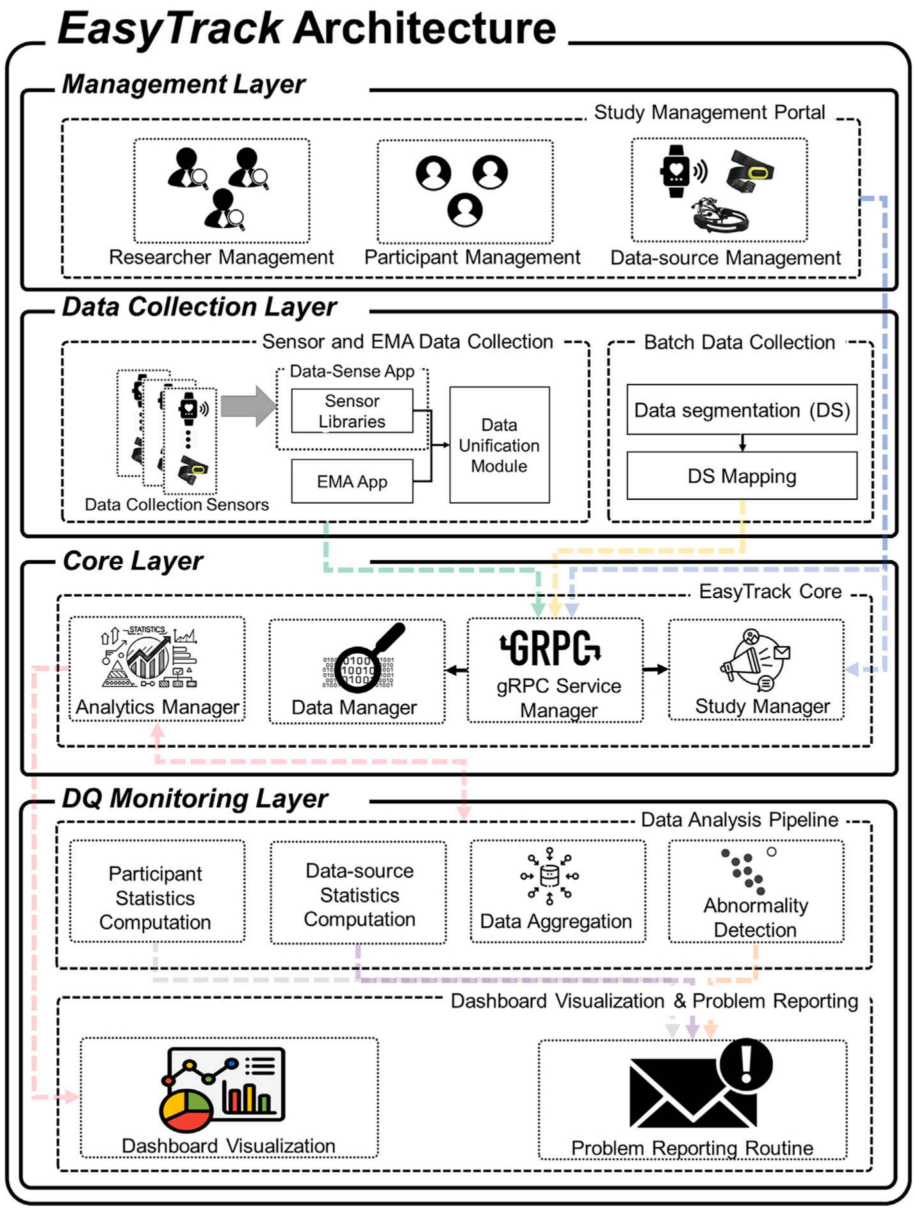


**Fig. 1.** Four-Layer Software Architecture of the *EasyTrack* platform.

#### 2.1.2. Data collection layer

This layer is responsible for acquiring data from participants and structuring it before transmission to the back-end for storage. It supports two operational modes corresponding to typical data types encountered in mHealth studies:

- Small-to-medium size data (typically up to ∼1 MB): real-time streams from smartphones and wearable sensors, such as accelerometers, GPS, physiological signals, and EMA responses.
- Large-size data (typically >1 MB): high-volume modalities including audio recordings, medical imaging, and video files.

The distinction reflects the ingestion strategy used by the platform rather than strict size boundaries; intermediate data sizes (e.g., hundreds of kilobytes) are handled by the same streaming ingestion pipeline used for sensor and EMA data. These two modes allow the platform to support both lightweight sensor streams and high-throughput media uploads in longitudinal mHealth deployments.

#### 2.1.3. Core layer

The core layer provides the platform's main processing logic, exposing functionality through APIs and managing database operations. It uses two database systems: a wide-column store (Apache Cassandra) for large-scale time-series data, and a relational database (e.g., PostgreSQL) for metadata and study configurations. Incoming operations are routed to the appropriate storage system based on workload characteristics. The layer comprises four components: the gRPC service manager,[1] study manager, data manager, and analytics manager. Additional implementation details are provided in Section S4.

The platform exposes its data ingestion and management functionality through gRPC-based APIs defined using Protocol Buffers. The corresponding interface definitions (`.proto` files) are available in the project repository and can be used to generate client libraries for different programming languages. This design enables EasyTrack to remain agnostic to specific data collection applications, allowing developers

[1] https://www.toptal.com/developers/grpc/grpc-vs-rest-api.

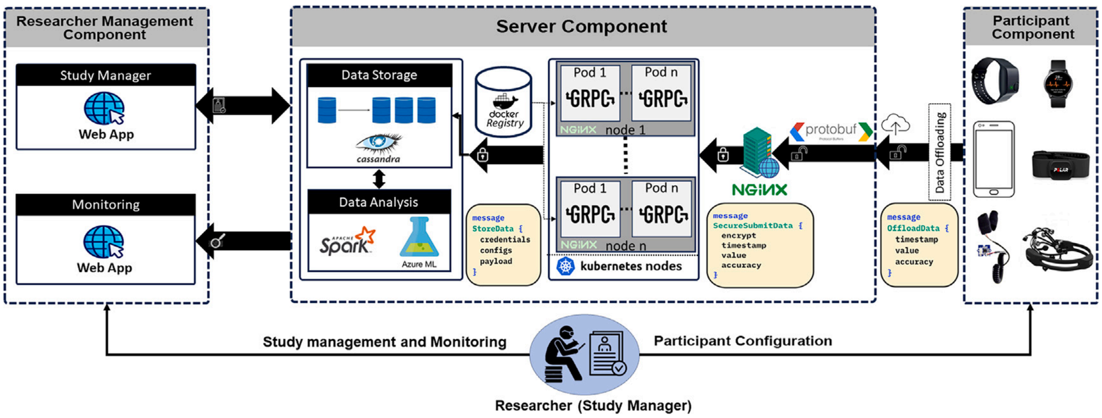


**Fig. 2.** *EasyTrack* functionality with participant, server and research management components.

to implement custom sensing clients that communicate with the server using the provided API definitions.

### 2.1.4. Data quality monitoring layer

The data quality (DQ) monitoring layer consists of a data analysis pipeline and an output component that includes dashboard visualizations and automated issue reporting. The analysis pipeline processes incoming data and provides summary outputs to the analytics manager, while the dashboard enables researchers to inspect data completeness and trends. A problem-reporting routine detects missing or anomalous data [11,35] and notifies researchers when intervention is needed. Further details appear in Section S5.

## 2.2. Software deployment

*EasyTrack* is organized into three components that reflect the roles within an mHealth study workflow: (1) a participant component running on user devices to collect and transmit sensor and EMA data, (2) a server component deployed in the cloud to handle data ingestion, processing, storage, and data quality checks, and (3) a researcher management component that provides a web-based interface for study configuration, participant management, and data quality visualization.

The interaction of these three components is illustrated in Fig. 2, and each component is described in detail below.

### 2.2.1. Participant component

This component includes the participant-facing applications used for data acquisition and upload within *EasyTrack*. The data-sense application collects sensor streams and uploads them to the server component; EMA applications gather self-reported information for specific study designs; and the WearOS application streams smartwatch data. Together, these tools enable the reliable transfer of heterogeneous data from participant devices to the backend infrastructure.

The participant applications currently support Android-based smartphones and WearOS smartwatches. These applications collect multimodal sensing streams from smartphone and wearable sensors, including inertial measurements (e.g., accelerometer), location data (GPS), device context signals, and wearable physiological measurements such as heart rate and activity data. In addition, ecological momentary assessment (EMA) responses entered by participants are collected through mobile questionnaires. Data are temporarily stored on the device and transmitted to the EasyTrack backend through gRPC-based APIs for secure and efficient data transfer. The platform has been used in prior mHealth research, including a longitudinal passive sensing study for depression severity prediction [17].

Implementation details, installation instructions, and tutorials for the participant applications are provided through their respective project repositories and documentation pages.

### 2.2.2. Server component

For scalability, *EasyTrack* is deployed in cloud-based environments. The platform is designed to support long-running longitudinal studies involving heterogeneous sensing streams from smartphones and wearable devices, which generate high-frequency multimodal data and variable ingestion workloads. To ensure reliable data ingestion, fault isolation, and flexible scaling across deployments, the backend adopts a containerized microservices architecture.

To efficiently stream large volumes of data, we employ Google Protocol Buffers[2] and secure communication through NGINX configured as a bastion reverse-proxy server.[3] Scalability is achieved through a microservices architecture using Docker[4] with horizontal autoscaling via Kubernetes,[5] initially deployed with three nodes that expand dynamically based on workload. The Kubernetes load balancer (kube-proxy)[6] routes incoming requests to microservice pods running gRPC-based modules for authentication, study management, and data transfer. Data are stored using multiple systems depending on workload characteristics: Apache Cassandra[7] for horizontally scalable time-series sensor data, relational databases for metadata and configuration requiring strong consistency,[8] and Apache Parquet[9] for optimized columnar storage. For large-scale analytics, the platform integrates Apache Spark for ETL and machine learning pipelines and Azure Synapse Analytics[10] for managed data warehousing, with additional implementation details provided in Section S6.

### 2.2.3. Researcher management component

This component provides a user interface for researchers to configure studies and monitor data collection. It consists of two modules: (1) a **Study Management Portal**, implemented using Gunicorn as a Web Server Gateway Interface (WSGI) application with Azure RDB for data retrieval and persistence, and (2) a **Data Quality Dashboard**, built with Node.js and the PM2 process manager,[11] which provides real-time visualization of data completeness and collection status across participants and sensors.

[2] https://protobuf.dev/.
[3] https://www.nginx.com/.
[4] https://www.docker.com/.
[5] https://kubernetes.io/.
[6] https://kubernetes.io/docs/concepts/services-networking/service/.
[7] https://cassandra.apache.org/.
[8] https://www.ibm.com/topics/relational-databases.
[9] https://parquet.apache.org/.
[10] https://azure.microsoft.com/products/synapse-analytics.
[11] https://pm2.keymetrics.io/.


(a) Campaign configuration

(b) Schema management

(c) Dashboard: intelligent visualization of data collection with problem detection.

**Fig. 3.** (a) Campaign configuration user interface, where the researcher can name their campaign, specify the start and end dates for data collection, and select the parameters. (b) Schema management user interface, where the researcher can define new data sources. (c) Web-based dashboard for analyzing data collection from a single participant. The Y-axes of the bar charts represent the hourly amounts of data collected from a particular data source, and X-axis is the time.

The dashboard is intended for monitoring and inspection of incoming data streams and does not perform analytical processing or knowledge extraction from the collected data.

## 3. Illustrative example

We demonstrate the practical use of *EasyTrack* through real-world longitudinal mHealth deployments that required sustained, multimodal data collection under unconstrained, in-the-wild conditions. These studies involved heterogeneous data sources, including smartphone sensors, wearable devices, and ecological momentary assessments (EMA), and posed challenges such as gradual data degradation, silent sensor failures, and participant non-compliance that often remain undetected in conventional pipelines.

Fig. 3 illustrates how *EasyTrack* supports these deployments, from study configuration to ongoing monitoring. Researchers define study parameters and sensing configurations through a unified management interface (Fig. 3a and b), which abstracts low-level schema and storage details while ensuring that all collected data conform to a consistent structure. This design enables rapid study setup and reuse across deployments with different cohorts and sensing modalities.

Once studies are active, *EasyTrack* provides continuous visibility into data collection behavior. As shown in Fig. 3(c), the dashboard aggregates longitudinal sensor and EMA streams at multiple temporal resolutions and highlights emerging data quality issues, such as missing data, irregular sampling, and non-compliance. These indicators allow researchers to distinguish participant-level behavior from system-level sensing failures and to intervene during deployment rather than after study completion.

Across multiple deployments, ranging from small pilot studies to cohorts exceeding 700 participants, this integrated workflow enabled consistent study management, scalable ingestion, and real-time data quality monitoring without architectural changes. Together, these deployments illustrate how *EasyTrack* functions as a shared research infrastructure that supports reproducible longitudinal mHealth studies by coupling scalable system design with continuous, deployment-time data quality assessment.

## 4. Impact

*EasyTrack* simplifies and improves large-scale mHealth data collection by coupling automated data quality (DQ) monitoring with interoperable data ingestion. This enables researchers to conduct robust longitudinal studies, such as analyzing behavioral dynamics with reduced data-cleaning overhead and immediate visibility into data gaps. Its interoperability further supports cross-platform integration, enabling secondary analyzes that combine datasets from heterogeneous sources. The platform has demonstrated practical value across four diverse deployments, including deployments for short-term depression detection (767 participants) [16], real-world stress detection (381 participants) [17], a longitudinal user-behavior study (300 participants) [7,14,15], and a college-student monitoring study (44 participants) [18]. Even studies with moderate cohort sizes can generate substantial ingestion workloads due to continuous multimodal sensing streams from smartphones and wearables, where individual participants may produce high-frequency sensor data over extended study durations. Across these studies, *EasyTrack* reduced manual troubleshooting through automated anomaly detection and intuitive visualization, while batch upload capabilities enabled seamless integration of externally collected datasets. These features facilitated faster study iterations and smoother scaling from pilot deployments to larger cohorts. Beyond academic research, *EasyTrack* is applicable to corporate wellness programs, remote patient monitoring, and large-scale behavioral analytics. Its primary impact lies in advancing reproducibility, transparency, and collaboration in longitudinal mHealth research. While the platform provides robust data ingestion, monitoring, and visualization, it currently relies on manual interpretation of complex longitudinal patterns and predefined rules for anomaly detection. These limitations motivate future extensions, such as integrating language-model–based assistants into the dashboard to support adaptive data summarization, anomaly explanation, and study-level decision support during ongoing deployments.

## 5. Conclusions

This paper presents *EasyTrack*, a research-oriented platform designed to enable reliable, scalable, and interoperable longitudinal data collection in real-world mHealth studies, where heterogeneous sensing configurations and long deployment periods frequently lead to data degradation. By integrating configurable study management, high-throughput data ingestion, and continuous data quality monitoring within a unified four-layer architecture, *EasyTrack* explicitly treats data quality as a first-class concern throughout the data collection lifecycle. Through real-world deployments across diverse participant populations and sensing modalities, together with systematic scalability evaluation, we demonstrate that the platform can sustain longitudinal sensing workloads while providing timely visibility into missingness, irregular sampling, and anomalous data patterns during ongoing studies. Collectively, these results show that *EasyTrack* moves beyond isolated sensing or storage solutions by offering an integrated research infrastructure that supports reproducible, large-scale mHealth research under real-world conditions.

## CRediT authorship contribution statement

**Alfred Malengo Kondoro:** Writing – review & editing, Writing – original draft, Software, Methodology, Conceptualization. **Kobiljon Toshnazarov:** Software, Resources, Methodology, Investigation. **Muhammad Salman:** Writing – review & editing, Visualization, Supervision. **Donggeun Oh:** Writing – review & editing, Writing – original draft, Software, Resources. **Yugyeong Jung:** Supervision. **Uichin Lee:** Supervision. **Youngtae Noh:** Supervision.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Youngtae Noh reports that financial support was provided by Hanyang University. Youngtae Noh reports a relationship with Hanyang University that includes employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. RS-2023-00236325), as well as the Ministry of Climate, Energy & Environment (MCEE) of the Republic of Korea (No. RS-2025-02654073).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at doi: 10.1016/j.softx.2026.102706.

## References

[1] Kumar D, Jeuris S, Bardram JE, Dragoni N. Mobile and wearable sensing frameworks for mhealth studies and applications: a systematic review. ACM Trans Comput Healthc 2020;2:1–28.
[2] Xu Z, Zhang J, Haroutounian S, Liu H, Cao Z, Messner GR, Alaverdyan HB, Ahuja S, Koshy R, Hanns J, Frumkin M, Rodebaugh TL, Lu C. Incorporating uncertainty in predictive models using mobile sensing and clinical data: a case study on persistent post-surgical pain. Proc ACM Interact Mob Wearable Ubiquitous Technol 2025;9. https://doi.org/10.1145/3729488
[3] Murnane EL, Glazko YS, Costa J, Yao R, Zhao G, Moya PML, Landay JA. Narrative-Based visual feedback to encourage sustained physical activity: a field trial of the WhoIsZuki mobile health platform. Proc ACM Interact Mob Wearable Ubiquitous Technol 2023;7. https://doi.org/10.1145/3580786
[4] Ikäheimonen A, Triana AM, Luong N, Ziaei A, Rantaharju J, Darst R, Aledavood T. Niimpy: a toolbox for behavioral data analysis. Softwarex 2023;23:101472. https://doi.org/10.1016/j.softx.2023.101472. https://www.sciencedirect.com/science/article/pii/S2352711023001681.
[5] Narziev N, Goh H, Toshnazarov K, Lee SA, Chung K-M, Noh Y. STDD: short-term depression detection with passive sensing. Sensors 2020;20:1396.
[6] Xu X, Liu X, Zhang H, Wang W, Nepal S, Sefidgar Y, Seo W, Kuehn KS, Huckins JF, Morris ME, et al. GLOBEM: Cross-Dataset generalization of longitudinal human behavior modeling. In: Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies, vol. 6; 2023. p. 1–34.
[7] Lee Y, Song W, Song J, Noh Y. Extending EV battery lifetime: digital phenotyping approach for departure time prediction. Proc ACM Interact Mob Wearable Ubiquitous Technol 2024;8. https://doi.org/10.1145/3699725
[8] Donckt JVD, Vandenbussche N, Donckt JVD, Chen S, Stojchevska M, Brouwer MD, Steenwinckel B, Paemeleire K, Ongenae F, Hoecke SV. Mitigating data quality challenges in ambulatory wrist-worn wearable monitoring through analytical and practical approaches. Sci Rep 2024;14. https://doi.org/10.1038/s41598-024-67767-3
[9] Widhalm D, Goeschka K, Kastner W. An Open-Source wireless sensor node platform with active Node-Level reliability for monitoring applications. Sensors (Basel, Switzerland) 2021;21. https://doi.org/10.3390/s21227613
[10] Alwan A, Ciupala M, Brimicombe A, Ghorashi S, Baravalle A, Falcarin P. Data quality challenges in large-scale cyber-physical systems: a systematic review. Inf Syst 2021;105:101951. https://doi.org/10.1016/j.is.2021.101951
[11] Cho S, Ensari I, Weng C, Kahn M, Natarajan K. Factors affecting the quality of Person-Generated wearable device data and associated challenges: rapid systematic review. JMIR Mhealth Uhealth 2021;9. https://doi.org/10.2196/20738
[12] Khwaja M, Vaid SS, Zannone S, Harari GM, Faisal AA, Matic A. Modeling personality VS. Modeling personalidad: in-the-wild mobile data analysis in five countries suggests cultural impact on personality models. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2019;3:1–24.
[13] Solouki MA, Angizi S, Violante M. Dependability in embedded systems: a survey of fault tolerance methods and Software-Based mitigation techniques. IEEE Access 2024;12:180939–67. https://doi.org/10.1109/access.2024.3509633
[14] Kondoro AM, Lee Y, Song J, Noh Y. Investigating user perceptions to mitigate dropout in longitudinal passive sensing studies. In: Proceedings of the extended abstracts of the CHI conference on human factors in computing systems, CHI EA '25. New York, NY, USA: Association for Computing Machinery; 2025. https://doi.org/10.1145/3706599.3719959.
[15] Lee Y, Hwang J, Kondoro AM, Song J, Noh Y. Enabling Delayed-Full Charging Through Transformer-Based Real-Time-to-Departure Modeling for EV Battery Longevity arXiv:2512.07723, 2025.
[16] Lee Y, Noh Y, Lee U. Data processing pipeline of Short-Term depression detection with Large-Scale dataset. In: 2023 IEEE international conference on big data and smart computing (BigComp). IEEE; 2023. p. 391–2.
[17] Akbarova S, Im M, Kim S, Toshnazarov K, Chung K-M, Chun J, Noh Y, Kim Y-A. Improving depression severity prediction from passive sensing: Symptom-Profiling approach. Sensors 2023;23. https://doi.org/10.3390/s23218866.
[18] Toshnazarov K, Lee U, Kim BH, Mishra V, Najarro LAC, Noh Y. SOSW: stress sensing with Off-the-Shelf smartwatches in the wild. IEEE Internet Things J 2024;11:21527–45. https://doi.org/10.1109/JIOT.2024.3375299
[19] Toshnazarov K, Lee TH, Noh Y. WildStress: exploring rich situational contexts for stress detection in the wild. In: 2023 IEEE international conference on big data and smart computing (BigComp). IEEE; 2023. p. 381–2.

[20] Ranjan Y, Rashid Z, Stewart C, Conde P, Begale M, Verbeeck D, Boettcher S, Dobson R, Folarin A. RADAR-base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices. JMIR Mhealth Uhealth 2019;7. https://doi.org/10.2196/11734

[21] Onnela J-P, Dixon C, Griffin K, Jaenicke T, Minowada L, Esterkin S, Siu A, Zagorsky J, Jones E. Beiwe: a data collection platform for High-Throughput digital phenotyping. J Open Source Softw 2021;6:3417. https://doi.org/10.21105/joss.03417

[22] Ferreira D, Kostakos V, Dey AK. AWARE: mobile context instrumentation framework. Front ICT 2015;2:6.

[23] González-Pérez A, Matey-Sanz M, Granell C, Díaz-Sanahuja L, Bretón-López J, Casteleyn S. AwarNS: a framework for developing context-aware reactive mobile applications for health and mental health. J Biomed Inform 2023;141:104359.

[24] Hossain SM, Hnat T, Saleheen N, Nasrin NJ, Noor J, Ho B-J, Condie T, Srivastava M, Kumar S. Mcerebrum: a mobile sensing software platform for development and validation of digital biomarkers and interventions. In: Proceedings of the 15th ACM conference on embedded network sensor systems; 2017. p. 1–14.

[25] Vandenberk T, Storms V, Lanssens D, Cannière HD, Smeets C, Thijs I, Batool T, Vanrompay Y, Vandervoort P, Grieten L. A Vendor-Independent mobile health monitoring platform for digital health studies: development and usability study. JMIR Mhealth Uhealth 2019;7. https://doi.org/10.2196/12586

[26] Vaidyam A, Halamka J, Torous J. Enabling research and clinical use of Patient-Generated health data (the mindLAMP platform): digital phenotyping study. JMIR Mhealth Uhealth 2021;10. https://doi.org/10.2196/30557

[27] Otorkpa OJ, Alao JO, Olaiya AP. Enhancing monitoring and evaluation of digital health interventions in sub-saharan africa: big data, mhealth, and dashboards. J Glob Health 2025;15. https://doi.org/10.7189/jogh.15.03013

[28] Pulantara I, Wang Y, Burke L, Sereika S, Bizhanova Z, Kariuki J, Cheng J, Beatrice B, Loar I, Cedillo M, Conroy MB, Parmanto B. Data collection and management of mhealth, wearables, and internet of things in digital behavioral health interventions with the awesome data acquisition method (ADAM): development of a novel informatics architecture. JMIR Mhealth Uhealth 2023;12. https://doi.org/10.2196/50043

[29] Akbar A, Kousiouris G, Pervaiz H, Sancho J, Ta-Shma P, Carrez F, Moessner K. Real-Time probabilistic data fusion for Large-Scale IoT applications. IEEE Access 2018;6:10015–27. https://doi.org/10.1109/access.2018.2804623

[30] Rad ZS, Ghobaei-Arani M. Data pipeline approaches in serverless computing: a taxonomy, review, and research trends. J Big Data 2024;11:1–42. https://doi.org/10.1186/s40537-024-00939-0

[31] Xiong H, Huang Y, Barnes LE, Gerber MS. Sensus: a cross-platform, general-purpose system for mobile crowdsensing in human-subject studies. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, UbiComp '16. New York, NY, USA: Association for Computing Machinery; 2016. p. 415–26. https://doi.org/10.1145/2971648.2971711.

[32] Kim J, Lee D, Lee U. AffectStream: kafka-based real-time affect monitoring system using wearable sensors. Softwarex 2025;31:102325. https://doi.org/10.1016/j.softx.2025.102325. https://www.sciencedirect.com/science/article/pii/S2352711025002912.

[33] Ramírez J, Ramírez J, Rubiano A. Muscle fatigue monitoring and analysis system: myosense. Softwarex 2025;32:102360. https://doi.org/10.1016/j.softx.2025.102360. https://www.sciencedirect.com/science/article/pii/S2352711025003267.

[34] Abdolkhani R, Gray K, Borda A, DeSouza R. Recommendations for the quality management of Patient-Generated health data in remote patient monitoring: mixed methods study. JMIR Mhealth Uhealth 2021;11. https://doi.org/10.2196/35917

[35] Salman M, Son J-H, Choi D-W, Lee U, Noh Y. DARCAS: dynamic association regulator considering airtime over SDN-enabled framework. IEEE Internet Things J 2022;9:20719–32.