University of Trento

Department of Information Engineering and Computer Science

Doctoral Thesis

~ · ~

A Quality-Aware Calendar Dashboard for Personal Data Reflection: Design, Implementation, and Empirical Evaluation

Advisor
Prof. Fausto Giunchiglia

Author
Ivan Kayongo

Final examination date: April 1, 2026

Dedication:

I dedicate this dissertation to my cherished son, Josiah; my beautiful daughter, Josiella; and my exceptional wife, Phiona, whose love, patience, and understanding have sustained me throughout this endeavour. Thank you for having endured my absence and supporting me with unwavering strength through the many demanding moments of this journey.

I also want to thank my mother for always believing in my potential and being a guiding influence in my life. Her encouragement and prayers have carried me farther than words can express.

To each of you, this achievement is as much yours as it is mine.

Acknowledgments

First and foremost, I thank God for giving me wisdom, grace, and resilience throughout this journey. This work would not have been possible without His guidance and strength, which sustained me through every challenge and milestone.

I wish to express my deepest gratitude to my PhD advisor, Professor Fausto Giunchiglia, whose steadfast guidance, intellectual generosity, and unwavering support have shaped this thesis from its inception to completion. His insightful critiques, patience, and dedication to my scholarly development have been instrumental to both my personal growth and the successful completion of this work. I am truly honoured to have had the privilege of learning under his mentorship.

I am also profoundly grateful to my colleagues and fellow researchers, whose collaboration and stimulating discussions have greatly enriched this endeavour. The shared experiences, thoughtful debates, and mutual support throughout this journey have made the process not only more productive but also profoundly rewarding.

My heartfelt thanks go to my family and friends for their immeasurable support. Their steadfast belief in me, boundless patience, and constant encouragement have been a source of strength and motivation. Each, in their own unique way, has contributed to this achievement, and for this, I remain forever thankful.

To all who have stood by me throughout this long journey—thank you.

Acknowledgement of Generative AI Assistance

While writing this thesis, generative AI tools like ChatGPT, DeepSeek, QuillBot, Copilot, and Grammarly were used in a limited way to help with the writing and presentation of the work.

These tools were used for:

- language refinement (improving clarity, grammar, and academic tone),
- formatting assistance (e.g., structuring tables, cross-references, and ensuring stylistic consistency),

- editorial support (identifying redundancy and suggesting structural improvements), and
- code assistance (supporting the development and debugging of implementation scripts).

All substantive intellectual contributions – including the research design, methodology, data collection, analysis, and interpretation – are solely my own. No generative AI tools were used to produce original research contributions, results, or scientific claims.

All outputs generated by these tools were critically evaluated, revised, and validated before inclusion. I retain full responsibility for the accuracy, integrity, and originality of the content presented in this thesis.

Abstract

Personal informatics systems are increasingly collecting multimodal data from smartphones and wearables, including sensor streams, self-reports, and behavioural logs. However, two limitations constrain their usefulness for reflection. First, data quality remains opaque: systems treat all observations as equally reliable despite missing data, delayed responses, or sensor noise, preventing users from distinguishing genuine behavioural patterns from artefacts of data collection. Second, planning and reflection are structurally separated: calendars represent planned activities, while dashboards summarise past behaviour, making it difficult to relate what was planned to what actually occurred. As a result, users lack the ability to reason about their behaviour in a way that is both temporally integrated and grounded in the reliability of the underlying data.

This thesis explores how the calendar can be reconfigured to support reflection over heterogeneous personal data. Rather than treating it solely as a scheduling tool, a calendar-based interface was developed that brings together planned activities, sensor-derived behaviour, and self-reports within a shared temporal view. This allows users to see not only what was intended, but also what actually occurred and when. To support interpretation, the system also exposes missing entries, delayed responses, and sensor gaps through visual indicators, making the reliability of the data visible at the moment of interpretation.

This work makes three contributions. First, we introduce a temporal integration approach that aligns heterogeneous data streams within a shared calendar structure, enabling direct comparison between plans and observed behaviour. Second, we formalise a computational data quality framework for personal informatics, operationalising metrics such as question delivery success, response latency, answer duration, and sensor completeness. Third, we report results from two evaluations: (a) a controlled study (N=130) showing that quality-aware visualisation improves the accuracy of behavioural reflection, and (b) a six-week field deployment (N=73) demonstrating that real-time monitoring of data quality supports timely intervention, which was associated with improved data completeness. In the field deployment, researchers were able to detect data gaps and intervene within 24 hours, with response rates increasing following reminders, helping sustain participant engagement over time.

Keywords: personal informatics, self-tracking, behavioural data, sensor logging, ubiquitous computing, self-reflection, mobile devices

Contents

Glossary

Context-Awareness – The ability of a user to understand their behaviour according to sensed or inferred contextual information, such as time, location, movement, or device usage.

Experience Sampling Method (ESM) – A research method in which participants receive short surveys throughout the day to capture momentary states, behaviours, or experiences.

Personal Context – The set of environmental, behavioural, temporal, and social conditions that describe a user at a specific moment.

Time Diaries – Structured logs or systems used to record participants' daily activities, often supported by predefined activity categories.

Nomenclature

Table 1: Abbreviations used throughout the thesis

Abbreviation	Meaning
API	Application Programming Interface
ESM	Experience Sampling Method
EMA	Ecological Momentary Assessment
DB	Database
DTO	Data Transfer Object
UI	User Interface
UX	User Experience
GPS	Global Positioning System
RT	Reaction Time (sometimes interchangeable with `rxn`)
DRN	Response Duration (ESM answer duration metric)
SSID	Service Set Identifier (Wi-Fi network name)
Wi-Fi	Wireless Fidelity (IEEE 802.11 standard)
RSSI	Received Signal Strength Indicator
BLE	Bluetooth Low Energy
OS	Operating System
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
REST	Representational State Transfer
CRUD	Create, Read, Update, Delete
JSON	JavaScript Object Notation
CSV	Comma-Separated Values
URL	Uniform Resource Locator
URI	Uniform Resource Identifier
JWT	JSON Web Token
OAuth	Open Authorization (authentication protocol)
CORS	Cross-Origin Resource Sharing
ORM	Object–Relational Mapping
JPA	Java Persistence API
MVC	Model–View–Controller
SPA	Single-Page Application
KPI	Key Performance Indicator

Table 2: Ethics, governance, and regulatory abbreviations

Abbreviation	Meaning
UNCST	Uganda National Council for Science and Technology
MUIRB	Makerere University Institutional Review Board
GDPR	General Data Protection Regulation (EU)
DPA	Data Protection and Privacy Act (Uganda, 2019)
IRB	Institutional Review Board
EU	European Union

Table 3: Symbols and variables used in data analysis

Symbol	Description	Unit
uid	User identifier	–
eid	Experiment identifier	–
sent	Number of questions sent	count
success	Successfully delivered questions	count
failed	Questions not delivered	count
answers	Number of answered questions	count
rxn	Reaction time (sent → opened)	ms
drn	Answering duration (opened → submitted)	ms
t	Timestamp	ms or ISO 8601
lat	Latitude	degrees
lon	Longitude	degrees
freq	Frequency of an event or activity	count

Table 4: System-specific terms and components

Term	Description
`timediariesquestions`	Database table containing all ESM-style questions
`timediariesanswers`	Database table containing participant responses to ESM questions
`pushController()`	Navigation method used to transition between screens in the Android application
Heatmap / Stacked bar / Word cloud	ECharts visualisation types used in the dashboard interface
Experiment dashboard	Aggregated view presenting experiment-level summaries and key metrics
Participant dashboard	User-facing view showing individual participation status and collected data
Context dimensions	Set of contextual factors (e.g., activity, location, mood, companions, sensors) used in analysis

Table 5: Temporal concepts used in the dashboard design

Term	**Description**
Past	Historical period representing completed activities, answered ESMs, and logged sensor data used for reflection and analysis
Present	Current moment capturing ongoing activities, recently triggered questions, and real-time or near-real-time contextual signals
Future	Prospective period representing scheduled events, pending ESM prompts, and anticipated activities used for planning and preparation
Past calendar	Calendar view showing completed events enriched with contextual dimensions and sensor summaries
Present state	System state reflecting active participation status and currently available interactions
Future calendar	Calendar view displaying upcoming events and questions awaiting response
Temporal continuum	Integrated representation linking past records, present behaviour, and future intentions within a single dashboard

Chapter 1

Introduction

1.1 Motivation: The Promise and Challenge of Personal Data

The way individuals organise their daily lives depends fundamentally on how they understand their own behaviour. Studies have shown that people track data not just to record events but also to compare behaviour over time and identify deviations from routine [40]. By reflecting on this data, users can identify patterns in their daily routines or gauge their progress towards their goals [41]. This reflective insight supports comparison to goals or expectations, enabling people to recognise discrepancies between their behaviour and their desired states and adapt accordingly; such feedback mechanisms are central to theories of self-regulation and behaviour change [52]. To sustain this process of self-regulation, individuals therefore require clear and reliable representations of their behaviour that make patterns, deviations, and influences of the environment understandable over time.

Opportunities for self-understanding have been made available by the proliferation of modern cellphones and wearable technologies. These devices are capable of continuously recording behavioural data through the use of sensors. For example, GPS location can track movement through physical space; accelerometers can capture physical activity and rest; app usage logs can reveal patterns of digital engagement; and screen events can indicate periods of attention and idleness [39]. These sensors, in contrast to previous approaches that rely simply on user self-reports, allow for the real-time capturing of behavioural and contextual signals as they occur. Their purpose is to provide what academics refer to as ”ecologically valid” evidence of ordinary life [37].

When this sensor data is combined with user self-reports collected through experience sampling methods (ESM), they form what has been described as ’big-thick’ personal data: information that is both large in scale and rich in contextual meaning [34]. The ’big’ dimension refers to the volume, velocity, and variety of passively sensed data, whereas the 'thick' dimension captures the subjective meaning and context that only humans can provide: mood, motivation, social context, and the reasons behind behaviour [12]. Together, these complementary data types promise a holistic view of daily life that neither can provide alone.

The value of such integrated data has been demonstrated across multiple domains. In behavioural science, continuous sensing enables observation of real-world activity as it naturally occurs, offering evidence that complements and often strengthens laboratory findings [30]. In digital health, moment-by-moment sensing supports personalised interventions that move beyond simple user self-reporting toward timely, situational support [31]. Within human–computer interaction, such data underpin the development of context-aware systems that adapt to users' immediate contexts as well as their longer-term behavioural patterns [4]. Across these fields, the basic idea is that rich, continuous personal data can facilitate deeper understanding and more responsive technologies.

However, realising this promise requires addressing several challenges. Sensor streams vary in sampling rate, reliability, and semantic formation [56], and are often collected at intervals determined by study design or user initiative. Each data source exhibits distinct failure modes; for instance, GPS signals degrade indoors; accelerometers can impose significant battery overhead at high sampling frequencies; and self-reports depend on user compliance and sustained attention. Without careful alignment and contextualisation, the resulting data may be more confusing than informative, which could increase cognitive load and heighten the risk of misinterpretation, a limitation documented across studies of personal informatics systems [22, 16].

1.2 The Persistence of Fragmentation in Personal Data Systems

Despite decades of research in ubiquitous computing and personal informatics, three fundamental gaps still exist in how personal data is collected, integrated, and presented to users.

The Contextualisation Gap

Digital calendars provide a structured representation of time and planned activities. They show what a person intended to do, including meetings scheduled, deadlines set, and appointments arranged. Yet calendars remain blind to what actually happened, i.e., whether the meeting occurred, how the person felt during the meeting or what they did afterwards. Personal informatics systems, on the other hand, capture rich behavioural traces but lack the intentional, forward-looking temporal framing that calendars provide. For example, a fitness tracker records steps taken but not whether they were part of a planned workout or incidental movement, or a mood log captures emotional states but not how they relate to scheduled commitments. This separation prevents users from asking questions that are both reflective and causal: 'Do my late-night meetings affect my sleep?', 'How do my daily commitments influence my mood?', 'Did

I actually follow my planned study schedule?' Without integration of the two, these questions remain unanswerable.

Unlike standalone dashboards, calendars already encode temporal intent, making them uniquely suited for aligning planned and observed behaviour within a shared time structure. This built-in temporal organisation—the division of life into days, weeks, and months—mirrors how humans naturally perceive and remember their experiences. By leveraging this existing framework, a calendar-based approach can reduce the mental effort required to integrate disparate data sources, potentially supporting more intuitive and effective reflection.

The Data Quality Understanding Gap

A more tricky problem concerns the reliability of the data itself. Personal data is usually noisy or incomplete, whereas sensors at times malfunction, lose signal, or are disabled by users. User self-reports too may be submitted late, answered hastily, or skipped entirely [9]. Yet in reality, most systems treat all incoming data as equally trustworthy. Few assess or visualise data quality—sensor reliability, data missingness, or metadata such as response delay—at all. As a result, users may base decisions on skewed or incomplete information without realising it. For instance, a person reviewing their weekly step counts cannot tell whether a low day reflects genuine inactivity or simply a day they left their phone at home on the charger. A researcher analysing mood patterns cannot distinguish between genuine emotional variation and artefacts of when users chose to respond. The critical question of—"Can I trust this insight?"—remains not only unanswered but unasked.

Experience sampling studies frequently encounter situations where participants do not respond within the intended time frame, which can affect the accuracy of the data. Passive sensing datasets also often suffer from incomplete coverage, stemming from differences in device models, operating system versions, and user behaviour. These issues are not merely random fluctuations; they reflect systematic factors that can influence the reliability and interpretability of the collected data [30, 9].

The Architectural Gap

Making sense of heterogeneous temporal data depends on a set of underlying technical considerations. Data from different streams need to be synchronised despite varying sampling rates; for instance, aligning a 50Hz accelerometer signal with user self-reports requires temporal precision and careful handling of missing values. It is also important to retain quality-related metadata throughout the pipeline—from data collection to processing and visualisation—so that reliability can be assessed at the point of interpretation. At the same time, visualisations should convey the data without overwhelming users, a challenge that still exists in personal informatics [2]. In practice, many research platforms address data quality only after collection, treating it as part of data cleaning rather than embedding it within collection and presentation [24, 48].

Taken together, these limitations point to a central issue: **individuals lack access to integrated,**

quality-aware representations of their own behaviour that support trustworthy reflection and informed decision-making.

1.3 The Opportunity: Calendars as Temporal Scaffolds

One approach to addressing these gaps is to use the digital calendar as an integrative temporal framework. Calendars have long been used to organise planned activities and commitments, supporting coordination, planning, and shared understanding across individuals and groups [50]. In their digital form, they are part of everyday computing environments, synchronising schedules and reminders across devices and contexts. They are also widely used, making them a familiar tool for personal information management [57].

But despite these advantages, current digital calendars have limitations. They mainly focus on planned events and usually don't consider the behavioural and contextual information that shapes our daily lives. A calendar may record a scheduled meeting, but not whether it was attended, how the user felt during the meeting, or whether poor sleep from the night before had an impact. This information often exists (in fitness trackers, mood logs, and location histories) but remains distributed across separate applications, each with its own interface, data model, data quality and end goal.

In this thesis, we propose extending the calendar from a scheduling tool to a quality-aware, context-enriched reflective dashboard through the integration of multimodal personal data into its temporal structure. This approach would allow users to see not only what was planned, but also what actually occurred, when it occurred, and how reliable the underlying data is. By bringing planned events, observed behaviour, and contextual signals together within a shared temporal framework, the calendar can support forms of reflection and causal reasoning that fragmented systems alone cannot provide.

Illustrative Scenarios

To ground the discussion, consider three scenarios that motivate the need for an integrated, quality-aware calendar dashboard.

Sleep and late meetings. A graduate student experiences persistent fatigue each Friday. His wearable device indicates poor sleep on Thursday nights, but the cause remains unclear until sleep data is aligned with his calendar. The pattern becomes obvious: every Thursday includes a late seminar, and each Friday shows corresponding sleep disruption. Without this combined view, the pattern would remain hidden; with it, the student gains actionable insight—a specific commitment to modify rather than a vague sense of inconsistent sleep.

Workload and physiological strain. An office professional's calendar appears fully booked but manageable. Her smartwatch data, however, shows an increased heart rate in the afternoons, with spikes during certain meetings, while her phone records indicate extended screen time late at night. Each of these observations alone is unclear, but when combined with her calendar, the connection between her schedule and her physical stress becomes clear. This combined view allows for informed actions, such as reorganizing meeting times and protecting time for relaxation after work, rather than relying on guesswork.

Physical activity, social context, and data quality. A third scenario underscores the role of data quality. A user's weekends consistently show high physical activity and elevated well-being reports. However, mood logs are often submitted hours after events, raising uncertainty: does the reported mood reflect the actual social experience, or has it been influenced by later fatigue or stress? By visualising not only the content of self-reports but also their timing (e.g., response latency indicators), the dashboard allows users to weigh evidence appropriately, distinguishing genuine behavioural patterns from artefacts of data collection.

1.4 Research Questions

This thesis addresses three research questions, each corresponding to a different facet of the work: design, empirical understanding, and methodology.

RQ1 (Design Inquiry): *How can a calendar-based interface integrate multimodal personal data—sensor streams, self-reports, and quality metadata—into a unified temporal view that supports both participant reflection and researcher monitoring?*

This question is answered through the design and implementation of the dashboard itself. Chapter 4 presents the design rationale, including the temporal structure (past, present, future), the five-dimensional context model, and the integration of quality indicators. Chapter 5 describes the technical architecture. The usability findings from both evaluations (Chapter 6) provide evidence that the design achieves its goals of supporting reflection and monitoring.

RQ2 (Empirical Inquiry): *How do users perceive and interpret quality indicators when reflecting on personal data presented in a calendar dashboard?*

We solve this through a controlled study with 130 participants (Section 6.2). Participants interacted with the dashboard, viewed synthetic personal data with embedded quality indicators, and completed a structured questionnaire. The Likert-scale responses (Appendix .2) and qualitative feedback reveal which aspects of the dashboard supported clarity and reflection and which caused confusion. Key findings include high ratings for location and activity visualisation, mixed responses to social context reflection, and lower clarity for material context (objects/tools).

RQ3 (Methodological Inquiry): *In a longitudinal field deployment, can a real-time researcher dashboard enable proactive management of participant engagement and data quality?*

This was answered through a six-week study deployment with 73 university students in Uganda (Section 6.3). The dashboard was used to monitor response rates, detect data gaps, and trigger researcher interventions. Findings demonstrate that real-time monitoring enabled detection of disengagement within 24 hours, that email reminders were followed by recovery in response rates, and that the dashboard revealed participant mental models (e.g., deactivating the app when questioning stopped). These operational insights contribute methodological knowledge for researchers conducting longitudinal mobile studies.

1.5 Approach

To answer these questions, we developed a calendar-based system and evaluated it in controlled and real-world settings. The system integrates three core components:

1. **A unified temporal framework** that aligns planned events (calendar), recorded behaviour (sensor streams), and self-reported context (ESM responses) within a single timeline. This framework operationalises the model of context presented in Chapter 3, where each moment is described through five dimensions: activity, location, social context, internal state, and objects or tools used.

2. **Explicit quality indicators** derived from the data quality framework, including question delivery success, response latency, answer duration, and sensor completeness. These indicators are computed during data collection, stored alongside the primary data, and visualised in the interface through graphs.

3. **Dual-role interfaces** supporting both participant reflection and researcher monitoring. The participant interface presents personal data within calendar views and graphs, enabling users to navigate and compare planned and recorded activities over time, enriched with contextual information. The researcher interface provides aggregate views of data quality across participants, enabling the detection of missing data, declining engagement, or sensor failures.

The system is built within the iLog ecosystem [58], which provides the infrastructure for multimodal data collection from smartphones and wearables. iLog collects sensor data (location, accelerometer, screen status, Bluetooth, and Wi-Fi) and administers experience sampling questionnaires, storing all data with timestamps and associated quality metadata. The dashboard accesses this data through REST APIs, providing visualisation and interaction for both researchers and participants.

1.6 Scope and Limitations

The scope of this work is defined by several design choices:

Population. The evaluation considers two groups: participants recruited through Prolific[1] (N=130) and university students involved in a six-week field deployment (N=73). The findings may not generalise to other populations, such as clinical groups or children.

Data types. The focus is on data commonly available from consumer smartphones, including location (GPS, Wi-Fi), motion (accelerometer, step counter), device interaction (screen status, app usage), and self-reports (ESM responses, daily diaries). Physiological sensors (e.g., heart rate, skin conductance), environmental sensors (e.g., air quality, noise), and smart home data are not considered.

Duration. The longest deployment spans six weeks. Longer-term studies may reveal different patterns of engagement, data quality, and reflection.

1.7 Contributions

This thesis makes three specific contributions:

1. **A formalised data quality framework for personal informatics.** We operationalise metrics for question delivery success, response latency, answer duration, and sensor completeness. Using data from a six-week field deployment (N=73), we provide evidence that question response improved after researchers encouraged poorly performing participants to improve their performance. This establishes simple, automatically measurable indicators as valid proxies for data reliability, enabling quality-aware collection and analysis.

2. **A scalable temporal integration architecture.** We demonstrate a system design that aligns heterogeneous data streams with differing sampling rates while preserving quality metadata throughout the pipeline. The architecture successfully ingested and synchronised sensor data, self-reports, and calendar events from 73 participants over six weeks, supporting interactive data querying and visualisation. We report implementation details, performance characteristics, and lessons learned that can inform future systems.

3. **Empirical evidence from two evaluations.** We report results from:

   - A controlled study (N=130) testing whether quality visualisation improves reflection. Participants viewed 70 days of synthetic personal data and completed tasks requiring them to distinguish reliable patterns from artefacts of missing or low-quality data.
   - A six-week field deployment (N=73) testing whether real-time researcher monitoring improves data completeness. Findings show that monitoring enables rapid intervention and maintains engagement: real-time monitoring allowed researchers to detect data gaps within 24 hours, and visual inspection of response trends showed recovery following email reminders. Average daily response rates increased following researcher contact, demonstrating that the approach helps sustain participation.

[1] https://www.prolific.com/

Collectively, these contributions advance the state of knowledge in personal informatics by: (a) providing validated metrics for data quality that can be adopted by other researchers; (b) demonstrating a feasible architecture for quality-aware personal data systems.

1.8 Thesis Structure

The remainder of this thesis is organised as follows:

Chapter 2: Background and Related Work reviews the literature on personal informatics, context-aware systems, and data quality, identifying the specific gaps this thesis addresses and situating the work within the broader research landscape.

Chapter 3: Modelling Personal Data presents our theoretical framework for representing personal context and data quality. It defines the five-dimensional context model (activity, location, social context, internal state, objects) and formalises metrics for question delivery success, response behaviour, and sensor completeness.

Chapter 4: Dashboard Calendar presents the design and development of our proposed quality-aware calendar dashboard. It details the resulting system and how the different user roles are supported.

Chapter 5: System Architecture. It details the technical implementation of the proposed system, outlining its architecture, core components, data storage and processing mechanisms, and its interaction with external services.

Chapter 6: System Evaluation This chapter reports on two studies, describing for each the design, procedures, participant characteristics, and results, and concludes with a discussion of limitations and implications.

Chapter 7: Conclusion summarises the thesis contributions, examines its limitations, and outlines directions for future work.

1.9 Related Publications

Portions of this work have appeared in the following peer-reviewed publications:

1. Kayongo, I., Zhao, H., Malcotti, L., & Giunchiglia, F. (2024). A methodology and system for big-thick data collection. *Informatik Festival.* `https://dl.gi.de/server/api/core/bitstreams/346e0026-1454-4bf5-81cd-3ccca1cbed61/content`
2. Kayongo, I., Malcotti, L., Zhao, H., & Giunchiglia, F. (2025). A methodology and a platform for high-quality rich personal data collection. *The European Journal on Artificial Intelligence*, 38(4), 474–495.
doi10.1177/30504542513336l5

3. Kizito, M., Kayongo, I., Nyende, H., et al. (2025). MakOne: Behavioural data of university students' smart devices in Uganda. *Proceedings of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 815–819. doi10.1145/3714394.3756183

4. Busso, M., Bontempelli, A., Malcotti, L., Meegahapola, P., Kun, P., Diwakar, S., Nutakki, C., Rodas Britez, M. D., Xu, H., Song, D., Ruiz Correa, S., Mendoza-Lara, A.-R., Gaskell, G., Stares, S., Bidoglia, M., Ganbold, A., Chagnaa, A., Cernuzzi, L., Hume, A., Chenu-Abente, R., Alia Asiku, R., **Kayongo, I.**, Gatica-Perez, D., de Götzen, A., Bison, I., & Giunchiglia, F. (2025). DiversityOne: A multi-country smartphone sensor dataset for everyday life behaviour modeling. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 9(1), Article 1. doi10.1145/3712289

Chapter 2

Background and Related Work

Understanding the current research landscape in context-aware systems, personal informatics, and multimodal data collection is essential for situating this work and clarifying its contribution. Over the last decade, the proliferation of smart devices has transformed how personal data is captured, interpreted, and integrated into daily life. Advances in human–computer interaction (HCI), behavioural science, and ubiquitous computing have produced a variety of systems that support reflection, behaviour change, and context-aware interaction [47, 3]. Yet, important gaps remain, particularly in how personal data is represented, how contextual signals are embedded into everyday tools such as calendars, and how data quality is communicated to users.

This chapter first examines the conceptual foundations of personal context, particularly within big-data frameworks (Section 2.1), focusing on definitions, classification, and applications in ubiquitous computing. While this literature provides a theoretical basis for understanding contextual cues, these cues are rarely operationalised in personal information management tools that structure everyday life. Section 2.2 reviews existing calendar and dashboard systems, evaluating their approaches to data integration and visualisation while highlighting limitations in supporting meaningful reflection and behaviour awareness. Section 2.3 surveys personal informatics systems across domains, analysing their data collection strategies, visualisations, and constraints. Finally, Section 2.4 synthesises these insights into a diagnosis of the research gap and outlines the requirements driving this thesis: extending the iLog sensing platform [58] with a quality-aware calendar dashboard that integrates multimodal personal data into a unified, temporally anchored interface.

2.1 Big-Thick Data

Modern digital ecosystems continuously generate large streams of personal information through smartphones, wearables, IoT devices, and online services. This ”big data” is commonly characterised by the three Vs: volume, velocity, and variety [18]. Sensor readings (e.g., accelerometer and

GPS), application usage logs, and digital traces (e.g., click streams and transaction records) provide objective, longitudinal measurements of behaviour at scales previously impossible. Such data enables statistical analyses and predictive modelling, uncovering trends like the relationship between smartphone use and sleep disruption [33] or clustering mobility patterns in urban populations [38].

However, this quantitative orientation creates a structural problem for systems designed to support personal reflection. Big data captures *what* happened but strips away the contextual cues that explain *why*. A spike in heart rate could stem from anxiety or exercise; a location trace shows a user at a café, but not whether they were working, socialising, or avoiding work. When integrated into temporal interfaces such as calendars, this contextual sparsity becomes a fundamental limitation: the data is temporally anchored but semantically shallow, leaving users to infer meaning without the necessary interpretive grounding [13].

In contrast, thick data emphasises the depth of human experience. It is collected through interviews, participant observation, daily self-reports, and surveys [26, 12, 35]. Thick data reveals motivations, emotions, and social contexts, providing the interpretive richness that quantitative approaches miss. For example, self-reports can show how looming deadlines affect well-being beyond the number of hours worked. However, thick data introduces a complementary structural problem: it is temporally sparse and episodically collected, making it difficult to align with the continuous timelines that calendars provide.

This complementary pattern reveals a structural misalignment. Big data is continuous but semantically shallow; thick data is semantically rich but temporally sparse. Neither alone is sufficient for systems that aim to support reflection within a temporal framework. Recognising this, scholars have proposed integrating the two approaches under the concept of ”big-thick data” [12, 10]. This hybrid approach combines the pattern-finding power of big data with the interpretive richness of thick data, allowing computational models to detect trends while human insight provides context and interpretation.

For personal informatics and reflective tools such as calendar dashboards, big-thick data provides a practical framework. Quantitative streams identify longitudinal patterns such as declining sleep quality every Thursday night while qualitative inputs explain their causes, such as late work meetings or social commitments [16]. The challenge lies in operationalising this integration: creating representational structures that can align continuous sensor streams with sparse self-reports within a unified temporal framework. This is the problem to which we now turn.

2.1.1 Context

While big–thick data provides a rich account of behaviour, motivation, and situation, its heterogeneous nature, ranging from high-frequency sensor streams to sparse, qualitative self-reports, creates alignment challenges. To address this, an intermediate representational layer is required. In this thesis, we adopt the notion of *context* as this unifying abstraction.

The concept of context has evolved alongside shifts in computing paradigms. In early work within Knowledge Representation (KR) and Artificial Intelligence (AI), context was treated as a formal mechanism for constraining reasoning, allowing systems to interpret information relative to a bounded set of assumptions [45, 27]. With the emergence of ubiquitous computing, the focus moved toward situated interaction. Schilit et al. [54] described context in terms of observable environmental properties, such as location, nearby people, and surrounding objects. Brown et al. [14] extended this view to include temporal and social factors, emphasizing that context is not static but evolves with changes in time, roles, and settings.

Subsequent work introduced a more explicitly user-centred perspective. Dey et al. [20] framed context as encompassing the user's physical, social, emotional, and informational state. This broader interpretation was formalized by Dey and Abowd [1], who defined context as any information that characterizes the situation of an entity involved in an interaction. This definition remains influential because of its generality: it accommodates both externally observable factors and internal, subjective states.

Building on this foundation, this thesis adopts a multi-dimensional model of personal context. Following Giunchiglia et al. [28, 29] and grounded in the earlier framework by Abowd et al. [1], we operationalise context through five core dimensions. These dimensions serve both as an analytical lens for reviewing existing systems and as a design framework for the proposed solution:

Activity (What): The actions or tasks in which the individual is engaged (e.g., working, commuting, sleeping, socialising). This dimension captures intentional behaviour.

Location (Where): The spatial setting of the individual (e.g., home, workplace, café, park).

Social context (Who): The presence and roles of others (e.g., alone, with colleagues, with family). This reflects the interpersonal dimension of experience.

Internal state (How): The individual's subjective condition, including mood, stress, and energy levels. This captures experiential qualities that are not directly observable.

Objects (With what): The artefacts and tools in use (e.g., phone, laptop, book). This reflects the material context in which activities occur (this was added in our work during this thesis).

This representation provides a common structure through which heterogeneous data sources can be aligned. Sensor-derived data, such as location traces, can be interpreted alongside reported activities; subjective inputs, such as mood, can be situated within specific temporal and social settings. It also enables the analysis of discrepancies, for example, between planned and actual behaviour, by relating them to both environmental conditions and internal states. Throughout this chapter, we use this five-dimensional model as an evaluative lens to assess existing systems, identifying which dimensions they capture and, crucially, which they omit.

2.2 Existing Calendar-Dashboard Systems

Digital calendars have evolved well beyond their original role as simple scheduling tools. They now function as central hubs for organising both professional and personal life, prompting the emergence of calendar-based dashboards aimed at helping individuals make sense of their daily activities. In this expanded role, calendars support not only planning but also recall, coordination, and lightweight forms of life logging [57, 44].

At the same time, individuals rarely rely on a single system. Instead, they construct what Payne and Maximilien describe as ”personal ecologies” of tools, combinations of paper diaries, digital calendars, workplace systems, and mobile notifications that reflect their routines and constraints [50]. Within such fragmented environments, the idea of a unified calendar dashboard that consolidates diverse data streams is both a natural progression and a non-trivial design challenge.

Several systems attempt to reposition the calendar as a representation of lived experience rather than a record of planned events. This aligns with a central principle in personal informatics: data becomes meaningful only when situated within the context of everyday life [40]. In principle, a calendar dashboard could bridge intention and behaviour by bringing together scheduled events, observed activities, and reflective insights within a single interface.

In practice, however, most existing systems fall short of this goal. A common underlying assumption is that temporal data becomes self-explanatory once visualised. As a result, many dashboards prioritise clarity of presentation, through colour coding, categorisation, or aggregation, over mechanisms that support interpretation. While such representations are effective for planning, they offer limited support for understanding behaviour. For example, visualising a week containing 30 hours of meetings is descriptive, but it does not explain how this workload affected mood or productivity, what activities were displaced, or whether similar patterns recur over time.

To address these limitations, more recent systems incorporate additional data sources, such as activity logs, mobility traces, and self-reports. The expectation is that richer data will enable deeper insight. Activity River [2], for instance, integrates planned and recorded activities within a shared timeline. However, evidence from self-tracking research suggests that simply increasing data volume does not guarantee understanding. Users often struggle to interpret heterogeneous datasets or identify meaningful relationships without substantial effort [22]. As a result, dashboards that present multiple disconnected metrics, such as steps, sleep, and meetings, can overwhelm rather than inform.

This points to a more fundamental issue. Calendars inherently segment time into discrete units, yet everyday experience is continuous, overlapping, and context-dependent. When dashboards adopt the same fragmented structure, they risk presenting isolated data points rather than coherent narratives. The challenge, therefore, is not only to visualise data accurately but also to support interpretation: helping users relate activities to one another, understand their consequences, and situate them within broader patterns of behaviour.

2.2.1 Activity River

Activity River [2] explores the relationship between planned and actual behaviour through a timeline-based visualisation. The system places scheduled calendar events alongside manually logged activities, allowing discrepancies between intention and execution to be identified at a glance (see Figure 2.1).

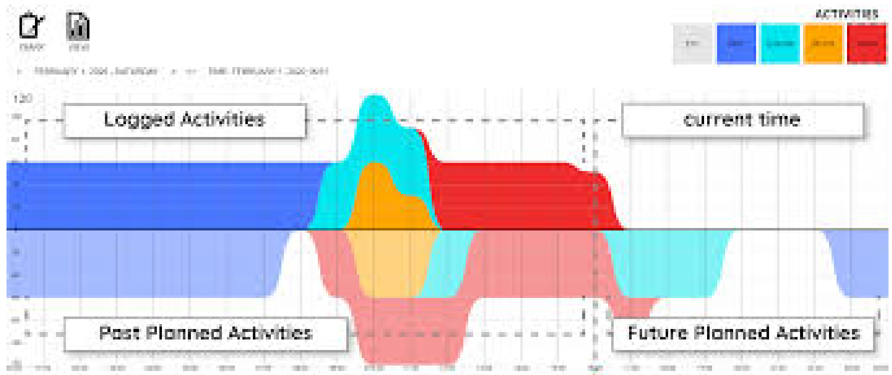


Figure 2.1: Activity River's side-by-side timeline visualisation. The figure illustrates how planned events (top) are juxtaposed with manually logged activities (bottom), making discrepancies between intention and execution visible. However, the absence of contextual layering—no location, social context, or internal state—limits interpretability. Source: [2]

Evaluated against our five-dimensional context framework, Activity River captures **activity** (both planned and logged) and, through manual entry, limited **objects**. It captures no information about **location**, **social context**, or **internal state**. More fundamentally, the system assumes that visualising temporal differences is sufficient to support reflection, yet prior work shows that even well-designed visualisations are difficult to interpret without additional contextual cues [16]. Activity River thus makes **temporal fragmentation** visible (discrepancies between planned and actual) but leaves users to infer the contextual causes of those discrepancies without support.

2.2.2 Exist.io

Exist.io [23] adopts a different strategy by aggregating personal data from multiple services and automatically identifying correlations between variables such as sleep, activity, and mood (see Figure 2.2). Since its introduction, it has developed into a comprehensive personal analytics platform organised around "attributes", individual variables that users track over time. These attributes may be populated automatically through integrations or entered manually and are typically recorded at a daily level.

Evaluated against our context framework, Exist.io captures **activity** (via fitness trackers), **location** (via Swarm and calendar integrations), and **internal state** (via manual mood entries).

Figure 2.2: Exist.io's correlation dashboard. The figure shows automated correlations between sleep, activity, and mood. Source: `https://exist.io`

However, it captures no **social context** and no **objects**. More critically, daily aggregation introduces what we term **temporal shallowness**: intra-day dynamics, essential for understanding behaviour-environment interactions, are lost. Reported correlations are presented without contextual grounding, and the system does not expose data quality, treating all relationships as equally reliable regardless of underlying sparsity or noise.

Table 2.1 summarises the range of services integrated into Exist.io, illustrating the breadth of its data aggregation approach.

Table 2.1: Services integrated by Exist.io

Category	**Integrated Services**
Fitness and Activity	Fitbit, Oura, Misfit, Withings, Garmin, Health Connect, Apple Health, Google Fit
Productivity	RescueTime, GitHub (commits), Todoist (tasks completed)
Location and Events	Swarm by Foursquare, Google Calendar, iCal
Media Consumption	Last.fm, Spotify (via Last.fm), Apple Music (via Last.fm)
Weather	Apple Weather
Social	Mastodon

2.2.3 Mainstream Calendar Platforms

Mainstream platforms such as Google Calendar, Apple Calendar (Figure 2.3), and Microsoft Outlook have begun to incorporate features that gesture toward reflective use, including basic summaries of time allocation and limited integrations with activity data. However, these additions remain peripheral to their primary function as scheduling tools.

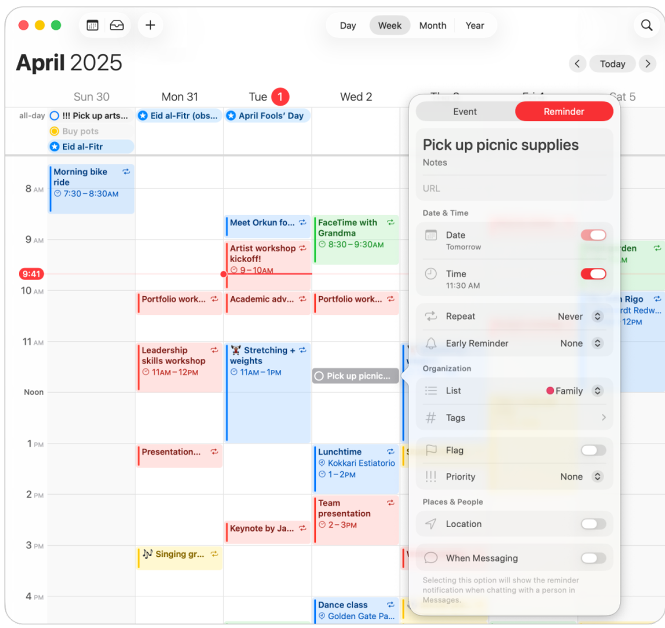


Figure 2.3: Apple Calendar's conventional event-based interface. The figure illustrates the assumption that scheduled events adequately represent activity. No contextual dimensions—location beyond event venues, social context, internal state, or objects—are captured. Source: https://support.apple.com

Evaluated against our context framework, mainstream calendars capture only scheduled **activity** and, optionally, **location** (event venues). They capture no **social context** beyond attendee lists, no **internal state**, and no **objects**. Where reflective features exist, they are typically coarse and aggregated, providing only high-level summaries that overlook the structure of everyday life, where activities overlap, plans evolve, and significant portions of time remain unstructured. This reflects a foundational assumption: that scheduled events are an adequate proxy for lived activity, which systematically excludes the contextual richness necessary for meaningful reflection.

2.2.4 Limitations of Current Calendar-Dashboard Systems

Across the systems reviewed, three interconnected limitations emerge that we will use throughout the remainder of this chapter:

Temporal fragmentation. Most systems represent time as discrete, non-overlapping units. While this simplifies interaction, it fails to capture the continuity and overlap inherent in everyday activity. Interruptions, multitasking, and deviations from plans are either compressed or omitted entirely. This reflects a deeper mismatch: the structure imposed by calendar representations does not align with how activities are actually experienced.

Contextual sparsity. Even in systems that aggregate multiple data sources, contextual dimensions, as defined by our five-dimensional framework, remain only partially represented. When related data is distributed across separate systems, questions such as whether late-night phone use affects next-day well-being become structurally unsupported.

Quality opacity. None of the reviewed systems makes data quality explicit at the point of interpretation. Missing entries, delayed responses, and unreliable measurements are treated identically to complete and accurate data. This **quality opacity** obscures the reliability of observed patterns, making it impossible to distinguish genuine behaviour from artefacts of data collection. Given the known variability in both sensor data and self-reports [9, 11], this omission fundamentally limits the validity of user interpretations.

2.3 Personal Informatics Systems

2.3.1 Conceptual Foundations

Personal informatics (PI) systems have developed from simple self-tracking tools, such as step counters and diet logs, into more complex platforms that integrate multiple streams of contextual information, including activity, location, and social interaction. Their core purpose remains consistent: to support the collection, monitoring, and interpretation of personal data in ways that enable self-knowledge and behaviour change [22].

More recent work has shifted attention from data collection to reflection, understood as the process through which users interpret their data and relate it to their experiences. Elsden et al. [21] frame this in terms of three interrelated dimensions of selfhood: the experiential self (ongoing lived experience), the remembered self (retrospective interpretation), and the narrating self (the construction of personal identity over time). From this perspective, personal data becomes meaningful not as isolated measurements but as material for constructing coherent accounts of one's life.

This view is further developed in Epstein et al.'s [22] "Lived Informatics" model, which characterises self-tracking as an ongoing, situated practice embedded in everyday routines. Reflection is not a discrete step but a continuous process of engaging with data, questioning patterns, and relating observations to personal goals and circumstances.

Despite these conceptual advances, their translation into practical systems remains limited. Cho et al. [15], in an analysis of personal informatics applications, found that most tools provide limited support for contextualisation or goal-oriented interpretation. Users are typically presented with data but receive little assistance in understanding its implications. As a result, the burden of interpretation remains largely with the user, a gap that worsens when data lacks the temporal and contextual structure necessary for coherent narrative construction.

2.3.2 Physical and Mental Wellbeing Applications

Digital tools for physical and mental well-being represent one of the most visible applications of personal informatics. These systems combine passive sensor data, such as heart rate, movement, and sleep patterns, with active user inputs, including mood logs and self-reported activities.

Physical Wellbeing

Health and fitness platforms such as Google Fit (Figure 2.4) and Apple Health are widely used to track activity levels, physiological signals, and behavioural trends [17].

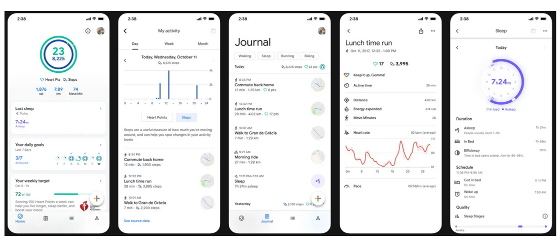


Figure 2.4: Google Fit's activity summary interface. The figure illustrates the platform's focus on aggregated metrics, step counts, heart rate, and sleep duration, presented as isolated trends. No integration with mood, social context, or planned activities is provided. Source: https://www.google.com/fit

Evaluated against our context framework, these platforms capture **activity** (step counts, exercise types) and **internal state** only indirectly through heart rate. They capture no **social context**, and no **objects**. Their primary limitation is **contextual sparsity**: physical activity is recorded in isolation from the situations, work stress, social settings, and environmental conditions that shape it. A user can see that the step count decreased but cannot determine whether this reflects illness, workload, or a conscious choice.

Mental Wellbeing

Applications targeting mental well-being typically rely more heavily on self-reported data, such as mood diaries, journaling, or stress ratings, sometimes complemented by lightweight sensor inputs. These tools aim to help users identify emotional patterns and develop greater awareness of their internal states [43].

Daylio [19] is a representative example. It enables users to log daily mood using simple visual scales, alongside activities such as exercise, social interaction, or leisure (Figure 2.5).

Figure 2.5: Daylio's mood tracking interface. The figure shows how emotional states are reduced to five-point scales with activity tagging. No integration with passive sensor data is provided, resulting in **contextual sparsity**: mood is recorded in isolation from the objective behavioural context that could explain it. Source: https://daylio.net

Evaluated against our context framework, Daylio captures **activity** (via manual tagging) and **internal state** (mood ratings). It captures no **location**, no **social context**, and no **objects**. Data collection relies entirely on manual input, with no integration of passively sensed contextual data that could enrich interpretation. This creates **temporal shallowness**: entries document only the moments users choose to record, introducing systematic bias and omitting continuity of experience. A user can see that mood was low on a particular day but cannot connect this to objective measures of sleep, activity, or social interaction from the same period.

2.3.3 Digital Behaviour and Attention

A growing class of personal informatics applications focuses on digital well-being, particularly in relation to smartphone and application usage. These systems monitor interaction patterns, such as screen time, device unlock frequency, and application usage, to support reflection on digital habits. ActionDash (Figure 2.6) provides detailed summaries of user interaction, while commercial platforms such as Google Digital Wellbeing and Apple Screen Time offer similar functionality.

Empirical studies suggest that self-monitoring can improve awareness of smartphone use [32].

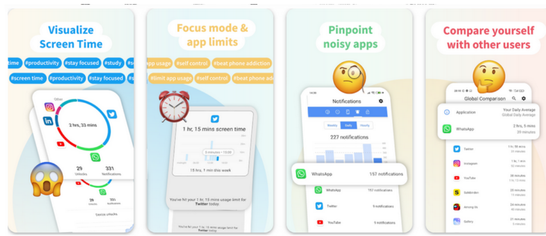


Figure 2.6: ActionDash's screen time summary. The figure illustrates detailed metrics of digital behaviour, app usage, and notifications, presented in isolation. No integration with physical activity, mood, or planned activities is provided. Source: https://actiondash.com

However, evaluated against our context framework, these platforms capture only **objects** (devices and applications) and, indirectly, **activity** (app categories). They capture no **location**, no **social context**, and no **internal state**. Their limitations are threefold: **contextual sparsity** (digital behaviour disconnected from physical and emotional dimensions), limited personalisation (generic facts rather than tailored guidance), and what we term the **awareness–action gap**: users are shown *what* they are doing but given little support in translating awareness into sustained change.

2.3.4 Environment and Context Awareness Platforms

In contrast to domain-specific applications, environment and context-aware platforms aim to capture a broader, continuous view of human behaviour. These systems typically operate in the background, using smartphone and wearable sensors to collect high-resolution data on movement, location, device interaction, and, in some cases, physiological and social signals.

AWARE [24] (Figure 2.7) is a prominent example. It transforms a smartphone into a sensing platform capable of capturing detailed contextual data, combining passive sensing with occasional user input.

Evaluated against our context framework, AWARE captures **location**, **activity** (via motion and app usage), and **objects** (device interactions). However, it captures no **social context** and no **internal state** beyond what is manually entered through surveys. Its primary limitation is not **contextual sparsity** but **reflective opacity**: while data is rich, the platform lacks a user-facing reflective interface. Participants are positioned as data sources rather than beneficiaries of insight.

Beiwe [48] extends this approach within digital phenotyping, focusing on continuous measurement of behavioural markers relevant to mental and physical health. It collects high-resolution data from smartphone sensors, including GPS, accelerometer, call logs, and voice samples (Figure 2.8). Designed for clinical research, Beiwe prioritises data security and configurability but provides no feedback to participants.

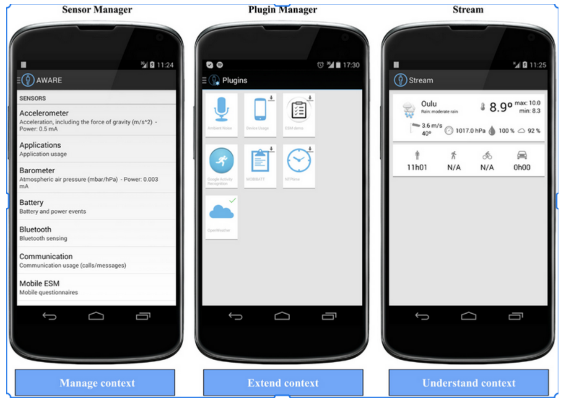


Figure 2.7: AWARE's data collection interface. The platform captures rich multimodal data—location, motion, Bluetooth, and app usage—but provides no user-facing reflective interface. This illustrates the data-rich but reflection-poor side of the utility gap. Source: [24]

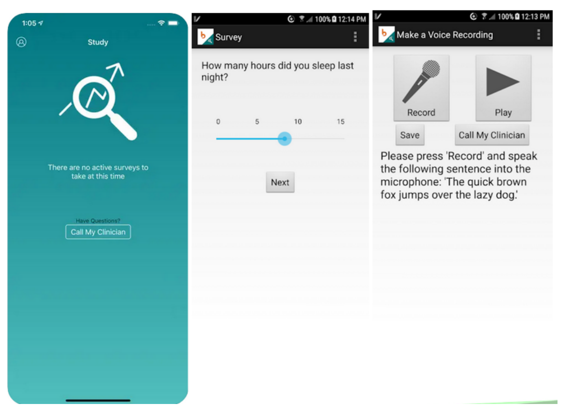


Figure 2.8: Beiwe's research configuration interface. Like AWARE, the platform prioritises data collection over user reflection, creating a utility gap where rich data is collected but no value is returned to participants. Source: [48]

RADAR-base [53] supports large-scale, real-time data collection in clinical contexts. It is an open-source platform for remote monitoring of chronic diseases. By integrating data from wearables, smartphones, and clinical systems, researchers can continuously monitor patients (Figure 2.9). Just like AWARE and Beiwe, RADAR-base focuses on data collection for research rather than participant reflection.

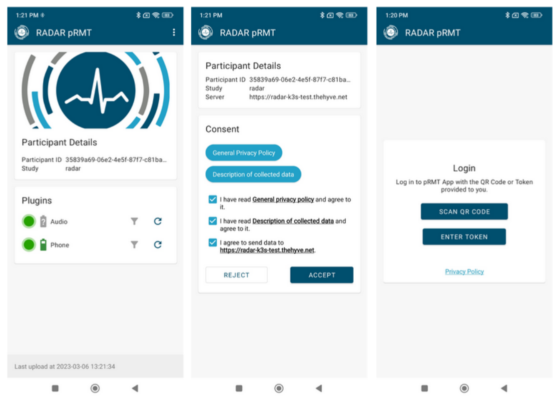


Figure 2.9: RADAR-base's clinician dashboard. The platform enables continuous monitoring of chronic conditions but lacks participant-facing reflection tools. Source: [53]

DEMAware2 [55] is an ambient-assisted living framework for dementia care. It integrates home sensors, wearables, and audio-visual analysis to build detailed semantic models of patient activity. Clinicians can monitor behavioural patterns and detect early symptoms (Figure 2.10).

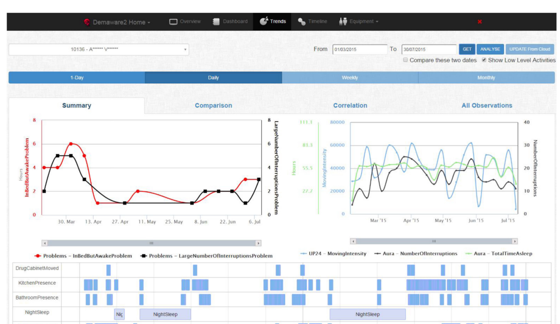


Figure 2.10: DEMAware2's clinical dashboard. While rich in sensor integration, the platform raises privacy concerns and provides no participant-facing interface. Source: [55]

Across these research platforms, a consistent pattern emerges. They address **contextual sparsity** by capturing rich multimodal data across multiple dimensions. However, they introduce a complementary limitation: **reflective opacity**. Even when data is collected at high resolution, no mechanism exists for participants to engage with or derive meaning from that information. This creates the first side of what we term the **utility gap**: systems that collect rich data provide no reflective value to users.

2.3.5 Habits and Routines Applications

A final category of applications focuses on habit formation and behaviour change through structured programs, reminders, and motivational design. These tools aim to translate awareness into action, addressing the awareness–action gap identified above.

Fabulous [1] adopts an approach to behaviour change by guiding users through structured ”journeys”—step-by-step programs designed to help them build healthy routines incrementally. It maintains user engagement through motivational design and gamification elements. Its key features include personalised coaching of users for common goals such as improved sleep, mindfulness practice, and productivity enhancement sessions. Its reminders and behavioural nudges support users' habit formation through the highly engaging visual design and gamification mechanics, which encourage continued participation.

By prompting users to explicitly ”promise” themselves that they will complete tasks and using specific sensory rituals—such as unique sounds or countdowns—it creates a psychological space that transforms mundane tasks into meaningful rituals, thus showing that lasting behaviour change does not come from a person's willpower alone but from carefully designed systems that offer regular support and reinforcement.

Journey [2] serves as a counterpoint to action-oriented trackers by focusing on the ”inner work” of habit formation through self-reflection. Instead of pushing for task completion, it uses journaling to help users process their thoughts and identify the underlying emotions that drive their behaviour. By providing a ”safe container” through secure, multimedia-rich entries, it encourages users to develop a personal narrative that links their daily habits to long-term mental clarity and emotional regulation.

The main challenge for Journey is the significant mental effort it demands from the user. While the app automates simple data like location or weather, the core value still relies on manual, deep-dive writing, which can lead to ”journaling fatigue” when life gets busy. Additionally, the insights it uncovers tend to exist in a vacuum; because the app has limited integration with broader clinical health platforms or physical sensors, it can be difficult for users to see the direct, objective link between their emotional reflections and their physical health data.

Evaluated against our context framework, these platforms capture **activity** (habit completion, journal entries) and, through self-report, **internal state**. They capture no **location**, no **social context**, and limited **objects**. Their strengths lie in supporting reflection and behaviour change, addressing the awareness–action gap that other systems neglect. However, they exemplify the second side of the **utility gap**: they provide reflective value but collect shallow, primarily self-reported data, with no integration of passive sensing that could contextualise or validate user inputs.

[1] Fabulous: https://thefabulous.co/
[2] Journey: https://journey.cloud/

2.3.6 Summary: Common Patterns Across Systems

Across the landscape of personal informatics systems, a systematic pattern emerges. Table 2.2 summarises how each system category addresses—or fails to address—the five dimensions of context.

Table 2.2: Contextual Coverage Across System Categories

System Category	Activity	Location	Social	Internal	Objects
Mainstream Calendars	Scheduled	Partial	—	—	—
Activity River	Manual	—	—	—	Manual
Exist.io	Auto/Manual	Auto	—	Manual	—
Physical Wellbeing (Google Fit)	Auto	—	—	Indirect	—
Mental Wellbeing (Daylio)	Manual	—	—	Manual	—
Digital Behaviour (ActionDash)	Indirect	—	—	—	Auto
Research Platforms (AWARE)	Auto	Auto	—	Manual	Auto
Habits Apps (Fabulous, Journey)	Manual	—	—	Manual	—
Our System	Auto/Manual	Auto	Auto	Manual	Auto

Three interconnected limitations recur across all categories:

1. **Temporal fragmentation**: Systems either aggregate data to coarse daily levels (Exist.io) or depend on user-initiated logging (Daylio), losing within-day dynamics essential for understanding behaviour-environment interactions.
2. **Contextual sparsity**: No system captures all five contextual dimensions. Even research platforms that collect rich sensor data omit **social context** and **internal state**, while consumer applications that capture internal state do so without objective behavioural correlates.
3. **Quality opacity**: No system makes data reliability visible to users. Missing entries, delayed responses, and unreliable measurements are indistinguishable from valid observations.

These limitations converge on a structural problem: the **utility gap**. Research platforms collect rich data but provide no reflective value to participants. Consumer applications provide reflective value but collect shallow data. Mainstream calendars provide temporal structure but capture only planned events. No existing system combines the data richness and quality tracking of research platforms with the reflective temporal structure that calendars provide.

2.4 Synthesis: Toward an Integrated Solution

The preceding reviews reveal a systematic pattern. Each family of systems addresses part of what is needed for meaningful self-reflection, but their limitations are complementary. Calendar systems offer temporal structure but lack rich behavioural data. Personal informatics systems collect rich behavioural data but lack integrated temporal interfaces and exhibit contextual sparsity. Neither family addresses quality opacity.

2.4.1 The Utility Gap

We term this structural problem the **utility gap**:

> *Systems that collect rich data provide no reflective interface; systems that provide reflective interfaces collect shallow data.*

Table 2.3 summarises this gap across system types, introducing the three diagnostic concepts developed throughout this chapter: **temporal fragmentation**, **contextual sparsity**, and **quality opacity**.

Table 2.3: The Utility Gap: Diagnosing System Limitations

System Type	Temporal Structure	Contextual Coverage	Quality Visibility	Reflective Integration
Mainstream Calendars	Strong	1 dimension	None	Low
Consumer PI (Exist.io, Daylio)	Weak (daily/manual)	2-3 dimensions	None	Moderate
Research Platforms (AWARE, Beiwe)	Continuous	3-4 dimensions	Backend only	None
iLog (Foundation)	Continuous	4 dimensions	Tracked	Basic dashboards
Our System	Continuous (Calendar)	5 dimensions	Visible	High

2.4.2 The iLog Foundation

The iLog platform [58] provides the sensing infrastructure on which this thesis builds. It captures:

- **Passive sensor data:** Location (GPS), motion (accelerometer), Bluetooth proximity, Wi-Fi connectivity, and device usage.
- **Structured self-reports:** Experience sampling method (ESM) surveys capturing the five components of context and contextual sensor information.
- **Quality metrics:** Sensor availability, response latency, and data completeness tracked at collection time.

iLog already addresses the data side of the utility gap. It provides the data richness of research platforms and the quality tracking that neither research platforms nor consumer applications make visible. It also includes basic participant-facing dashboards—a step beyond AWARE or Beiwe.

2.4.3 Why a Calendar-Based Interface?

The calendar provides a natural solution to these limitations. It offers a familiar temporal structure that aligns with how people naturally experience time—dividing life into days, weeks, and months. This addresses **temporal fragmentation** by providing a continuous timeline onto which heterogeneous data streams can be aligned.

By integrating iLog's rich, quality-tracked data within a calendar interface, we can simultaneously address **contextual sparsity** (by visualising multiple dimensions together) and **quality opacity** (by making reliability indicators visible alongside the data). The calendar becomes not merely a scheduling tool but a unified temporal framework for integrating multimodal personal data.

2.4.4 From Gaps to Requirements

The synthesis above yields a clear set of requirements that no existing system meets:

1. **Data richness:** The system must capture multimodal behavioural data alongside subjective experience. This is provided by iLog.
2. **Quality awareness:** Data reliability must be visible to users.
3. **Temporal structure:** The system must align planned activities, observed behaviour, and contextual information within a continuous timeline. This requires a calendar-based interface.
4. **Contextual integration:** Data must be organised not just along a timeline but also in relation to the five dimensions of context. This requires visualisation that makes these dimensions visible together.
5. **Reflective interface:** The system must support meaning-making, not just data presentation. It must help users connect different aspects of their lives, understand relationships between actions and outcomes, and construct coherent narratives from their data.

2.4.5 The Present Work

These requirements converge on the central contribution of this thesis: extending iLog with a quality-aware calendar dashboard that makes rich, quality-tracked personal data meaningful within a temporal framework that mirrors how people naturally experience their lives.

iLog provides the foundation—data richness. What is missing is a reflective layer that integrates this multimodal data into a coherent temporal narrative, aligns planned activities with observed behaviour across all five contextual dimensions, and makes quality visible in a way that supports interpretation.

This thesis addresses this missing layer through the design, implementation, and evaluation of a quality-aware calendar dashboard built on the iLog sensing infrastructure. Drawing on the big-thick data framework established in Section 2.1, we integrate multimodal sensor streams with structured self-reports within a unified temporal representation. The five-dimensional context model developed in Section 2.1.1 provides the representational structure that aligns these heterogeneous data types.

By extending iLog with a reflective calendar interface that makes both data and its quality visible, and that aligns planned events with observed behaviour across all five contextual dimensions, we aim to bridge the utility gap and support meaningful, quality-aware reflection. The following chapters describe the design, implementation, and evaluation of this system.

Chapter 3

Modelling Personal Data

3.1 Introduction

Smart devices have become deeply integrated into the fabric of everyday life, generating continuous, multi-modal streams of personal data that reflect our behaviours, contexts, and routines. These streams and self-reports capture what researchers have termed ”big-thick” data: information that is both large in scale and rich in contextual meaning [34]. Beyond its volume and variety, this data forms a continuous temporal record of daily life, offering potential for self-understanding and behavioural insight.

Realising this potential, however, requires addressing two fundamental challenges. First, raw sensor data and self-reports are heterogeneous in format, sampling rate, and reliability. Second, to be interpretable, this data must be organised within a framework that preserves its temporal structure and contextual meaning. This chapter presents the conceptual and formal foundations for addressing these challenges.

As described in 2.1.1, we apply the concept of context, which enables the integration of objective sensor traces with subjective self-reports. We then formalise a data quality framework that defines metrics for assessing the reliability of both self-reported and passively sensed data. These frameworks together provide the theoretical and methodological foundation for the calendar dashboard described in subsequent chapters.

3.1.1 A Five-Dimensional Context Model

To operationalise the integration of big and thick data, in this work, we adopted and extended a structured representation of context originally proposed by Giunchiglia et al. [28] and refined in subsequent work [42]. The original model comprised four core dimensions: *what* (activity), *where* (location), *when* (time), and *who* (agent). We extended this model with a fifth dimension: *with what*—the set of artefacts or tools with which the user is interacting.

The incorporation of objects and tools as a distinct contextual dimension reflects the growing recognition that tools are not merely external aids but actively shape cognitive processes and behaviour [49]. By explicitly modelling the artefactual dimension, the framework captures the mediating role of objects in structuring action, attention, and social interaction. This five-dimensional representation is both compact, reducing the complexity of real-world circumstances to a finite set of facets, and expressive, capturing the essential elements that define a human activity.

Formally, we define the **situational context** C as a tuple:

$$C = \langle WE, WA, WI, WO, WU \rangle \tag{3.1}$$

where each dimension is defined as follows:

- WE (spatial context): a linguistic description of the location where a person is situated. This can be obtained from hardware sensors (e.g., GPS or Wi-Fi) or from self-reports. In iLog, this is labelled through the question ”WhEre are you?”
- WA (activity/event context): a description of the activities currently performed by the person, which may occur sequentially or in parallel. Sources include accelerometers (for physical activity), software logs (for online activity), or self-reports from users. In iLog, this is captured by the question “WhAt are you doing?”
- WI (internal context): a description of the internal states or processes occurring inside the person (e.g., mood). In our setup, this is represented by the question “What mood are you In?”
- WO (social context): a description of the people accompanying the person, if any. This can be inferred from Bluetooth connections or user self-reports. In iLog, this is asked via the question, “WhO are you with?”
- WU (object/tool context): a description of the objects or tools used by the person. These may be detected using sensors (Bluetooth, RFID, Wi-Fi) or elicited through user self-reports. For example, the question ”Which Utensils are you using?” captures this dimension.

This five-dimensional model not only formalises the multiple facets of personal situation context but also provides a practical framework for integrating diverse data sources within our system. As illustrated in the example in 3.2, the different parts of context can be instantiated using sensor data, user self-reports, or a combination of both, giving a comprehensive representation of everyday situations.

3.1.2 Life Sequences and Temporal Structure

A person's life is not composed of isolated situational contexts but a sequence of them over time. We define a *life sequence* S as an ordered sequence of situational contexts covering a period of interest:

$$S(p) = \langle C_1(p), C_2(p), \ldots, C_n(p) \rangle \tag{3.2}$$

where $C_i(p)$ is the ith situational context of person p. For modelling purposes, we assume that at any given time a dominant situational context can be identified. While human activity may involve overlapping physical, digital, and social processes, the model represents the context most salient to the activity under analysis. At any moment, a person can be in only one primary location; thus, the life sequence is a non-overlapping temporal partition. This simplifying assumption ensures temporal consistency and computational tractability while preserving expressive power.

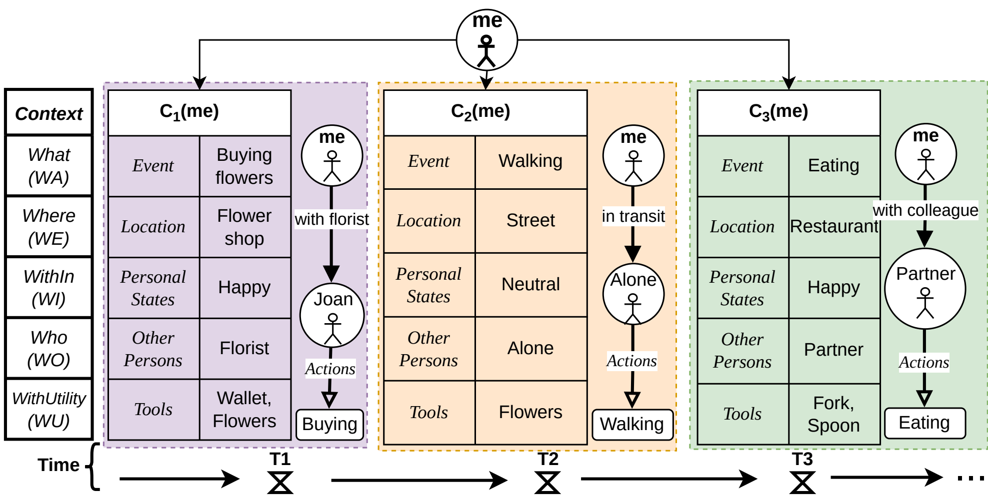


Figure 3.1: An example life sequence showing three situational contexts across an evening: (T1) buying flowers, (T2) walking, (T3) dining.

Figure 3.1 illustrates a life sequence from an evening. The activities unfold as follows:

- **T1:** At a flower shop, purchasing flowers from a florist, using a wallet, feeling happy.
- **T2:** Walking on the street toward a restaurant, alone, holding flowers, feeling neutral.
- **T3:** At the restaurant, dining with a partner, using cutlery, feeling happy.

Mapping T1 to the five-dimensional model:

- WA (activity): *buying flowers*
- WE (location): *flower shop*
- WI (internal state): *happy*
- WO (social context): *florist*
- WU (objects): *wallet, flowers*

This example demonstrates how the context model can represent the richness of everyday experience. In practice, such representations are constructed from real sensor data and self-reports.

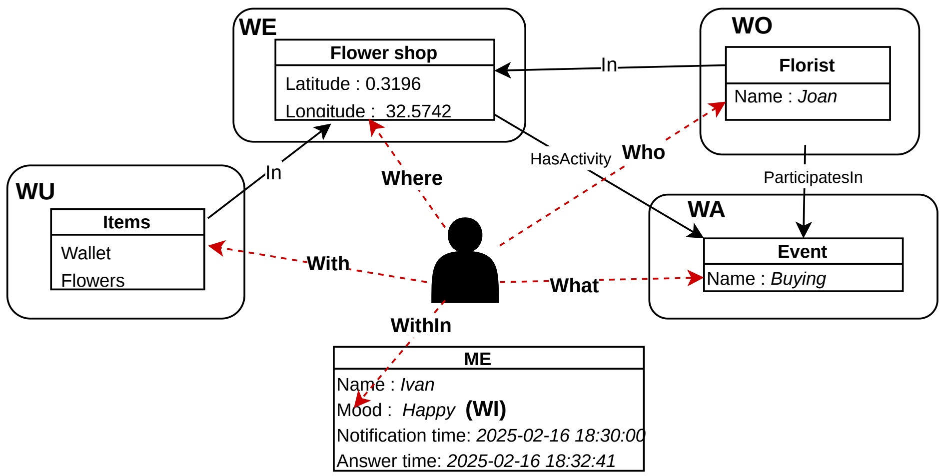


Figure 3.2: Knowledge graph representation of situational context T1 from Figure 3.1. Nodes represent entities; edges represent relations; attributes provide additional detail.

Figure 3.2 shows how this situational context can be represented as a knowledge graph [29]. The graph comprises: (i) a subgraph for each contextual dimension; (ii) nodes for entities (person, location, objects); (iii) attributes and values for each node (name, mood, timestamps); and (iv) edges representing relations between entities (location containment, co-presence, activity participation).

Table 3.1 shows a representative sequence extracted from the MakOne dataset [36], demonstrating how the model captures authentic daily activity.

3.2 Data Quality Framework

A data-driven system is only as reliable as the data it processes. Data quality includes dimensions such as accuracy, completeness, consistency, timeliness, and fitness for purpose [5]. In personal informatics, maintaining high-quality data is challenging, as data is collected in dynamic,

Table 3.1: Example life sequence from MakOne dataset (participant M008, 2024-04-15)

Time	WA (Activity)	WE (Location)	WI (Mood)	WO (Social)	WU (Objects)
08:15–09:00	Commuting	Matatu (Bus)	Neutral	Alone	Phone, Backpack
09:00–11:30	Lecture	Classroom	Focused	Classmates	Laptop, Notebook
11:30–12:15	Eating	University Cafeteria	Happy	Friends	Phone, Food tray
12:15–14:00	Studying	University Library	Focused	Alone	Laptop, Textbook
14:00–16:00	Group Work	Study Room	Neutral	Colleagues	Whiteboard, Laptop
16:00–17:30	Socialising	Campus Common Area	Happy	Friends	Phone, Coffee
17:30–18:30	Commuting	Walking	Tired	Alone	Phone, Headphones
18:30–20:00	Relaxing	Home	Neutral	Family	TV, Phone

heterogeneous, and often noisy environments from sources that are prone to error. Self-reports are particularly susceptible to recall bias, inattention, and response fatigue [9, 25]. The quality of sensor data is often affected by hardware constraints, signal loss, and missing values resulting from device settings or user permissions [11].

In this thesis, data quality is considered a fundamental issue. We incorporate data quality monitoring measures across data collection, processing, and presentation.

3.2.1 Data Sources and Their Quality Challenges

Our system collects data from two primary sources, each with distinct quality challenges:

Self-Reported Data

Quality issues in self-reports stem primarily from human factors. Frequent questioning can become intrusive, leading users to ignore prompts, provide rushed answers, or disable the application all together [9, 11]. Recall bias affects responses completed long after the event [51]. Response fatigue reduces engagement over time [25], whereas user inattention leads to careless or inconsistent responses [25, 7, 8].

Studies have shown that response timing can serve as a useful, non-intrusive proxy for answer quality in mobile self-report settings. Bison et al. [9, 8] found that shorter response latencies and answer durations correlate with higher answer accuracy. In this work, we operationalise answer quality not only as availability (whether a question was answered) but also as timeliness and engagement, using response latency and completion time as indicators.

Sensor Data

Sensor data quality is influenced by technical and environmental factors: hardware limitations, battery optimisation, signal interference, and user permission changes. Common problems affecting sensor data quality include missing samples, irregular sampling intervals, and noisy measurements [39]. Missing data is a significant problem because gaps can disrupt the continuity needed for accurately reconstructing context, which can then introduce systematic bias [11].

A key distinction from self-reports is that missing sensor data is often invisible to the user, making systematic detection and handling essential for any scalable system that incorporates human data.

3.2.2 Quality of Question Delivery

The reliability of the data collection process depends on the integrity of question delivery. Even if users are willing to respond, data cannot be collected if questions are not sent, delivered correctly, or duplicated unnecessarily. Faulty delivery can distort datasets, burden participants, and undermine temporal consistency.

To formalise these aspects, we introduce the following measurements:

Let $Q(p)$ denote the total number of questions the system attempts to send to participant p:

$$Q(p) = Q_R(p) + Q_L(p) \tag{3.3}$$

where $Q_R(p)$ are questions successfully received and $Q_L(p)$ are questions lost due to delivery failures.

The aggregate questions received across all P participants is:

$$Q_R = \sum_{p=1}^{P} Q_R(p) \tag{3.4}$$

Similarly, the total number of questions lost is given by:

$$Q_L = \sum_{p=1}^{P} Q_L(p) \tag{3.5}$$

For temporal monitoring, let $Q(p, I)$ be the number of questions delivered to participant p during interval I. Therefore, the number of questions, Q(I), delivered in interval I are:

$$Q(I) = \sum_{p=1}^{P} Q(p, I) \tag{3.6}$$

Average questions received and lost per participant provide normalised views of the system performance:

$$\overline{Q_R} = \frac{1}{P} \sum_{p=1}^{P} Q_R(p), \quad \overline{Q_L} = \frac{1}{P} \sum_{p=1}^{P} Q_L(p) \tag{3.7}$$

These measures enable monitoring of question delivery quality at both system and participant levels, helping identify whether delivery problems are systemic or participant-specific.

3.2.3 Quality of Self-Reported Answers

Let $A(p, q)$ indicate whether participant p provided a valid answer to question q within the allowed time window:

$$A(p, q) = \begin{cases} 1 & \text{if answered within the allowed time window} \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

Let Q_p denote the set of unique questions delivered to participant p.

The number of valid answers provided by participant p is:

$$A(p) = \sum_{q \in Q_p} A(p, q) \tag{3.9}$$

The total number of valid answers across the experiment is given by:

$$A = \sum_{p=1}^{P} A(p) \tag{3.10}$$

Temporal Responsiveness

Consider $T_R(p, q)$ as the response latency for question q answered by participant p, measured from delivery to opening, and $T_C(p, q)$ as the completion time from opening to submission. The total response latency and total completion time for participant p are:

$$T_R(p) = \sum_{q \in Q_p} T_R(p, q), \quad T_C(p) = \sum_{q \in Q_p} T_C(p, q) \tag{3.11}$$

Total interaction time, D(p), captures both attention delay and interaction effort:

$$D(p) = T_R(p) + T_C(p) \tag{3.12}$$

Temporal Distribution of Responses

For a time interval I (e.g., hour, day), let $Q_p^{(I)} \subseteq Q_p$ be questions delivered to the participant p during I. The number of valid answers in the interval I is:

$$A_I = \sum_{p=1}^{P} \sum_{q \in Q_p^{(I)}} A(p, q) \tag{3.13}$$

This allows for the analysis of engagement patterns over time, including changes in response rates throughout the day.

3.2.4 Quality of Sensor Data

Let $S(p, s)$ be the expected number of samples for a participant p from the sensor s (based on configured sampling frequency) and $E(p, s)$ the observed samples. With P participants and M sensor modalities, the total sensor data collected per participant is:

$$E(p) = \sum_{s=1}^{M} E(p, s) \tag{3.14}$$

Total sensor data across the experiment is:

$$E = \sum_{p=1}^{P} \sum_{s=1}^{M} E(p, s) \tag{3.15}$$

Missing sensor data for participant p is:

$$S_L(p) = \sum_{s=1}^{M} (S(p,s) - E(p,s)) \tag{3.16}$$

The completeness ratio across all participants is:

$$C_S = \frac{\sum_{p=1}^{P} \sum_{s=1}^{M} E(p,s)}{\sum_{p=1}^{P} \sum_{s=1}^{M} S(p,s)} \in [0,1] \tag{3.17}$$

Participant-level completeness is:

$$C_S(p) = \frac{\sum_{s=1}^{M} E(p,s)}{\sum_{s=1}^{M} S(p,s)} \tag{3.18}$$

High-quality sensor data requires completeness (minimal missing intervals), continuity (consistent sampling rate), and accuracy (faithful representation of the underlying situation). These metrics enable systematic identification of recurring issues—data loss from network problems, power-saving restrictions, or permission changes—and support targeted interventions to maintain dataset integrity.

3.2.5 Example: Measuring GPS Sensor Data

To illustrate how these metrics are applied, we can consider a scenario where a participant takes part in an experiment in which the GPS sensor is set to record location every 10 seconds over one hour.

- **Expected number of GPS samples to collect:**

$$S(p) = 360 \text{ (60 minutes} \times \text{6 samples per minute)}$$

- **Actual number of GPS samples collected:**

$$E(p,s) = 300$$

- **Sample Loss:**

$$S_L(p) = S(p) - E(p,s) = 360 - 300 = 60$$

The completeness ratio for this participant is:

$$C_S(p) = \frac{E(p,s)}{S(p)} = \frac{300}{360} \approx 0.83$$

This value shows that roughly 17% of the GPS data was absent, potentially attributable to signal degradation, device deactivation, or application disruptions. These metrics are valuable for researchers, enabling them to pinpoint data deficiencies, refine data collection methodologies, or supplement absent data with alternative sensors, such as Wi-Fi positioning or accelerometer data, thereby ensuring the preservation of a high-quality dataset.

In the proposed calendar-dashboard system, these metrics are calculated and embedded as part of the data displayed in the different views; for example, Figures 5.3 and 5.6

3.3 Chapter Summary

This chapter has presented the conceptual and formal foundations for modelling personal data in a quality-aware system. We introduced a fifth dimension to the already existing four-dimensional context model that provides a structured representation for integrating heterogeneous data streams while preserving semantic meaning. We formalised a data quality framework with metrics for question delivery, answer quality, and sensor completeness. These frameworks together establish the theoretical and methodological foundation for the calendar dashboard described in the following chapters. The next chapter describes how these models are operationalised in the dashboard interface.

Chapter 4

Dashboard Calendar

4.1 Awareness and Reflection via the Dashboard

Mobile and wearable technologies now collect multimodal personal data continuously: sensor streams, self-reports, and contextual metadata. This data offers unprecedented opportunities for insight, but its value depends entirely on how it is structured, interpreted, and communicated to users. Raw data alone does not produce understanding; it must be transformed into representations that align with how people naturally perceive, remember, and reason about their lives.

Historically, individuals used analogue artefacts such as journals and calendars to document experiences and reflect on life events [44]. These tools provided temporal organisation but lacked integration across behavioural, contextual, and physiological dimensions. Digital dashboards can extend this tradition by aggregating heterogeneous data streams into unified visual environments. By combining calendar views, timelines, graphs, and contextual annotations, they enable structured reflection across multiple temporal scales [40, 6, 46].

A central challenge in personal informatics is how fragile the data foundation is. Despite advances in designing for user reflection, demonstrated by Choe et al. [16] and Epstein et al. [22], the potential for meaningful sense-making is often undermined at the source. Data from ubiquitous sensors and self-reports is routinely compromised by gaps, noise, and contextual ambiguity, problems that are frequently invisible to users. This creates a dual risk: it leads to misinterpretation of personal patterns and erodes the trust necessary for long-term system engagement.

This thesis seeks to address this limitation by integrating data-quality awareness directly into a calendar-based dashboard. Rather than treating reliability as a purely back-end concern, the system exposes indicators of completeness, timing, and consistency within the same interface used for behavioural reflection. The dashboard provides visual indicators of data quality, enabling users to reflect not only on their behavioural patterns but also on the reliability of the data underlying those patterns.

The dashboard serves a dual purpose: it is both a tool for participants to understand their daily lives and an observational tool for researchers to monitor data quality and participant engagement throughout experiments. By linking personal experiences with data-quality feedback, the system fosters a deeper understanding of when, how, and why personal data may vary, empowering individuals to make more informed interpretations and sustain high-quality data-collection practices over time.

This chapter presents the design of the proposed calendar dashboard system. Section 4.2 introduces the conceptual framework of life monitoring, including the notion of life awareness and the temporal model (past, present, future) that organises it. Section 4.3 then describes the dashboard interface, detailing how these concepts are operationalised through user roles, visual encoding strategies, and integrated components.

4.2 Life Monitoring

4.2.1 Life Awareness

Life awareness refers to an individual's capacity to reflect on past experiences, monitor present routines, and anticipate future activities. It is a continuous process through which people maintain understanding of how their actions, contexts, and emotions interrelate over time. Within personal informatics, life awareness emerges when raw data collected about people and contextual cues are transformed into meaningful insights that support reflection, sense-making, and self-knowledge [40, 22].

This process can be understood as a continuous loop between data, reflection, and action. Individuals gather personal data, examine it to uncover patterns, and use this understanding to inform future behaviour. As Li et al. [40] proposed in the *Stage-Based Model of Personal Informatics*, this unfolds across stages from preparation and collection to integration, reflection, and action. Epstein et al. [22] further emphasise that self-tracking and reflection are not isolated activities but part of an evolving relationship between the individual and their data.

Our context-aware dashboard operationalises life awareness by translating raw personal data into interpretable narratives. It integrates information from diverse sources within a unified temporal framework that mirrors how individuals naturally perceive their daily lives. Beyond individual reflection, the dashboard also enables researchers to monitor the quality of the collected data, detect missing samples, and visualise engagement patterns. In this case, awareness operates on two levels: personal (behavioural insight) and observational (continuous evaluation of data reliability and participant interaction).

4.2.2 Temporal Stages: Past, Present, and Future

The dashboard organises data according to three temporal stages that structure human experience: *past*, *present*, and *future*. The *past* records completed actions and routines, providing a basis for evaluating progress and recognising patterns. The *present* reflects the immediate context in which decisions and actions occur. The *future* captures intentions and plans that guide anticipation and decision-making.

Each planned activity is represented as a plan P_i with a start time t_s and end time t_e ($t_s < t_e$). The state of a plan at time t is;

$$\text{State}(P_i, t) = \begin{cases} \textit{Future}, & \text{if } t < t_s \\ \textit{Present}, & \text{if } t_s \leq t < t_e \\ \textit{Past}, & \text{if } t \geq t_e \end{cases}$$

Transitions are monotonic: *Future* → *Present* → *Past*. No reverse transition is permitted, ensuring temporal consistency and preserving historical integrity.

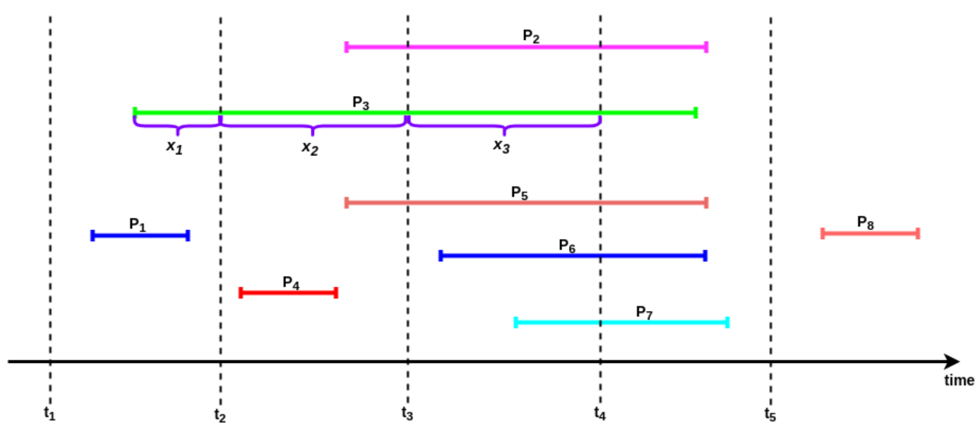


Figure 4.1: Plans and their states along the timeline, illustrating transitions between *past*, *present*, and *future*

Figure 4.1 illustrates multiple plans and their progression. At $t_1$all, plans are future. At t_2, P_1 has completed (past), while P_3 is ongoing (present). At t_3, P_1 and P_4 are past; P_2, P_3, P_5 are present. At t_4, P_6 and P_7 enter present; P_8 remains future. By t_5, all plans have moved to the past. During the present stage, the elapsed portion $[t_s, t_t]$ represents accumulated execution history, enabling partial evaluation prior to completion.

Future Stage

The *future* stage includes all plans where $t < t_s$. Activities exist as intentions, defined when researchers configure the experiment. Each plan is enriched with metadata including timestamps, expected responses, and associated sensor streams. For participants, visualising future plans enables anticipation of upcoming prompts and sensing intervals, supporting preparation and integration into daily routines. For researchers, the future stage provides pre-execution validation, enabling confirmation that experimental configurations are correctly deployed before data collection begins.

Present Stage

The *present* stage encompasses plans where $t_s \leq t < t_e$. Activities transition from intention to execution, generating observable behavioural and sensor data. The dashboard's Present view (Figure 5.8) shows participants their active tasks in real time, helping them manage study commitments and become more aware of their activity patterns as they occur. For researchers, this stage provides real-time visibility into participant engagement and protocol adherence, enabling timely intervention when issues arise, such as missed responses or inactive sensors, before data quality is affected.

Past Stage

The *past* stage includes all plans where $t \geq t_e$. Completed plans become immutable records, forming a stable basis for analysis and interpretation. For participants, the past stage enables reflection on behaviour over time: identifying recurring patterns, reviewing consistency of responses, and observing how routines change across days, weeks, or months. For researchers, the past stage provides the complete dataset for analysis, enabling assessments of data quality, including answer completeness, response-time distributions, and sensor reliability.

By modelling these three interconnected stages, the dashboard operationalises a continuous intention–action–reflection loop. Participants can anticipate actions, observe them in real time, and later reflect on what they have done. Researchers can verify study execution, monitor adherence, and assess data quality throughout the study lifecycle.

4.3 Dashboard Interface

4.3.1 Design Philosophy

The dashboard is designed as an interactive environment for managing, visualising, and analysing personal and contextual data. The design prioritises clarity, role-based functionality, and transparent communication of data quality. The calendar was chosen as the central organising structure because it aligns with how people naturally organise time and integrates past, present, and future activities within a single coherent view.

The system supports two distinct user roles, each with dedicated interfaces tailored to their specific needs:

Researcher Role

Researchers design experiments, configure study parameters, and monitor the ongoing data collection during the course of the experiment. Their key functionalities include:

- Creating, configuring, and scheduling experiments or plans. This includes specifying the data collection frequency, duration, and data types.
- Monitoring data quality metrics: The dashboard helps researchers to view and monitor the timing of question delivery, response rates, and completeness of sensor streams.
- Communicating with participants based on observed activity patterns to encourage engagement or address issues.

Participant Role

Participants contribute active data through their questionnaire responses, as well as passive data via continuous sensor streams. Their responsibilities encompass the following:

- They receive and respond to questions or tasks that are scheduled within the experimental framework.
- They are tasked with reviewing personal events and the associated contextual information, which is facilitated through the integrated calendar view.
- They also manage privacy settings and permissions for data collection, thereby maintaining control over personal information.

A single user can take on both roles at the same time, which allows for self-monitoring and participation in experiments.

4.3.2 Integrated Components

The dashboard comprises several visual components that work together to provide a great user experience:

- **Calendar Component:** The calendar, as already discussed, serves as the central temporal interface for reporting and reflecting on events, tracking events, and viewing planned future activities. Events are organised into past, present, and future stages 4.2.2 (with past and present combined in a single view due to interface complexity). Each event is annotated with contextual icons and quality indicators, enabling reflective analysis within a familiar temporal structure (see 3.1.1).
- **Graphs and Charts:** These are included to provide quantitative summaries of behavioural and contextual variables such as response rates, mood levels, activity durations, or sensor data completeness. Through interactive visualisations, users can observe trends, identify correlations, and detect deviations from typical patterns. Heatmaps offer aggregated views of activity intensity across time intervals, highlighting periods of high engagement or irregular submission patterns (see Figures 6.2 and 5.10).
- **Map Component:** A map view is featured to give an interactive spatial visualisation of participants' activities. Geolocation markers indicate the locations of events, while interactive navigation enables exploration of location-based patterns. Temporal overlays allow users to observe how spatial patterns change over time, while cluster representations highlight frequently visited locations or spatial hotspots.
- **Quality Indicators:** Integrated visual cues such as colour coding, icons, or annotation layers instantly reveal the completeness, consistency, and any potential problems with the data. These indicators shift data integrity from a back-end concern to a central aspect of how we understand things. This allows users to assess how reliable the patterns they observe in their data are.

4.3.3 Integration and Interaction

Dashboard components are interconnected to provide a coherent user experience, for instance:

- Selecting a temporal segment in the timeline highlights corresponding events on the map, linking temporal and spatial representations (Figure 5.9).
- Filtering by contextual dimensions (mood, activity, social presence) updates all visualisations in real time.
- Researchers can contact participants directly via the dashboard when data anomalies are detected, enabling timely intervention (Figure 4.2).

4.3.4 Visual Encoding of Contextual Data

The five dimensions of context (Section 3.1.1) are visually encoded to support quick comprehension and comparison:

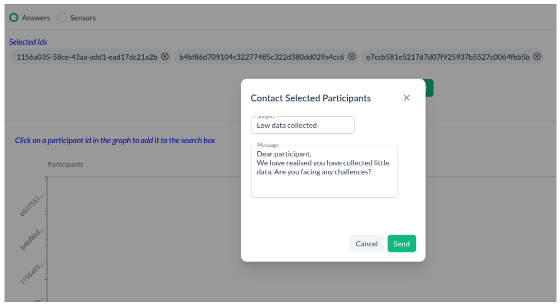


Figure 4.2: A researcher contacting participants with low response rates to improve data collection compliance

- **Location (WE):** Is represented by labelled markers on maps or location icons on the calendar.
- **Mood (WI):** Uses emoticons representing a 5-grade emotional scale.
- **Activity (WA):** Activities are represented by icons corresponding to specific tasks or actions.
- **Social Context (WO):** In the social context, we use symbols or avatar thumbnails to show the presence of other people.
- **Objects/Tools (WU):** These include small icons for relevant objects associated with the activity.

4.3.5 Spatial Context Visualization

The map component provides an interactive visualisation of activities within their spatial context. Features include geolocated markers, interactive navigation, temporal overlays, and cluster representations showing areas with higher concentrations of activities, as shown in Figure 4.4.

4.3.6 Design Requirements

The dashboard was designed to satisfy the following requirements, derived from the gaps identified in Chapter 2:

- **Temporal Awareness:** To represent past, present, and future events, enabling users to contextualise activities over time.

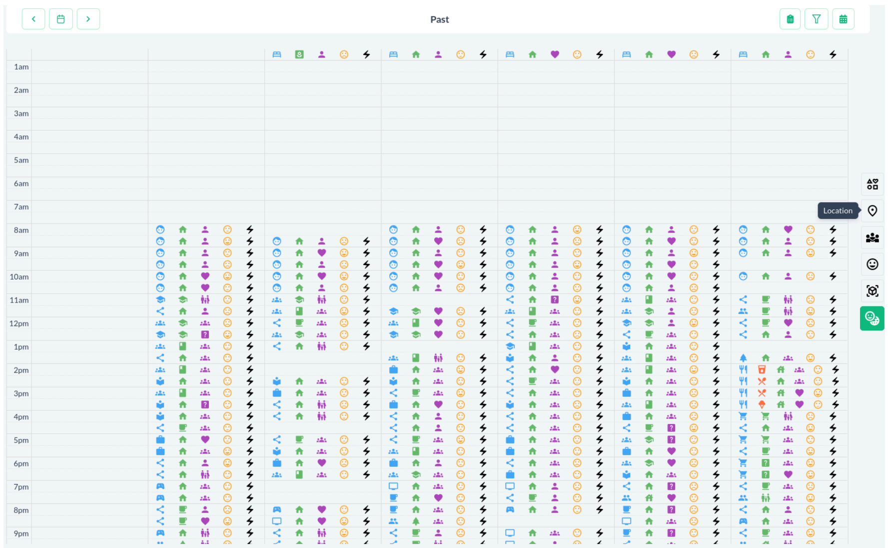


Figure 4.3: Contextual dimensions displayed in the dashboard. Each event includes icons representing location, activity, mood, social presence, and objects used.

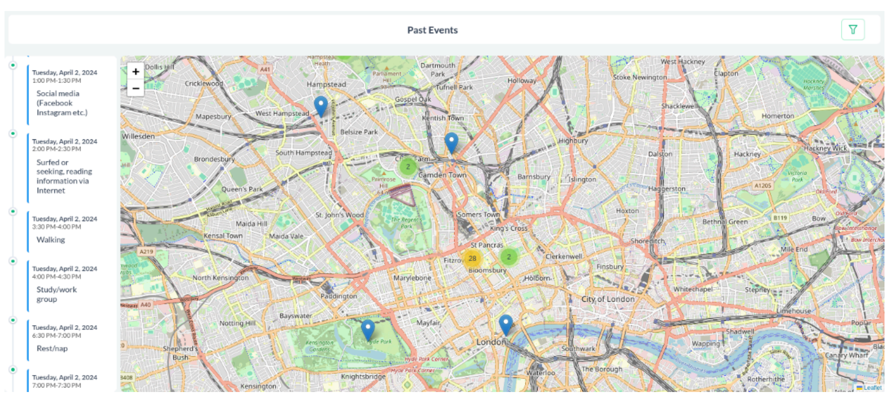


Figure 4.4: Spatial contexts displayed in the dashboard, showing geolocated markers, temporal overlays, and cluster representations.

- **Contextual Integration:** To incorporate all five contextual dimensions (location, activity, internal state, social context, and objects/tools) of data representation.
- **Data Quality Feedback:** To make completeness, consistency, and potential issues visible through explicit indicators.
- **Participant Engagement:** Support reflection, ongoing monitoring, and anticipation of future events.
- **Research Oversight:** Enable monitoring of experiment adherence and detection of anomalies by the researcher.
- **Interactive Visualisation:** Support exploration through visualisation views, including calendars, graphs, charts, and maps.
- **Role-Based Functionality:** Provide distinct capabilities for both participants and researchers.

4.4 Chapter Summary

This chapter has presented the design of the quality-aware calendar dashboard. We introduced the concept of life awareness and operationalised it through a temporal model comprising past, present, and future stages. We described the dashboard interface, including its dual-role architecture, integrated components (calendar, graphs, maps), and visual encoding of the five-dimensional context model. The next chapter details the system architecture that implements these design decisions.

Chapter 5

System Architecture

The proposed calendar dashboard is designed to enable continuous monitoring of participant behaviour and contextual data while generating meaningful insights for both researchers and participants. This goal is achieved by integrating and synchronising multimodal data streams and meta-contextual indicators into a single temporally aligned framework.

For researchers, this architecture provides tools to dynamically monitor data quality, participant engagement, and protocol adherence in real time, supporting adaptive study management and methodological refinement. For participants, the system provides an interface that promotes behavioural pattern recognition and informs daily choices. The architecture is designed to handle the full complexity of personal data by addressing several core requirements:

- scalable ingestion of high-volume, heterogeneous data streams;
- flexible modelling and representation of multidimensional behavioural and contextual data
- robust, real-time mechanisms for assessing data reliability;
- and visualisation approaches that provide analytical depth without imposing excessive cognitive load.

The key system requirements and constraints are as follows:

- **Reliable Data Collection:** The system must consistently capture, transmit, and store scheduled questions, tasks, and sensor measurements with minimal loss, duplication, or interruption.
- **Participant Privacy and Security:** All personal data, including sensor readings and self-reported information, must be securely managed and stored in compliance with relevant privacy regulations (e.g., GDPR).

- **Real-Time Data Monitoring and Feedback:** Researchers should be able to monitor ongoing experiments, assess participant compliance, and detect anomalies in the data in real time, enabling timely interventions.
- **Scalability:** The architecture must accommodate numerous simultaneous participants and extended studies without decline in performance or bottlenecks.
- **Contextual Awareness:** The platform should integrate multidimensional contextual signals (e.g., location, physical activity, mood, social presence, and environmental context) to support both human behaviour interpretation and modelling.
- **Visualisation Support:** Dashboard interfaces must provide rich temporal, spatial, and analytical visualisations, including calendars, maps, graphs, and heatmaps, to enhance reflection, awareness, and overall data understanding.
- **Cross-Platform Accessibility:** Participants should be able to access and use the system from any device they choose, including smartphones, tablets, and standard web browsers. This ensures widespread availability of the system and a consistent usage experience for everyone.
- **Data Quality Assurance:** It is necessary that the system undergoes continuous evaluation and reporting concerning data completeness, response times, internal consistency, and other pertinent quality indicators. This is essential to facilitate researchers' assessment of the reliability of the data that has been collected.

5.1 Architecture Overview

Figure 5.1 presents the component-level architecture of the dashboard system, showing the interactions between users, the dashboard, and supporting systems.

The dashboard's design rests on a modular, layered architecture, a choice for ensuring scalability, ease of use, and maintenance. This structure also streamlines data collection and contextual comprehension. At its core, the system is constituted of three primary components: the dashboard front-end, the back-end, and data storage. These elements interact via clearly defined APIs, which foster modularity and a clear delineation of user roles.

5.2 System Components

5.2.1 Supporting Systems

The dashboard operates within a large ecosystem of interconnected systems that collectively support data collection, management, and user engagement. These supporting systems provide the necessary infrastructure to ensure that data flows reliably to foster collaboration and scalability. These are discussed in the preceding sections.

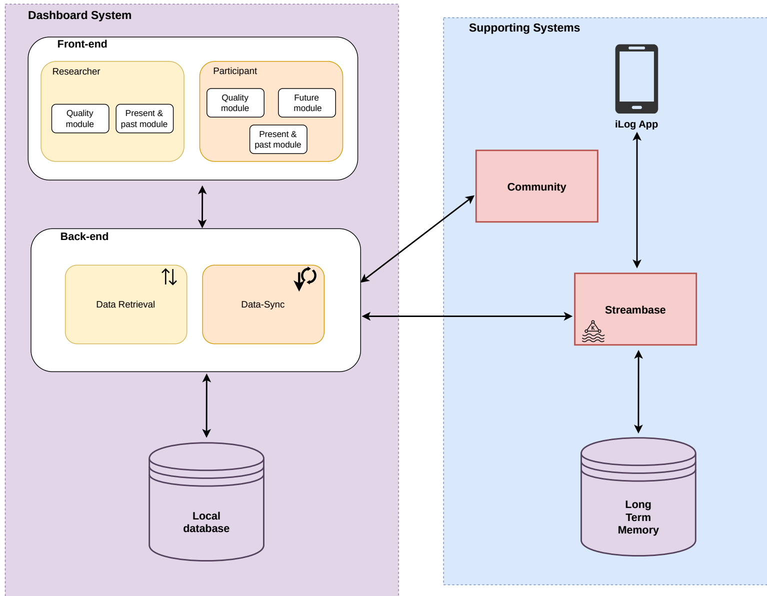


Figure 5.1: Logical architecture of the dashboard system

iLog App

The iLog application [58] is the primary data collection tool installed on participants' smartphones. It is responsible for delivering scheduled questions, capturing user responses, and collecting sensor data from the participant's smartphone. The application communicates with the back-end in real time, ensuring that data is synchronised and reliably stored. In addition, iLog provides participants with task notifications, thereby linking the experimental design configured by researchers with participants' lived experiences.

Community Platform

The DataScientia Community operates as a collaborative space associated with the dashboard ecosystem. It serves several functions:

- **Onboarding and Orientation:** New participants are first directed to the welcome page, where they can find guidelines and tutorials. These resources clarify the study's goals, explain how to use the dashboard, and guide participants on the best ways to participate.
- **Support and Feedback:** Participants are welcome to flag any issues, pose questions, and request clarification on this platform. Researchers then sift through these exchanges, looking for patterns of recurring problems, misunderstandings, or technical glitches. This ongoing dialogue is key to preserving data quality and keeping participants engaged.

- **Motivation and Engagement:** Community elements, like discussion boards and shared accomplishments, help keep participants involved throughout the entire course of the studies.
- **Knowledge Sharing:** Researchers can share insights, participation guidelines, data summaries, and best practices that participants can use to reflect on their data and behaviour patterns.

By integrating this Community module, the system enhances both the human side of the experiment, supporting engagement, clarity, and motivation, and the technical side, helping to communicate and fix any technical problems.

Streambase

The Streambase component acts as an external middleware layer that facilitates the transmission of high-frequency data streams collected. It primarily functions as a buffering and routing means between the data collection infrastructure and the dashboard back-end. Incoming data produced by the iLog application is forwarded here and delivered to the data synchronisation service (5.2.2).

In combination, these supporting systems, iLog, Community, and Streambase, form the foundation that enables the dashboard to function as an integrated platform for both reflection and research. They bridge the gap between the raw data collection and participant engagement, ensuring reliable integration between the data collection infrastructure and the calendar dashboard.

5.2.2 Core Dashboard Components

Front-end Architecture

The dashboard's front-end serves as the main way users interact with the system. It is web-based, which allows access from different devices, and it changes based on the user's login role. The system supports two main roles: researcher and participant. When a user logs in, the system authenticates their credentials and loads an interface tailored to their role, ensuring both usability and role-based access control.

We discuss the role-based interfaces as follows:

Researcher Interface The researcher interface supports continuous oversight of participant activity, data quality, and experiment progress. It combines high-level summaries with participant-level diagnostics and configurable quality controls.

At the participant level, the interface provides a structured, tabular overview of collected data, summarising key indicators such as enrolment time, number and percentage of completed

Participant Id	Date joined	No. of Answers	%age of Answers	Avg Completion Time (sec)	Avg Reaction Time (sec)	UnAnswered Questions	Sensor Data (1 unit/hr)
c21e29f8f5cbe420d2ef48b2d8e9f37eebb74b37	2024-03-11 11:39:29	830	97.41784037558685	13.69	10180.88	0	1182
f398a1835d783a3379a9d117e7273b7086ab3905	2024-03-08 16:00:54	813	95.4225352112676	9.04	22101.24	8	1092
bafa9740f87e312a2d1c67190a303e83f98e4f07	2024-03-14 13:54:04	789	92.6056338028169	14.16	27686.57	7	582
d46273cf311adfcbc1c1079987d040b749807f71	2024-03-10 09:47:33	783	91.90140845070422	13.4	6684.77	0	1073
d24efefe59e6b8cc0e22c223b5c2dcfe20d3bcde	2024-03-15 15:19:05	776	91.07981220657277	14.59	26412.76	0	1382
b8e22dc0d8c731924b339ae0040332f1d29dba89	2024-03-18 14:22:42	762	89.43661971830986	12.81	2767.86	0	410
bd6cddf9a54a04998f212d2f16c5f7ea1c40173e	2024-03-16 20:08:56	759	89.08450704225352	12.05	21977.5	0	1043
fb4d508e2aa067b5cd33294f4fe41003de545541	2024-03-07 11:22:26	729	85.56338028169014	13.9	15029.22	0	266
eb9d9eecbccdfbb5871122ac9ffc9142a5965b4d	2024-03-15 09:30:05	722	84.74178403755869	11.24	8305.55	0	414
b91384850a70c809b57a797cdd6db92be42bcf1d	2024-03-10 11:53:36	719	84.38967136150235	13.86	22123.86	27	958

Figure 5.2: Participant-level data monitoring interface showing aggregated answer metrics, response behaviour, and sensor data availability.

answers, average completion and reaction times, unanswered questions, and associated sensor data volume. As shown in Figure 5.2, this view supports sorting and filtering across multiple quality-related dimensions (e.g., participants that meet quality limits, question loss, duplicated data, or expired questions), allowing researchers to rapidly identify anomalous behaviour, low engagement, or missing data streams during active data collection.

These indicators are computed based on the formulations introduced in Sections 3.2.2 and 3.2.4. In particular, question loss is derived using Equation 3.5, while the number of questions received is computed using Equation 3.4.

In parallel, the dashboard offers an experiment-level summary view that aggregates statistics across participants, questions, answers, and sensors (Figure 5.3). This includes metrics such as total and consented participants, the proportion meeting predefined quality thresholds, the number of questions generated and their delivery status, average answer-response characteristics, and sensor availability. Temporal indicators, including the experiment progress, remaining duration of the experiment, and overall answer completion percentage, provide general information about the experiment's progress relative to the planned study timeline.

Significantly, the researcher interface integrates configurable data-quality parameters directly into the front-end, enabling the specification of acceptable answer ranges, response-time thresholds, and other quality constraints. These parameters ensure that quality assessment is applied consistently throughout the study. When participants fall below defined thresholds or exhibit missing or delayed input, the system enables researchers to initiate targeted communication, supporting corrective intervention during the data collection rather than post-exclusion after the experiment (Figure 4.2). This design strengthens data reliability, reduces attrition-related bias, and improves the overall robustness of the resulting dataset.

Within the researcher interface, data quality monitoring is structured around two complementary components: (i) a modality-specific heatmap visualisation for answer data (computed using

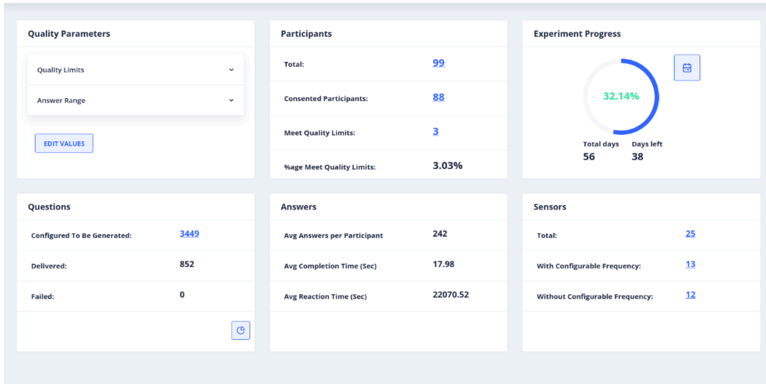


Figure 5.3: Experiment summary dashboard illustrating participant statistics, data quality indicators, experiment progress, and sensor configuration overview.

Equation 3.10) and sensor data (computed using Equation 3.15), and (ii) a configurable quality filtering interface that supports participant selection and comparative analysis.

Heatmap views are used to inspect either answer-level data or sensor-level data, with the researcher explicitly selecting the modality of interest. In this view, participants are arranged along one axis and questions or sensors along the other, while colour intensity encodes data availability or completeness. This representation enables rapid visual detection of missing data streams or uneven data coverage across participants (Figure 5.4).

Complementing this visual inspection, the Quality Monitoring module provides a dedicated filtering interface that allows researchers to define explicit quality constraints, including answer completion percentage, temporal ranges, response duration, and reaction time. Applying these filters returns a structured table of participants who meet or violate the specified criteria, from which individuals or subsets can be selected for further inspection (Figure 5.5A).

Once participants are selected through the filtering interface, the system supports direct comparison of their response behaviour using automatically generated visual summaries. These include comparative graphs showing the number of answers submitted, answer completion duration, and response time distributions across selected participants (Figure 5.5B). This comparative view enables researchers to distinguish behavioural differences from isolated anomalies and to contextualise quality issues within individual participation patterns.

These components work together to create a quality assessment workflow that is clear and includes exploratory visual diagnosis, rule-based filtering, and comparison at the participant level. By enabling researchers to move fluidly between these modes during active data collection, the quality monitoring approach reduces reliance on post-experiment evaluation and enhances the clarity and replicability of data quality determinations in multimodal studies.

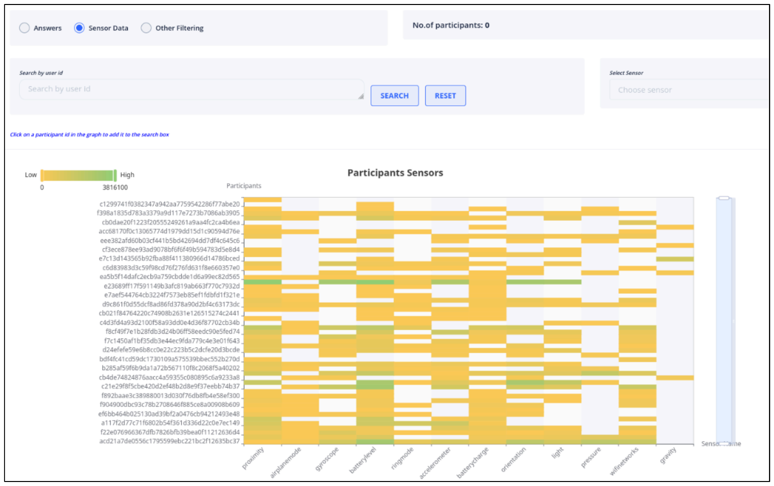


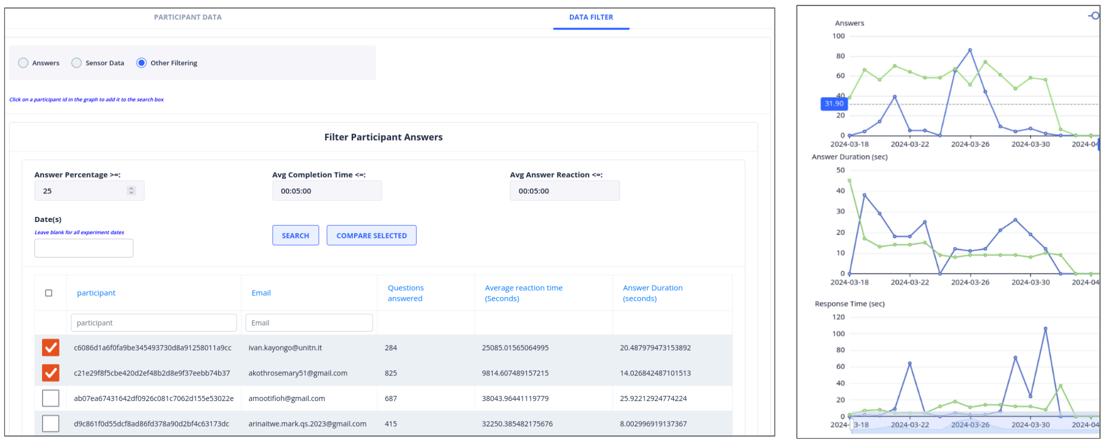
Figure 5.4: Modality-specific quality heatmap showing participant–question or participant–sensor data completeness.

Participant Interface The participant interface prioritises reflection and engagement. Unlike the researcher's view, which emphasises monitoring, this interface organises personal data along a temporal structure, past, present, and future, to foreground lived experience while integrating contextual visualisations such as activity logs, mood indicators, location traces, and experiment-related tasks.

Across all temporal modules, interaction is supported through lightweight visual encodings. High-level summaries offer an overall overview of the experiment and the data, whereas interactive components facilitate participants' exploration of more granular contextual data as required. This approach enables participants to understand their data while simultaneously gaining access to more comprehensive information when necessary.

The *Experiment Progress* view (Figure 5.6) exemplifies this design approach by synthesising temporal organisation and reflective feedback within a single view. The interface combines indicators of the present state (e.g., overall progress and remaining study duration), summaries of past participation (e.g., questions received, expired prompts, and the user's response behaviour), and forward-looking cues that guide future engagement through standard comparison. Metrics such as average response time and missed answers are presented as reflective signals rather than evaluative performance measures, enabling participants to contextualise their participation patterns without invoking surveillance or judgement.

Through this framing, raw participation metrics, such as question counts, response durations, and engagement totals, are transformed into a simple narrative of the participant's progress within the study. These metrics are also derived from the quantitative formulations defined in Section 3.2.2, including Equations 3.4, 3.5, 3.9, and 3.11.



(a) Rule-based filtering and participant selection table

(b) Comparative graphs of selected participants

Figure 5.5: Quality filtering and comparison interface supporting participant selection and cross-participant data comparison.

Future Module The future module supports planning, anticipation, and preparation for upcoming events and experiment-related tasks. Unlike the past and present modules, which use data that has already been collected or is currently being collected, this module visualises scheduled tasks or surveys before they happen. As shown in Figure 5.7, future events are displayed within a calendar-based layout, with distinct colours used to differentiate question types and task categories.

This representation allows participants to anticipate upcoming tasks and understand future participation requirements. Importantly, it establishes continuity across temporal states: events visible in the future transit into the present and subsequently into the past, reinforcing a sense of time rather than fragmented task prompts. From a research perspective, this module also enables monitoring of expected compliance by making upcoming participation requirements explicit to participants.

Present Module The present module foregrounds real-time awareness by highlighting the *current segment* of the participant's day. This view emphasises ongoing activities, locations, companions, and moods as they unfold, supported by visual cues that distinguish active contexts from completed or planned ones. Sensor-derived inputs, such as location updates or device interaction signals, are integrated alongside self-reported information, enabling participants to cross-check their subjective perceptions against system-captured data.

By making the current context visible, the present module promotes awareness while also supporting compliance with the experiment through timely reminders or notifications when tasks are due. Together, these features help bridge immediate lived experience with structured data collection (Figure 5.8).

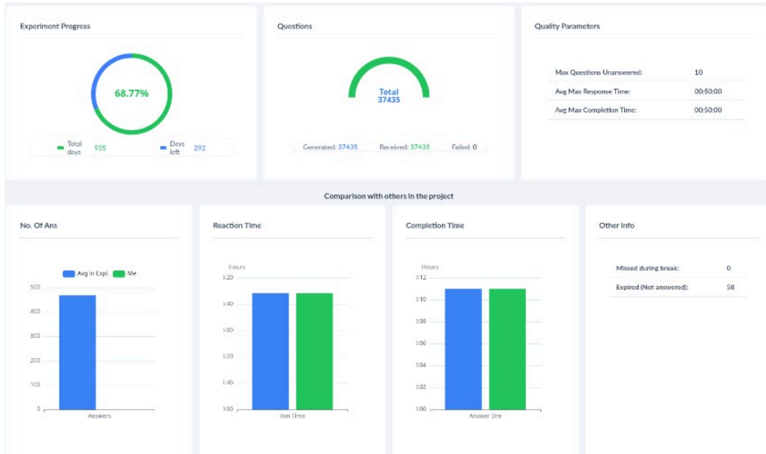


Figure 5.6: Participant-facing *Experiment Progress* dashboard integrating progress tracking, historical participation summaries, and comparative feedback within a temporal and reflective framework.

Past Module The past module represents previously recorded experiences in a calendar. Each calendar cell integrates the five context dimensions (section 3.1.1), activity, location, companions, mood, and objects, encoded as icons. This simple visual language allows participants to quickly see their daily activities. At the same time, the interface's interactive features enable a deeper understanding of the collected data.

When a user selects or hovers over an entry, expanded contextual information is displayed, encompassing descriptive annotations, sensor summaries, and experiment responses. This two-tiered presentation, which integrates overview icons with readily available detailed information, facilitates reflective sense-making while mitigating cognitive strain. By situating experiences along a temporal axis, the past module helps participants identify routines, disruptions, and longer-term behavioural patterns (Figure 5.8).

Beyond participant reflection, the past module serves as a core analytical resource for researchers. By revisiting completed temporal segments, researchers can assess data completeness, identify delayed or missing responses, and compare participant behaviour over time using derived metrics such as answer counts, completion duration, and response times. The temporal alignment of contextual events enables researchers to interpret quantitative quality indicators in relation to daily routines and situational factors, rather than in isolation (Figures 5.5 and 5.4).

Furthermore, we incorporate two additional components: a Map View and a Data Viewer, as discussed below.

Map View Complementing the temporal calendar views, included is a map view that provides a spatial perspective on participant experiences. The Map View provides a spatial representation of participants' activities by plotting recorded events and contextual data onto a geographic map.

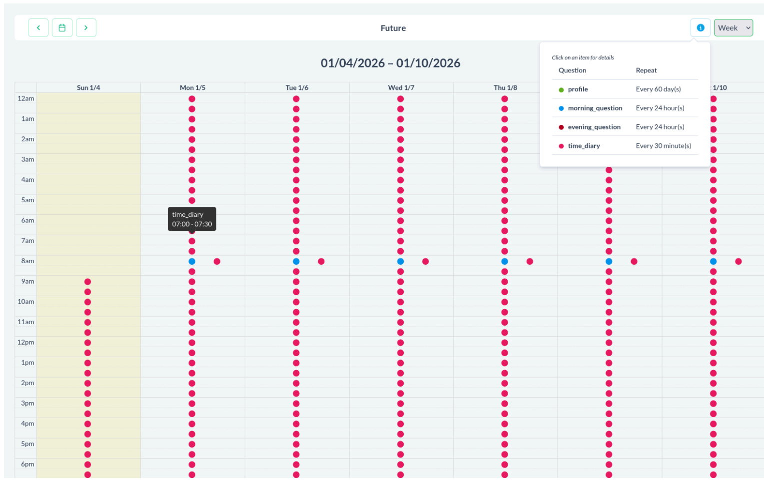


Figure 5.7: Future module showing scheduled experiment tasks and prompts visualised in a calendar-based layout with colour-coded question types.

This enables participants to examine where activities occurred and how spatial context relates to their behavioural patterns over time.

The interface, beyond the map, features a temporal activity panel. This panel, situated adjacent to the map, displays a chronological inventory of documented activities, each accompanied by its respective time frame. Once an activity is chosen from the list, the related location is highlighted and centred on the map. This feature allows for direct connections between events in time and their locations in space. As a result, this relationship encourages an exploration of the spatial and temporal patterns of specific activities.

At the top of the Map View, a set of filtering controls allows users to navigate the displayed data based on attributes such as activity type, time range, or other recorded metadata. These filters dynamically update both the activity list and the map, supporting focused inspection of subsets of data while maintaining a connection between temporal and spatial representations (Figure 5.9).

By enabling participants to explore their data spatially as well as temporally, the map view supports recognition of mobility patterns and place–context relationships. Together, the calendar and map views provide a holistic representation of lived experience across time and space.

For researchers, the Map View provides an aggregate spatial overview of participant distribution rather than a detailed activity-level trajectory. Participant data are visualised as spatial markers summarising the geographic distribution of active participants within an experiment.

This view is intended to support high-level monitoring tasks such as assessing geographic coverage, identifying spatial clustering, and detecting potential biases or gaps in participant recruitment. Unlike the participant-facing map view, the researcher's view does not expose

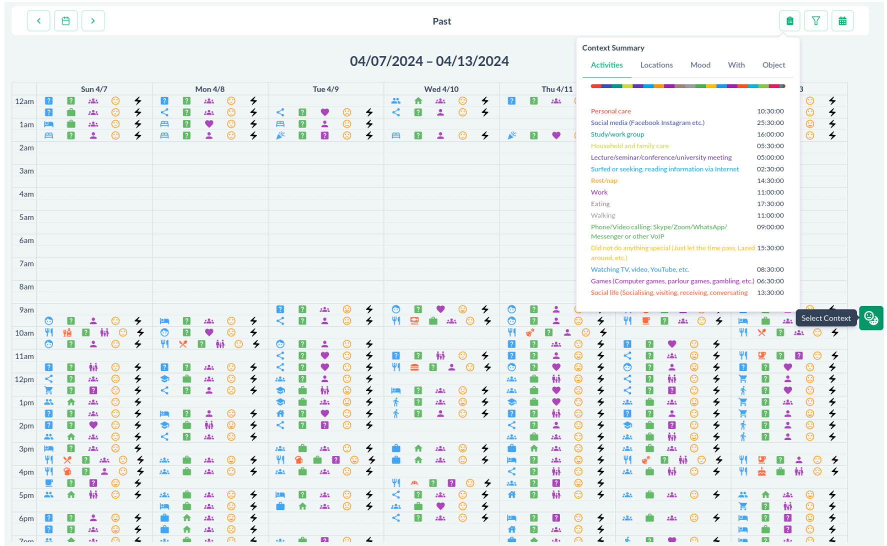


Figure 5.8: Participant interface showing present and past modules with calendar-based contextual encoding of activities, locations, companions, moods, and objects.

individual activity timelines or per-event navigation, thereby preserving abstraction and reducing the risk of over-interpretation at the individual level.

Filtering controls allow researchers to constrain the displayed distribution based on experiment-level parameters (e.g., date range or participation status), enabling focused inspection of spatial patterns while maintaining a population-level perspective.

Data Viewer The data viewer enables participants to explore behavioural patterns in relation to contextual information. Rather than presenting raw data in isolation, this component aggregates sensor readings, survey responses, and contextual annotations into interpretable visual summaries that evolve across time.

Multiple visual representations are supported, including temporal line charts (e.g., step counts and movement patterns), stacked distributions (e.g., food eaten, companions, and locations over days), and proportional summaries (e.g., dominant locations or social contexts). These views enable participants to identify trends, routines, and variability in their behaviour while maintaining a clear connection to the underlying temporal structure of their daily lives.

Interaction within the data viewer is designed around progressive disclosure and comparison across time windows. Participants can navigate between detailed views, adjust scopes, and visually inspect how different contextual dimensions co-occur or change. In doing so, the data viewer reinforces transparency and personal data ownership, helping participants interpret their behavioural data over time without requiring analytical expertise.

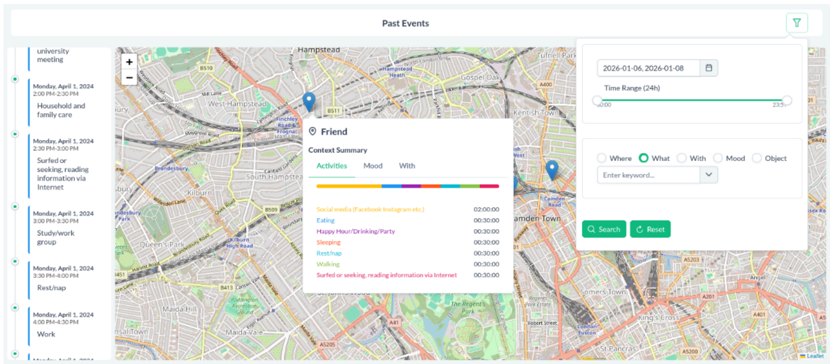


Figure 5.9: Map view providing a spatial representation of participant experiences, enriched with contextual information and interactive detail.

Figure 5.10 illustrates how aggregated summaries and temporal charts are combined to support progressive, participant-centred exploration of personal data.

Back-end Architecture

The back-end of the dashboard acts as the central layer that bridges the stored data collection from the iLog platform with the visualisation and interaction features of the front-end. Its main job is to ensure that incoming data is reliably captured, processed, and stored. It also provides efficient and flexible access to this data through well-defined query methods. It is composed of:

Data Synchronisation Service (Data-Sync) The data-sync component serves as the ingestion pipeline for all information generated during a running experiment. As participants interact with the system, the iLog application continuously produces different forms of data: answers to questions, logs of scheduled or completed tasks, contextual entries such as mood or activity, and sensor traces from mobile devices. This information is transmitted to the data-sync in real time, ensuring that no significant delays occur between data collection and storage.

The data-sync component not only helps in the storage of incoming data but also performs several intermediate steps to ensure that the data remains reliable and usable for subsequent analysis. First, all records are normalised into a consistent schema that allows heterogeneous inputs (e.g., free-text answers, categorical values, or continuous sensor streams) to coexist in the same storage environment. Second, the data is cleaned and filtered to remove duplicates, resolve inconsistencies, and address missing or corrupted data. Finally, metadata such as timestamps, participant identifiers, and experiment identifiers are attached to each entry, creating strong links between raw values and their experimental context.

In addition, the data-sync prepares pre-aggregated summaries of common statistics, such as the

(a) Temporal visualisations enabling inspection of trends and distributions across days and contextual dimensions.

(b) Overview summaries showing dominant moods, locations, companions, and activity-related sensor aggregates.

Figure 5.10: Participant-facing data viewer supporting reflective exploration of personal sensor data and self-reported context through complementary overview and temporal visualisations.

number of delivered questions, the proportion of answered items, and the number of missing sensor readings. These precomputed values enable faster retrieval, reducing the processing load on the retrieval service. Because the service is event-driven and operates continuously, researchers can expect the dashboard to display near-real-time reflections of participant activity and engagement.

Data Retrieval Service The data retrieval service acts as the primary gateway between the back-end and the front-end. Once data has been stored and indexed by the data-sync, this service makes it accessible to users through role-based queries. The retrieval layer is designed to give both participants and researchers access to the specific information they need, without introducing unnecessary complexity or affecting performance.

For participants, queries are lightweight and focused on personal data. The retrieval service provides access to individual timelines, daily mood distributions, activity patterns, and summaries of completed or pending experiment tasks. Since participants may consult the dashboard multiple times a day, queries are optimised for low latency, delivering results in microseconds even on mobile devices (refer to section 5.4.4). This responsiveness is crucial for maintaining participant engagement and encouraging reflection on personal behaviours.

In contrast, researchers require broader and more computationally intensive queries. These include aggregating data from all study participants, quality checks to find missing or duplicate records, comparing data over time across different parts of an experiment, and visual summaries like heatmaps and statistical graphs. To support these needs, the retrieval service uses indexing methods, caching strategies for frequently requested datasets, and the ability to filter data using complex parameters, such as date ranges, participant groups, or experimental conditions.

By separating data ingestion from access, the back-end architecture achieves both scalability and modularity. The continuous processing of incoming data ensures that stored records are always current. At the same time, the retrieval service efficiently handles both simple and complex queries. These two services together form the core of the dashboard, enabling researchers and participants to interact with information that is accurate, current, and contextually rich.

Storage Layer

This is the storage component where the large volumes of data collected during experiments are stored. Sensor measurements, contextual annotations, and question–answer responses are stored in relational tables indexed by time, which supports both data consistency and efficient querying.

The database is implemented using *TimescaleDB*, which provides indexing and partitioning mechanisms suited for large time-series datasets. To handle the demands of long-term studies and a growing number of participants, the database primarily organises data based on time. It also uses participant or experiment identifiers when needed. This design allows the system to scale as datasets grow, while still providing efficient access to both recent and historical data.

Several design considerations are implemented to guarantee quality and performance:

- **Data model:** Each record is time-stamped and associated with a participant ID and an experiment ID, which ensures complete traceability. Structured attributes, including contextual elements like location, mood, and activity, along with related objects, are stored. This structure allows for precise joins and aggregations.

- **Ingestion and validation:** Incoming data from the synchronisation service undergoes validation before its insertion. Duplicate entries are eliminated, missing values are flagged, and integrity checks are applied to ensure the reliability of the dataset.

- **Performance optimisation:** Indexing on time and participant identifiers supports efficient temporal queries and contextual segmentation. We use TimescaleDB's native compression to reduce storage overhead for long-term sensor data while maintaining queryability.

- **Retention and summarisation:** Policies are configured to retain raw, high-frequency sensor data for a limited window, after which summarised representations (e.g., averages, counts, and min/max) are stored to support scalable long-term analysis.

In summary, the storage layer is robust and scalable for the calendar dashboard system. By combining relational consistency with time-series performance, it ensures that both real-time synchronisation and historical analysis are supported within the same structured framework.

5.3 Implementation Details

This section describes the specific technologies, frameworks, and design patterns used to build the architectural components mentioned earlier. The implementation follows current software engineering practices, which helps ensure the system is scalable, easy to maintain, and reliable.

5.3.1 Front-end Technology Stack

The front-end is built using Angular 19, which uses a component-based architecture. This design choice supports modularity, makes maintenance easier, and allows for dynamic updates to the user interface. For a rich, interactive experience, we integrated several specialised libraries:

- **PrimeNG, PrimeIcons, PrimeFlex, TailwindCSS-PrimeUI:** These libraries form the primary user interface (UI) framework for the system. `PrimeNG` offers a comprehensive suite of prebuilt UI components, including advanced tables, form controls, modals, menus, and multi-select dropdowns, which accelerate development while ensuring a consistent design language. `PrimeIcons` provides a scalable vector icon set tightly integrated with `PrimeNG` components, which enabled us to enhance the visual clarity and user guidance within the interface. `PrimeFlex` and `TailwindCSS-PrimeUI` provide utility-based classes for responsive layouts, spacing, alignment, and typography, facilitating the construction of complex page layouts that adapt seamlessly across different screen sizes and devices. Together, these libraries enabled the rapid development of our beautiful UI for the calendar dashboard.

- **FontAwesome and Google Material Icons** were both brought into the mix. Beyond the icons already included with `PrimeIcons`, `FontAwesome` offered a broad and varied collection of scalable vector icons. These were used for actions, status indicators, and buttons, which are not part of the core `PrimeNG` suite. `Google Material Icons`, adhering to Google's Material Design guidelines, contributed even more consistent and modern iconography. This was particularly useful for standard UI patterns like navigation bars, floating action buttons, and feedback elements. The combination of these icon sets ensured consistency across the dashboard's components, ultimately improving usability and the overall user experience.

- **ECharts and ngx-echarts:** We also used the `ECharts` library for advanced, interactive visualisations, including line charts, bar charts, stacked bar charts, heatmaps, gauges, and word clouds. `ngx-echarts` is an Angular wrapper for `ECharts`, enabling reactive updates in response to dynamic sensor or questionnaire data. This setup was used for interactions, such as zooming, dynamic tooltips, animated transitions, and interactive legends, enabling intuitive exploration of large datasets and real-time insights from patterns and trends.

- **FullCalendar:** To visualise temporal events such as sensor triggers, survey responses, or participant activity logs, we integrated `FullCalendar`, which provides interactive calendar views. Its plugin architecture, including `daygrid`, `timegrid`, `list`, `interaction`, and multimonth, offers a wide range of flexible representations, from detailed daily schedules to aggregated monthly overviews.

- **Leaflet with ngx-leaflet and marker clustering:** Geospatial visualisations of participant locations and activity contexts are implemented using `Leaflet`, integrated into Angular components via `ngx-leaflet`. This approach allows reactive updates of maps in response to user interactions or incoming data streams. The system further incorporates `Leaflet's Marker Cluster` to aggregate dense point locations for more precise visualisation, and `Leaflet's Heatmap Library` to represent spatial activity intensity via a heatmap. These features provide an intuitive view of location-based patterns, improving the interpretability of movement and contextual data.

- **Tippy.js:** To enhance usability and reduce visual clutter, the system uses `Tippy.js` for interactive tooltips and popovers. These provide contextual information when hovering or clicking on UI elements, allowing users to access additional details, explanations, or controls without overwhelming the interface. This contributes to a cleaner, more guided, and user-friendly experience, particularly when working with dense datasets or complex visualisations.

The front-end, thanks to these libraries, offers a dashboard that's both visually striking and incredibly interactive. It presents complicated sensor and survey data in a way that's easy to understand and navigate. The technologies were chosen based on their maturity, community support, compatibility with Angular, and ability to handle real-time dynamic datasets. This ensured the system's maintainability and scalability.

5.3.2 Back-end Technology Stack

The back-end is constructed using Java and Maven, adhering to a modular, service-oriented architecture that distinctly separates data ingestion, processing, and storage functionalities. This architectural design promotes maintainability, scalability, and the capacity to independently develop or upgrade individual components without impacting the entire system. The back-end is responsible for managing data from diverse sources, such as mobile applications, sensors, and system logs. It validates the data to ensure both accuracy and consistency, and it efficiently persists this data for subsequent retrieval, analysis, and visualization.

The key technologies and frameworks employed within the back-end are as follows:

- **Apache Kafka:** Kafka functions as a real-time messaging and event streaming platform, specifically designed for handling high-frequency sensor data, questionnaire submissions, and system events. It offers fault tolerance, keeps messages in order, and allows for replaying them, all critical when dealing with data streams that are either time-sensitive or carrying a lot of information.

- **Flyway:** Is the tool of choice for managing database versioning and schema migrations within TimescaleDB. It makes sure every environment, from development to production, has the same database structure, which is key for continuous integration and deployment. Flyway also makes it easy to upgrade database schemas without losing any historical data or causing any downtime.

- **Jackson:** Handles the serialisation and deserialisation of JSON between Java objects and messages sent over Kafka or REST APIs. By acting as a connector between the object-oriented back-end and the data transport layer, Jackson streamlines data exchange, cuts down on repetitive code, and keeps message formats consistent.
- **Persistence Libraries (JPA/Hibernate):** These provide object-relational mapping (ORM) capabilities to map Java objects to database tables efficiently. JPA/Hibernate allows complex queries, batch inserts, and transactional operations to be expressed in Java, while optimising interactions with TimescaleDB for temporal, aggregate, and analytical queries.

5.3.3 Database Implementation

All data collected is stored in a TimescaleDB instance, a PostgreSQL-based time-series database optimised for high-frequency temporal data. The storage layer is built to deliver both speed and dependability, catering to analytics and long-term data storage needs. Key features include:

- **Structured Storage:** Sensor data, survey answers, and system logs are organised in structured tables, complete with clearly defined schemas. This setup allows for efficient querying, aggregation, and relational joins, all of which are essential for subsequent analytics.
- **Time-Series Optimisation:** TimescaleDB uses hyper-tables to automatically partition data based on time intervals. This design allows for the efficient execution of queries on extensive temporal datasets, and it scales effortlessly as data volumes increase.
- **Data Integrity and Consistency:** Database constraints, foreign key relationships, and transactional guarantees work together to ensure that all stored data is accurate, consistent, and reliable.
- **Scalability:** The storage layer is designed to handle both vertical scaling, which adds more resources to each node, and horizontal scaling, which adds more nodes. This allows it to manage increasing amounts of sensor and survey data without affecting performance or how quickly queries are processed.
- **Integration with back-end Services:** back-end services interact with TimescaleDB via JPA/Hibernate for efficient ORM-based data access. Flyway is used to ensure controlled schema migrations, reducing the risk of inconsistencies or downtime across development, testing, and production environments.

Collectively, these back-end and storage components ensure that the system can handle high-frequency, heterogeneous data streams, maintain data integrity and consistency, and provide reliable access for analytics, visualisation, and reporting.

5.3.4 Data Processing Pipeline

To improve the reliability and analytical value of the data, a structured data processing pipeline is used. This pipeline includes:

1. **Preprocessing:** This step involves cleaning, aligning, and formatting the raw data from sensors and questionnaires.
2. **Normalisation and Feature Engineering:** Involve standardising numerical ranges, encoding categorical variables, and creating new features. These steps are crucial for improving later analysis and modelling.

Data Flow and Synchronisation Data from supporting systems, such as the iLog app and the community platform, is sent using REST APIs and Kafka streams. This data is then processed and stored in TimescaleDB. This ensures real-time dashboard updates while keeping the system decoupled and scalable.

5.3.5 Security and Authentication

Security in the system is enforced at multiple levels to protect the sensitive participant information:

- **Authentication:** Users are verified via JSON Web Tokens (JWT), ensuring that only authorised users can access the system. Tokens include expiration and refresh mechanisms to maintain security without affecting usability.
- **Authorisation:** Role-based access control (RBAC) is designed to limit access to sensitive data and administrative tasks. Permissions are carefully managed at both the API and database levels, ensuring that unauthorised actions are effectively blocked.
- **Data Protection:** All data is encrypted in transit using TLS/HTTPS and at rest in TimescaleDB. Sensitive fields, such as participant identifiers, are pseudonymised before storage. Logging and audit trails serve to monitor who accesses sensitive data, ensuring accountability. Meanwhile, access to the database itself is secured with a password.

5.3.6 Error Handling and Logging

To guarantee system dependability, robust error handling and logging protocols are employed:

- Exceptions are captured across various layers, including the API, Kafka consumers, and database operations, and subsequently logged to aid in both debugging and auditing processes.

- Kafka consumer offsets are only tracked and committed once processing is successful. This approach guarantees either exactly-once or at-least-once delivery, depending on the specific requirements.
- Database transactions are designed to be atomic. Every operation within a transaction must either complete successfully in its entirety or fail completely. If something goes wrong, rollback procedures are triggered to safeguard against data corruption.
- System metrics, alerts, and structured logs are constantly monitored. This allows for the detection of any anomalies, performance slowdowns, or outright failures. The goal is to enable quick intervention and recovery when issues arise.

5.4 Privacy, Data Ethics, and Transparency Considerations

Given the sensitive and personal nature of the data collected, the system uses strict privacy, ethical, and transparency protocols to ensure responsible data management.

5.4.1 Consent Management

Participants provide informed consent before data collection, including a clear explanation of:

- The types of data collected
- How data will be processed, stored, and analysed
- Rights to withdraw consent and have data deleted

The system tracks consent status and enforces data processing restrictions accordingly.

5.4.2 Transparency and Reproducibility

We have documented most of the pipeline, i.e., data preprocessing, augmentation, and analysis procedures. Synthetic or augmented data is clearly marked to prevent misinterpretation. The system's design supports reproducible pipelines, allowing for independent verification of results while maintaining privacy.

5.4.3 Ethical Considerations in Synthetic Data

The creation of synthetic and augmented data serves to broaden analytical scope while safeguarding privacy. This generation process maintains the statistical characteristics of the source datasets, thereby precluding the disclosure of individual participant details. Ethical protocols are strictly adhered to, with the aim of preventing the introduction of bias or misrepresentation in subsequent analyses.

5.4.4 Performance Evaluation

To validate the system's scalability and query performance, we measured database query response times across datasets of varying sizes under realistic multi-user conditions. The datasets were generated using the synthetic data generator described in Section 6.2.2, with each dataset representing 3 participants and realistic sensor and self-report patterns. Metrics were captured using Prometheus monitoring[1] and Grafana instrumentation[2] at the TimescaleDB repository layer.

Test Environment

All tests were conducted on a dedicated database server with the following specifications:

- **CPU**: 12th Gen Intel(R) Core(TM) i7-1260P (16 cores)
- **Architecture**: x86_64
- **Memory**: 32 GB RAM
- **Storage**: NVMe SSD
- **Database**: TimescaleDB 2.x on PostgreSQL 14

Test Configuration

Each test simulated a realistic multi-study deployment with three concurrent users:

- **User A**: Accessed a large dataset of varying size (70, 100, 150, or 200 days)
- **User B**: Accessed a 70-day dataset (baseline)
- **User C**: Accessed a 70-day dataset (baseline)

All users executed dashboard queries simultaneously during each 30-minute measurement window.

For the 200-day dataset, we recorded two measurements:

[1] `https://prometheus.io/`
[2] `https://grafana.com/`

- **Cold cache**: Immediately after database restart, representing the worst-case initial load
- **Warm cache**: After 30 minutes of continuous use, representing steady-state operation

Results

Table 5.1 summarises aggregate query performance across all three concurrent users for each dataset size.

Table 5.1: Aggregate query performance under concurrent multi-user workload

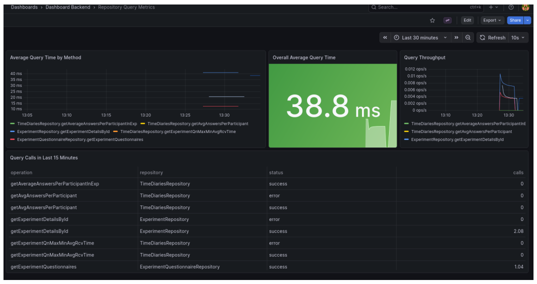
Metric	**70 Days**	**100 Days**	**150 Days**	**200 Days**
Average query time (warm)	38.8 ms	58.3 ms	131 ms	103 ms
Average query time (cold)	—	—	—	1,130 ms
95th percentile latency (warm)	52 ms	78 ms	187 ms	125 ms
95th percentile latency (cold)	—	—	—	2,450 ms
Peak latency (warm)	89 ms	124 ms	245 ms	250 ms
Peak latency (cold)	—	—	—	25,000 ms
Query throughput	0.012 ops/s	0.008 ops/s	0.008 ops/s	0.012 ops/s
Error rate	0.4%	0%	0%	0%

Analysis

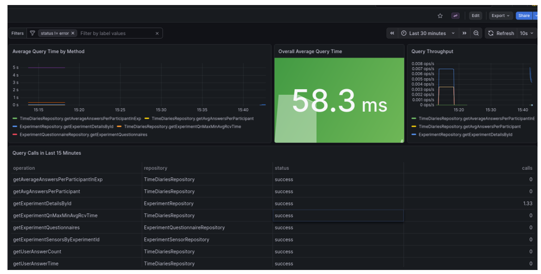
Warm cache results show that the architecture maintains sub-150 ms average query response times for all dataset sizes up to 200 days under steady-state conditions. The 200-day warm cache result (103 ms) is comparable to the 70-day baseline (38.8 ms), confirming that:

- TimescaleDB's hypertable partitioning scales effectively with data volume
- Indexing strategies maintain efficiency regardless of dataset size
- The system supports concurrent multi-user workloads without significant performance degradation
- With 32 GB of RAM, datasets up to 200 days remain fully cacheable in steady-state operation

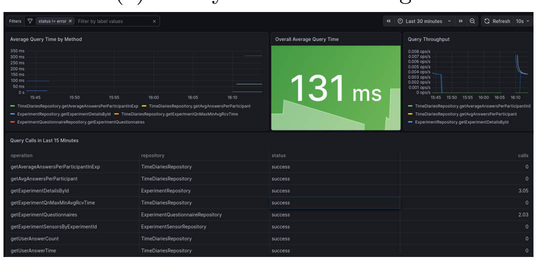
The cold cache measurement for the 200-day dataset shows the expected performance impact of initial database load. The 25-second peak reflects a case where a complex query required full dataset scanning before indexes and data were loaded into memory. This behaviour is typical for databases and does not affect steady-state operation. Once warmed, the 32 GB memory configuration is sufficient to maintain the working set for datasets up to 200 days.



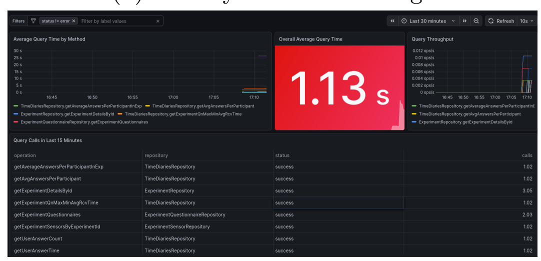
(a) 70 days: 38.8 ms average



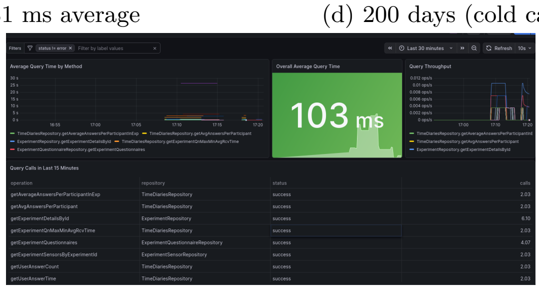
(b) 100 days: 58.3 ms average



(c) 150 days: 131 ms average



(d) 200 days (cold cache): 1,130 ms average



(e) 200 days (warm cache): 103 ms average

Figure 5.11: Grafana dashboard showing repository query metrics across dataset sizes under concurrent multi-user workload

Query throughput remained stable at 0.008–0.012 operations per second across all dataset sizes and cache states, indicating that the system's processing capacity is not saturated. Error rates were 0% for all warm cache runs, with only transient errors (0.4%) during the initial 70-day baseline setup.

Implications for Deployment

These results demonstrate that the system can support real-world deployments as follows:

- For studies up to 200 days, the dashboard remains responsive (average <150 ms) in steady-state operation
- A brief warm-up period (30 minutes) after database restarts is recommended
- The system handles concurrent studies of different durations without performance degradation

Conclusion

The evaluation confirms that the architecture meets the performance expectations for large-scale data. The system maintains interactive dashboard performance for datasets up to 200 days under realistic multi-user workloads, with cold cache queries requiring a short warm-up period. The tested hardware provides sufficient resources for deployments at this scale.

Chapter 6

System Evaluation

To evaluate the design and utility of the dashboard, two case studies were conducted. As stated earlier, the dashboard system has two user roles: researcher and participant (see 4.3); accordingly, the two case studies corresponded to these roles, and thus there were two dashboard evaluations.

6.1 Evaluation Framework and Objectives

We evaluated whether our *dual-role* calendar dashboard works effectively for both participants and researchers in real-life study conditions. Beyond basic functionality testing, we assessed whether the system meets its human-centred and data-quality goals. The evaluation focused on three main questions:

1. **Participant-Facing Utility:** To determine whether non-expert users can easily understand and navigate the participant dashboard, examining whether the integration of multimodal data (sensor and self-report) within a calendar-centric interface successfully enhances personal awareness and sense-making.
2. **Researcher-Facing effectiveness:** To assess how effective the dashboard is in a research study, specifically in enabling quality-aware data collection, sustaining participant engagement, and facilitating proactive study management.
3. **System Validation:** To test the overall system's robustness and design validity under both controlled conditions (using a synthetic dataset) and authentic, real-world conditions (6.3), thereby providing evidence for its generalisability and practical application.

The following two case studies were structured to address these objectives independently and collectively, providing a comprehensive assessment from both ends of the user spectrum.

6.2 Case Study 1: Participant Dashboard Evaluation

6.2.1 Introduction and Objectives

This case study was carried out to evaluate the participant section of the dashboard. It was used to examine how participants explore and interpret their personal data when it is presented through our dashboard interface. The primary goal of this study was to examine how combining sensor-based and self-reported data, contextualised within a calendar interface, can enhance participants' awareness of their daily routines, activity patterns, and overall life balance.

The study evaluated the participant calendar dashboard, which integrated time-use information, sensor streams (e.g., the number of steps counted and other passively collected signals), and self-reported diary entries into a unified interface. The primary objectives of the study were:

- To assess the dashboard's usability and clarity for non-expert users (participants) who are using it;
- To evaluate the degree to which the dashboard supports reflective sense-making of everyday patterns; and
- To examine user perceptions of data quality, interpretability, and trust when confronted with personal multimodal, 'big-thick' data.

The evaluation reported here draws on usage data from 130 study participants who were asked to review a synthetic 70-day dataset generated from a single user's sensor and answer data via the dashboard, then complete a structured questionnaire with both Likert-scale and open-ended questions (.1).

6.2.2 Study Design and Procedures

Materials: The Synthetic Dataset

Overview of the Synthetic Data Generation We used a synthetic dataset to evaluate the dashboard due to privacy concerns. A data generator code simulated realistic human routines by combining temporal context, mobility patterns, and daily life engagement. Each generated record represents a time-stamped diary entry, including activity, location, steps, mood, connectivity (Wi-Fi and Bluetooth), and use of phone applications.

The generator produced realistic personal-context data for use in testing. It simulated daily routines, movement, social interactions, and smartphone use, allowing us to visualise sensor-derived inferences and integrate self-reports without real participant data.

The design integrates three layers of real life:

- **Temporal regularity** – activities were tied to specific time slots (e.g., ”sleeping” during night hours and ”study/work” during daytime).
- **Context alignment** – each activity type had a limited set of semantically valid locations (e.g., ”lecture/seminar” occurs in ”work” clusters like university buildings).
- **Sensor coupling** – each event carried simulated sensor metrics such as GPS coordinates, step counts, screen states, and app usage.

Temporal and Activity Modelling The core of the data generator relied on time-context-activity scheduling. The twenty-four hours in a day were divided into minute-level timestamps, with each interval probabilistically associated with a dominant activity category (e.g., sleeping, working, commuting, studying, or eating).

Activity selection then followed a time-conditioned categorical probability model. For example, the likelihood of *sleeping* was considered mostly between 00:00–06:00, while *work* or *lecture/seminar* dominated mid-morning to early afternoon. Transitional activities (e.g., *travel* or *Break*) emerged between major blocks to simulate a realistic daily rhythm.

Each generated activity inherited semantic attributes, including:

- **Activity label** (e.g., *Lecture/Seminar*, *Social life*, *Work-related task)*);
- **Expected duration**, normally distributed within activity-specific bounds;
- **Energy usage (calories) and mobility potential**, informing step generation;
- **Social context**, affecting companion and mood fields.
- **internal state** e.g mood

This hierarchical structure ensured that the synthetic data mirrored patterns observed in real-time use studies.

Spatial and Mobility Simulation The spatial component of the synthetic dataset was grounded in real-world geographic coordinates derived from the **London metropolitan area**. We chose London because of its diversity of urban functions and daily mobility patterns. It provides a heterogeneous landscape comprising residential, educational, commercial, and recreational zones in close spatial proximity, making it well-suited for modelling a dataset with realistic daily transitions.

To contextualise each activity, we defined **semantic location clusters** corresponding to typical time-use environments. Each cluster comprised a set of geographic coordinates representing plausible venues for that activity type. The pipeline’s model dynamically selected one location within a cluster based on the current activity, the participant’s previous position, and the time.

Table 6.1: Representative location clusters and example coordinates in London.

Cluster	**Example Locations (latitude, longitude)**	**Representative Activities**
Home / Residence	51.5234 N, -0.1321 E (Camden Town); 51.5096 N, -0.1180 E (South Bank)	Sleeping, Personal care, Relaxing
University / Work	51.5246 N, -0.1342 E (UCL Main Campus, Bloomsbury); 51.5179 N, -0.0911 E (City office cluster)	Lecture/Seminar, Research, Meetings
Leisure / Social	51.5076 N, -0.1276 E (Trafalgar Square); 51.5094 N, -0.0760 E (Tower Bridge area); 51.5426 N, -0.1462 E (Camden Market)	Social life, Party, Shopping, Dining
Recreation / Outdoor	51.5064 N, -0.1657 E (Hyde Park); 51.5320 N, -0.1236 E (Regent's Park); 51.4684 N, -0.1698 E (Battersea Park)	Exercise, Sports, Walking
Cafés / Restaurants	51.5133 N, -0.1368 E (Soho cafés); 51.5008 N, -0.1247 E (Westminster eateries)	Eating, Socializing, Meetings
Shops / Errands	51.5143 N, -0.1410 E (Oxford Street); 51.5219 N, -0.1073 E (Angel Islington)	Shopping, Groceries, Errands

When an activity transition implied movement across clusters (for instance, from *Home* to *University* or *University* to *Leisure*), the generator computed a geographical distance between the previous and new coordinates using the **Haversine formula** [1]:

$$d_k = 2r \arcsin\left(\sqrt{\sin^2\frac{\Delta\phi}{2} + \cos\phi_1\cos\phi_2\sin^2\frac{\Delta\lambda}{2}}\right) \tag{6.1}$$

Where:

- d_k is the great-circle distance between two points on the Earth's surface (in kilometres),
- r is the Earth's mean radius, taken as 6,371 km,
- ϕ_1 and ϕ_2 are the latitudes of the first and second points (in radians),
- λ_1 and λ_2 are the longitudes of the first and second points (in radians),
- $\Delta\phi = \phi_2 - \phi_1$ is the difference in latitude,

[1] Haversine formula: https://en.wikipedia.org/wiki/Haversine_formula

- $\Delta\lambda = \lambda_2 - \lambda_1$ is the difference in longitude, and
- arcsin denotes the inverse sine function.

This distance directly influenced the number of steps attributed to the transition according to the following formula:

$$\text{steps} = \frac{d_m}{0.78} \times k + \epsilon \tag{6.2}$$

Where:

- d_m is the distance travelled, measured in meters,
- 0.78 represents the average step length in meters (based on adult walking averages),
- k is an activity-dependent scaling factor, typically ranging from 1.2 to 1.5 to account for walking intensity or terrain,
- ϵ denotes random Gaussian noise used to simulate intra-location movements (e.g., walking within classrooms, shops, or halls).

This design ensured that even short movements, e.g., around campus, yielded realistic step counts. At the same time, larger transitions, such as commuting from *Camden Town* to the *UCL campus* or travelling from work to leisure zones, produced proportionally higher step counts.

By basing all location coordinates in London, the model captured urban-scale variability and allowed visualisation of *activity flows* across a real metropolitan geography, enhancing the interpretability and credibility of the generated dataset.

App Usage and Digital Context Simulation To simulate digital engagement, each activity triggered a weighted random selection of phone applications likely to be used in that context.

For instance:

- *Social Life / Leisure* → Instagram, WhatsApp, TikTok;
- *Work / Study* → Zoom, Google Docs, YouTube (educational), Slack;
- *Entertainment / Relaxation* → YouTube, Netflix, Spotify;
- *Commuting / Waiting* → Music or messaging apps.

Each record included 1–3 applications randomly drawn from the context-relevant subset, capturing the diversity and multitasking behaviour typical in real-life smartphone use.

Affective Context and Companionship Every generated record also encoded **mood** and **companionship** to reflect socio-emotional dimensions of activity engagement. Mood was represented using five categorical emojis (😡, 😕, 😐, 🙂, 😀), corresponding to a simple ordinal affective scale from negative to positive valence.

The assignment of mood followed a weighted schema based on activity type:

- *Social or Leisure* activities favoured positive moods (🙂 – 😀)
- *Work or Study* activities clustered around neutral (😐 – 🙂)
- *Stressful or personal* tasks included negative moods (😡 – 😐).

Companion selection followed similar logic, mapping activities to likely social configurations (e.g., *alone*, *with family*, *with friends*, *with colleagues*).

Temporal Realism and Behavioural Dynamics To maintain temporal realism, each answer record included three critical timestamps:

1. **Notification time** – representing when a time-diary prompt is issued;
2. **Answer time** – when the participant responds;
3. **Delta** – the interval between opening a time diary to answer and when it is submitted.

Each sensor data record also had a timestamp linked to the exact time it was generated.

To mimic human variability, timestamps were jittered within empirically realistic intervals (1–30 seconds for notifications, 1–500 seconds for responses). Activity duration and inter-event intervals were drawn from normal distributions, consistent with observed human routines. The generator thus produced smooth temporal continuity with stochastic deviations, yielding data streams suitable for testing prediction, summarisation, and dashboard visualisation algorithms.

Implementation and Output Structure The data generator was used to generate a single user's data spanning a period of 70 days. The synthetic data records generated contained the fields shown in Table 6.2:

Sensor Data in the Synthetic Dataset To make sure that the synthetic dataset was realistic and made sense in context, we included multiple smartphone sensor streams to mimic those captured during experiments run on iLog. These sensors provide the basis for modelling digital interaction, mobility, social closeness, and environmental context. In practice, such data would be collected continuously by background services or dedicated applications. In this dataset, they were generated based on the activity type, time of day, and transitions between locations. They are discussed as follows:

Table 6.2: Synthetic data schema.

Field	Description
experimentid	Experiment or simulation identifier
instanceid	Unique instance per timestamp
instancetimestamp	Realistic datetime in ISO format
notificationtimestamp	Time when the notification or prompt was sent
answertimestamp	Time when the participant submitted a response
userid	Synthetic user identifier
latitude, longitude	Real geographic coordinates
activity, location, companion	Semantic activity context
mood	Emoji-coded affective state
steps	Estimated step count for the interval
appsused	List of applications accessed
answerduration, delta	Response and transition times

Screen Status The *screen status* sensor is among the most informative indicators of device use and attention. Real smartphones record this data through system-level broadcasts such as `SCREEN_ON`, `SCREEN_OFF`, or `USER_PRESENT`.

In our simulation, each activity period was annotated with a binary flag indicating whether the device was likely actively used (`SCREEN_ON`) or idle (`SCREEN_OFF`):

- **Realistic Modelling:** The probability of the screen being on was dependent on the activity type. For example:
  - Social and communication activities (e.g., chatting, using WhatsApp or Instagram) had frequent and longer `SCREEN_ON` intervals.
  - Work and study activities exhibited intermittent screen use, consistent with note-taking or reference-checking patterns.
  - Sleep, personal care, and passive transport periods were mostly logged as `SCREEN_OFF`.
- **Temporal Consistency:** Each activity segment incorporated brief bursts of `SCREEN_ON` to account for quick checks or notifications, even when engagement was low.

This simulation closely mimics real-world mobile interaction rhythms, in which screen events occur hundreds of times per day but vary with users' routines and contexts.

Bluetooth Sensor The *Bluetooth sensor* captures the proximity of connected and nearby devices through periodic scanning, returning a list of detected devices and their signal strengths (RSSI). In reality, this is used in contact tracing, social sensing, and indoor localisation.

For the synthetic dataset, this is represented in:

- Device Proximity Simulation: Each activity type was assigned a likelihood of detecting nearby devices. Social gatherings, lectures, and workplace activities resulted in higher device densities, indicating co-located users. Home or solitary tasks yielded fewer detections.
- Signal Strength Distribution: RSSI values were sampled from realistic distributions (-30 to -90 dBm), with stronger signals during close-range interactions.
- Dynamic Context: Temporal continuity was ensured by reusing the same ”neighbouring devices” within social clusters, creating persistent but variable proximity patterns.

This approach realistically mirrored how Bluetooth Low Energy (BLE) scans would appear in an urban or university setting, with fluctuating proximity intensities reflecting crowd density or mobility.

Wi-Fi Connectivity Wi-Fi data provides both *location cues* and *environmental stability*. In real devices, Wi-Fi scanning detects available networks (SSIDs), signal strengths, and connection events.

Within the synthetic dataset:

- **Location Anchoring:** Distinct Wi-Fi SSIDs were allocated to key spatial categories, including home, university, cafeteria, and public space. Consequently, a Wi-Fi transition was initiated when the user traversed these locations, thereby signifying a shift in context.
- **Connection Patterns:** Strong signal strength and long connection times indicated indoor, stationary situations, like working from home. In contrast, weak or unstable signals suggested movement or outdoor activities.
- **Temporal Realism:** Connection events were aligned with screen activity; users were typically connected while the screen was on during work or media sessions, but not while commuting.

This synthetic modelling provides location awareness without using GPS, reflecting how real-world context inference often relies on Wi-Fi network traces.

Step Counter Sensor The *step counter*, a feature built from accelerometer data, keeps track of the total steps taken since the device was last restarted. It’s a dependable sensor, often used to measure physical activity.

In the simulated dataset, this was represented by:

- **Movement Generation:** Step increments were calculated based on the *distance between consecutive locations*, the*type of activity*, and *duration* of the activity.. Walking, commuting, or exercise activities yielded thousands of steps over short intervals. Stationary activities often required minimal, if any, movement. Small-step bursts were accompanied by short intra-location transitions (e.g., moving between classrooms or rooms) to ensure realism.

- **Context Coupling:** Steps were correlated with changes in Wi-Fi networks or Bluetooth encounters, capturing the multi-sensor consistency seen in real data streams.

This modelling ensured that spatial movement is quantitatively represented, making the synthetic dataset physically plausible and compatible with activity recognition algorithms.

Integrated Sensor Logic Each sensor was designed to operate semi-independently while synchronising timestamps, enabling multimodal correlations. Table 6.3 summarises typical cross-sensor configurations for everyday contexts.

Table 6.3: Example of integrated sensor relationships across contexts.

Context	**Screen**	**Bluetooth**	**Wi-Fi**	**Steps**	**Interpretation**
Home at night	OFF	Few devices	Stable SSID	0–10	Sleep/rest
Lecture	ON/OFF	5–10 devices	University SSID	100–400	Study/work
Walking	OFF	Few devices	Transition	500–1500	Commuting
Party	ON (frequent)	15+ devices	Event SSID	200–800	Social engagement

In real smartphone data, sensor signals often change together. For example, increased steps may occur alongside Wi-Fi transitions or Bluetooth detections. The synthetic dataset thus achieves a high degree of *situational realism*, enabling its use for modelling context-aware systems.

Summary of Synthetic Dataset The generator produced a realistic dataset that could be used to test the dashboard without relying on real participant data. By embedding realistic mobility trajectories, digital engagement dynamics, and emotional context, the simulation effectively bridged the gap between purely synthetic data and empirical field studies. The resulting dataset provided an ecologically valid foundation for evaluating the dashboard's functionality and user experience under controlled yet contextually plausible conditions. Once the data had been generated and securely stored in the back-end database, the dashboard's front-end was integrated with the data layer, rendering the complete system ready for testing, visualisation, and participant interaction.

Recruitment and Participants

Participants were recruited via Prolific, an online research platform that connects researchers with verified participants for academic and commercial studies. The participants completed the study remotely on their own devices. After cleaning the collected data from the dashboard usage study, the analysis sample comprised 130 participants (final N = 130). The participants predominantly used laptop/desktop devices for the task (115 participants, 88.5%); mobile phones (9 participants, 6.9%) and tablets (6 participants, 4.6%) accounted for the remainder.

Participants varied in age and background; the dataset's demographic fields indicate a range of fluency and prior exposure to digital tracking tools (see Figure 6.1). More details are discussed in section 6.2.3.

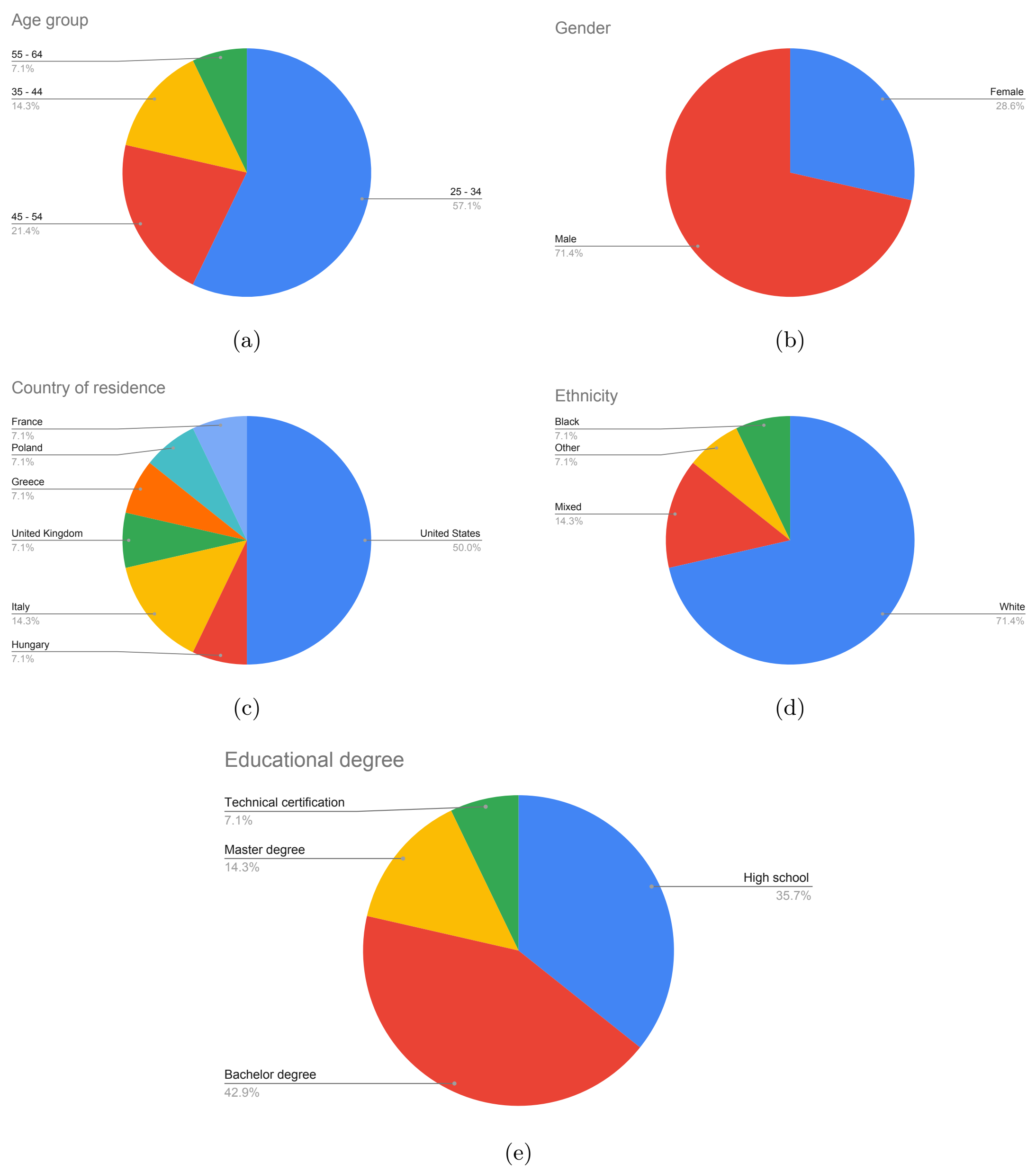


Figure 6.1: Participant demographics, showing the distribution of age groups, gender, country of residence, ethnicity, and educational attainment in the study sample.

Experiment Procedure

Dashboard Evaluation Procedure Each participant first interacted with the dashboard, exploring its features and visualisations at their own pace. During this interaction, user behaviour was tracked using Matomo, an analytics platform that captured click events during page navigation, page views, navigation paths, and time spent on different sections of the dashboard. This allowed us to measure how participants interacted with the dashboard and compare these behaviours with their questionnaire responses.

After the experiment, participants completed a structured questionnaire. This questionnaire was designed to gather both quantitative and qualitative feedback.

The questionnaire had three main parts:

1. Likert-scale items (1–5) addressing:
   - Usability, such as the ease of navigating to different locations and data views.
   - Reflective potential, including whether the visualisations supported self-reflection or understanding of social interactions.
   - Recommendability, for instance, the statement: ”I would recommend a dashboard like this to others.”
2. Contextual perception items, which assessed participants' views on the usefulness of the dashboard in relation to their material and social contexts, including which types of information they found most meaningful.
3. Open-ended prompts, inviting participants to describe:
   - The features they found most helpful,
   - Any confusing or unclear aspects, and
   - Comparisons to other tools or dashboards they have previously used.

For this study, we focused on key Likert-scale items that directly reflect user engagement and perceived utility:

- **Recommendation:** ”I would recommend a dashboard like this to others.”
- **Ease of navigation (locations):** ”It was easy to navigate the dashboard to find information about locations.”
- **Reflection (role of others):** ”The visualisations helped me reflect on the role of others in the activities.”

The full questionnaire used in the study can be found in Appendix .1.

By combining self-reported feedback with objective interaction data from Matomo, the evaluation captured both functional usability and actual engagement behaviour, enabling a richer understanding of how participants used and experienced the dashboard.

6.2.3 Results

Quantitative Findings

A total of 130 participants completed the dashboard evaluation. The sample included 71 males and 59 females, aged 20 to 75, from Europe and the Americas. In terms of education, the majority held at least a bachelor's degree (41%), followed by high school (22%), master's degree (20%), technical certification (12%), and a smaller proportion with doctoral degrees (3%).

Their employment status was mixed: 46 participants reported full-time employment, 12 reported part-time employment, and 8 indicated they were not in paid work (e.g., homemakers, retired, or disabled). 13 respondents were students at the time of participation.

Each participant interacted with the dashboard, explored its features, and completed the *post-task* questionnaire (.1); and their interactions while using the dashboard were logged using the Matomo analytics library [2].

During the experiment, participants were required to rate multiple aspects of the dashboard using 5-point Likert scales (1 = strongly disagree, 5 = strongly agree). The key items analysed include usability, reflection, trust, confidence, and overall recommendation. Table 6.4 summarises these ratings.

Table 6.4: Summary of Likert-scale responses for dashboard evaluation (N=130).

Item	**Mean**	**SD**	**Min**	**25%**	**Median**	**Max**
It was easy to navigate the dashboard to find information about locations.	4.30	0.84	1	4	5	5
The visualisations helped me reflect on the role of others in the activities.	3.78	1.19	1	3	4	5
I would recommend a dashboard like this to others.	3.46	1.36	1	3	4	5
I trusted the dashboard to represent the data in a meaningful way.	4.03	1.01	1	4	4	5
I felt confident using the dashboard to reflect on the data.	4.00	1.06	1	4	4	5

Overall, participants expressed high levels of confidence and trust in the dashboard's data representation, with mean scores above 4.0. Navigation scored the highest (M = 4.30, SD = 0.84), indicating that most participants found the interface intuitive and easy to use. Reflection on the role of others scored moderately high (M = 3.78), demonstrating that the visualisations encouraged self-reflection, though with some variation across users.

The recommendation item, with a mean of 3.46 and a standard deviation of 1.36, received a somewhat lower score. This shows that while participants generally viewed the system positively,

[2] https://matomo.org/

they saw it more as a tool for research or exploration, rather than something they would actively recommend to others.

Qualitative Insights

Open-ended responses provided additional explanations for the quantitative results. Thematic analysis revealed several recurring themes:

- **Clarity and visual appeal:** Many users praised the "colour-coded graphs", "clear icons", and "intuitive layout", describing the dashboard as "visually appealing" and "easy to understand at a glance". One participant noted: *"The visuals made it easy to see what I was looking at and how it compared to other information."*

- **Reflective depth and self-awareness:** Participants frequently mentioned that the dashboard helped them recognise patterns in their behaviour, mood, and social interactions. Comments included: *"It helped me reflect on how much time I spend alone vs. with others,"* and *"I could see how my mood changed based on activities and who I was with."*

- **Information overload and complexity:** A significant number of users reported feeling overwhelmed by the volume of data and visual elements. Responses such as *"There was so much information—it was very well segregated but a lot to pore over"* and *"It felt cluttered and too busy"* highlighted usability challenges.

- **Comparative value over existing tools:** Many users found the dashboard superior to existing applications like Apple Health or Google Fit, noting its comprehensive approach. They said, "This dashboard provides much more information than other trackers," and "It's more thorough and not just focused on physical health."

- **Desire for simplification and guidance:** Users recommended improvements such as "a tutorial with audio" and "clearer icons for moods". Some expressed that the dashboard "would need a 'how-to-use' guide" to be fully accessible.

- **Trust and data accuracy:** Trust was bolstered by the "detailed and consistent data presentation", though some questioned the feasibility of manual input: *"It would take a lot of time and effort to maintain this daily."*

A complete set of results is provided in Appendix .2 (Tables 1–6).

Matomo Usage Statistics

To quantify how participants interacted with the system during the study, we tracked their actions while using the dashboard through a tracking library called *Matomo analytics*. We then analysed the complete set of usage logs, which recorded page visits, actions, and time-on-page metrics for all 130 users. These logs allowed us to compute both *per-user engagement levels* and *page-level dwell times* across the entire experiment.

The average time each user spent while using the system was 818.85 seconds, or about 13.6 minutes. This shows that participants used the platform for a considerable amount of time, even though the interaction tasks were relatively simple. A page-level analysis revealed apparent differences in users' navigation and exploration of the platform. The pages with the highest cumulative time spent were as shown in Table 6.5:

Table 6.5: Common participant pages and total time spent

URL part	**Description**	**Cumulative Time spent**
`participant_data`	Page with graphs displaying all the collected data	33,059 seconds (≈ 551 minutes)
`participant_past_map`	Map page with locations and filtering by activities and other contextual dimensions	22,784 seconds (≈ 380 minutes)
`participant_past_calendar`	Page with past events on the calendar, showing contextual dimensions and sensor icons	22,637 seconds (≈ 377 minutes)
`participant_dashboard`	Experiment summary page shown when an experiment is selected	15,998 seconds (≈ 267 minutes)
`participant_calendar`	Calendar page containing future events and questions to answer	11,653 seconds (≈ 194 minutes)

Pages related to navigation or system entry, such as login pages, accumulated significantly less time (e.g., 4,410 seconds ≈ 74 minutes for the `/login` link), as expected.

The disproportionately high engagement with the `participant_data`, `past map`, and `past calendar` pages (see Table 6.5) demonstrates that users dedicated substantial time to exploring the historical and contextual data. These components contain the densest visual information, including timelines, geospatial maps, and summary statistics. Their dominant share of total time spent indicates that the personal data visualisation elements were central to participant engagement, aligning with the goals of supporting self-reflection and data awareness.

In contrast, the shorter "dwell times" on simpler interface elements, such as the main dashboard or landing pages, demonstrate that the navigation was efficient. This means participants could quickly access their data without unnecessary delays.

These Matomo logs show that participants actively explored the dashboard during the study, indicating that participants not only accessed the platform but also engaged deeply with its visual analytics features, particularly the sections that present historical and context-aware personal data.

Summary

Across all thematic categories, respondents generally rated the dashboard positively, with mean scores ranging from 3.5 to 4.5. The highest-rated items were related to **activity differentiation** (”I could distinguish between different types of activities,” $M = 4.54$) and **spatial understanding** (”The visualisation of locations helped me understand the broader context of activities,” $M = 4.35$), indicating strong performance in visual clarity and categorical distinction. Items concerning **material content** received the lowest ratings (means ranging from 3.62 to 3.68), showing that this aspect of the dashboard was less salient or less clearly communicated. **Recommendation and personal adoption intent** scored moderately ($M = 3.46$ and $M = 3.31$, respectively), reflecting a perceived utility that was tempered by considerations of complexity or integration into daily routine. Overall, the dashboard excelled in supporting **visual interpretation and contextual awareness**, while areas related to **object-based context and long-term engagement** present opportunities for refinement.

The study indicated that participants responded positively to the dashboard, with high ratings for usability, reliability, and perceived security. A lot of them said they appreciated how it helped with reflection and how it pulled together data from different sources. They typically felt it was better than other personal tracking tools out there. But also, there were some real concerns around information density and accessibility. Participants recommended that we could improve things by simplifying the design, making onboarding easier, and better accounting for the needs of neurodivergent users.

Looking at the Matomo logs, we saw that participants spent a good amount of time engaging with the data-rich parts of the dashboard. That tells us they were actively digging into the behaviour patterns of our synthetic user, a good sign that the system can support self-reflection and awareness. For future versions, we shall aim to make a better balance between analytical depth and clarity. The goal is to make the dashboard easier to navigate while keeping the reflective qualities that people found valuable.

Design Implications: These findings point to clear pathways for improving the dashboard. Future versions will prioritise *adaptive complexity*: a default ”summary view” with progressive disclosure of details, more filtering and customisation options, and improved onboarding (e.g., interactive tutorials). Furthermore, exploring alternative visual encodings for different data types, like material context, will be considered to help balance the depth-brevity trade-off, making the dashboard's reflective potential accessible to a broader range of cognitive styles without diminishing its overall power.

6.3 Case Study 2: Researcher Dashboard Evaluation with MakOne Dataset

Introduction

This case study describes the design, deployment, and evaluation of the MakOne study, a large-scale behavioural data collection project conducted at Makerere University in Kampala. The study investigates whether a real-time monitoring dashboard can support data quality, sustain participant engagement, and improve the reliability of longitudinal data collection. It also demonstrates how combining mobile sensing, self-report surveys, and dashboard-based quality feedback can support sustained data collection practices in an academic setting.

The data for this experiment was collected using the iLog mobile platform, which is available on Android and iOS devices. This platform combined passive smartphone sensing with ecological momentary assessments (EMAs). The accompanying monitoring dashboard enabled researchers to visualise data inflow in real time, identify inactive participants, detect sensor dropouts, and take corrective actions. This combination of mobile data collection and real-time visualisation allowed for proactive management of participant engagement while safeguarding the integrity and continuity of the dataset.

The study ran for six weeks, from March 18 to April 29, 2024, and formed part of a broader effort to build a high-quality behavioural dataset for understanding student life in resource-constrained contexts. The MakOne study provides a practical demonstration of how quality-aware data-collection principles, developed within the iLog framework, can be operationalised in real-world research involving diverse participants and dynamic data conditions. The resulting dataset is described in [36].

6.3.1 Participants and Recruitment

Participants were undergraduate, postgraduate, and graduate students drawn from several colleges of Makerere University, ensuring broad representation across academic disciplines, demographic groups, and daily routines. The sampling strategy aimed to capture variation in students' lives across faculties, including Computing, Engineering, Humanities, Business, and Health Sciences. This diversity was essential for generating a behavioural dataset that supports generalisable insights into student mobility, routines, and digital engagement patterns.

The recruitment process spanned a fortnight, leveraging various institutional communication channels to broaden its reach. Announcements were distributed via university noticeboards, student mailing lists, faculty WhatsApp groups, and official social media pages. Students who expressed interest scanned a QR code, which then redirected them to a specialised study webpage.

This webpage served as a centralised resource, offering all necessary information. Specifically, the webpage furnished exhaustive details concerning the study, encompassing:

- installation guidelines for both Android and iOS versions of the iLog application;
- technical specifications and a list of supported device models;
- explanations of the data acquisition methods, including passive sensing and ecological momentary assessments (EMAs);
- instructions for granting and sustaining the necessary permissions, privacy protections, data management protocols, and the rights of participants;
- and the eligibility requirements, which stipulated current student enrolment and possession of a compatible smartphone.

The eligibility criteria were inclusive; specifically, participants needed to be registered at Makerere University throughout the study's duration, possess a compatible Android or iOS device, and give informed consent as stipulated by the sanctioned ethical protocol. Students utilising older or less capable devices received troubleshooting assistance through the designated webpage and service desk, for example, instructions for modifying battery configurations or activating background permissions.

A total of 101 students were enrolled, though only 73 consented to their data being collected and kept, consistent with recruitment capacity and expectations for typical longitudinal attrition rates (30–50%). Participants represented a range of residential arrangements—including campus halls, private hostels, and off-campus dwellings—allowing the study to capture behavioural patterns across different spatial and social living contexts. This diversity was particularly valuable for analysing mobility behaviour, class attendance patterns, and differences in daily routines between on-campus and off-campus students.

The Makerere University Institutional Review Board approved the study after ensuring ethical standards were met. The study's website made it clear how data would be managed, that participation was entirely voluntary, and that anyone could drop out whenever they chose. Those who finished the study received a small payment. By centralising documentation and instructions through a single online resource, the study ensured that participants were well informed and able to self-manage installation and maintenance with minimal training/technical help.

6.3.2 Data Collection and Instruments

The iLog platform was used to collect both self-reported and passive sensor data. It was configured to function reliably under variable network connectivity, a critical feature in the study environment (Uganda). Data were stored locally on participants' devices and automatically uploaded when connectivity became available. The platform supported both Android and iOS devices, with adaptations to handle iOS background-task restrictions and Android battery-optimisation behaviours, ensuring continuous data capture without disrupting participants' regular device usage.

Questionnaires

Two types of questionnaires were administered to capture static and dynamic aspects of student life:

- **Profile/Baseline Survey:** This was administered at enrolment, i.e., when registering in the iLog app for the experiment. It collected essential demographic and contextual information, including age, gender, faculty, year of study, residential arrangements (campus hall, private hostel, or off-campus housing), and lifestyle details. This survey provided crucial context for interpreting subsequent behavioural and sensor data.
- **Ecological Momentary Assessments (EMAs):** These were delivered via in-app prompts; EMAs collected real-time contextual information about ongoing activities, social interactions, moods, and locations. Questions were adapted to local student life, covering experiences such as attending lectures, commuting, meals, studying, and leisure activities. EMAs were designed to be short and minimally intrusive, reducing participant burden while providing high temporal resolution.

Daily Diaries

The Daily Diaries module captures temporal variation in participants' routines and subjective experiences. It complements the profile survey and passive sensor streams by providing daily self-reported contextual data. Full details of the questionnaire are provided in Appendix .3.

The module consisted of four diary types:

- **Morning Question (Appendix 8):** A single-question diary administered each morning asking, "How was your day?" Participants responded using a five-point emoji-based scale ranging from 😀 very good to 😡 very bad. This provided a low-burden measure of overnight well-being and mood upon waking, offering baseline information contextualised by demographic and lifestyle data from the profile survey.
- **Evening Question (Appendix 9):** A single-question diary administered each evening using the same emoji-based scale as the morning question. This allowed participants to reflect on their day and provided complementary data to the morning check, enabling analysis of day-to-day mood fluctuations and their associations with activities, social interactions, and environmental context captured through passive sensing and EMAs.
- **Time Diaries (Appendix 7):** Triggered at 30-minute intervals, these questions captured participants' ongoing activity, location, social company, and mood. Time Diaries offered high temporal resolution and enabled detailed examination of daily routines, social patterns, and behavioural contexts across the day. They were carefully scheduled to minimise disruption, and participants could temporarily disable prompts during lectures, exams, or other essential commitments.

- **Snack Diaries (Appendix 12):** These were sent every two hours. The snack question recorded informal eating and drinking episodes outside main meals. Each prompt consisted of two sequential questions:

  1. *Occurrence of Snacks/Drinks:* Participants reported whether they had consumed any snacks or drinks in the previous two hours, selecting time-specific options such as *"No"*, *"Yes, between now and 30 minutes ago"*, or *"Yes, between 1.5 and 2 hours ago"*.
  2. *Type of Food or Drink:* If a snack was reported, participants selected all applicable items from a predefined multiple-choice list including 26 categories (e.g., confectionery, sandwiches, fruits, vegetables, dairy, meat, beverages, alcoholic drinks) and an "Other food" option.

  This two-step design captured both the timing and content of snacking behaviour, allowing analysis of correlations with mood, activities, and sensor-detected contexts.

By combining low-burden mood checks (morning and evening), high-resolution Time Diaries, and detailed Snack Diaries, the Daily Diaries module provided a comprehensive picture of participants' daily experiences. These self-reports, contextualised by demographic and lifestyle information from the profile survey, were integrated with passive sensor streams and EMAs within the iLog platform. The resulting multimodal dataset enabled quality-aware longitudinal analyses of student behaviour, social interactions, mobility, and well-being.

Sensor Data Streams

Passive sensing in the iLog app captured rich, multimodal data from both Android and iOS devices, providing a continuous, objective view of participants' behaviours and environments. These data streams, along with self-reported diaries and ecological momentary assessments (EMAs), allowed for detailed, cross-validated analyses of daily activities, movement patterns, and environmental factors. The sensors included:

- **Device Interaction:** This data included battery level, screen-on/off events, and touch interactions. These signals helped track phone engagement, detect periods of inactivity, and infer daily rhythms in device use. Platform-specific handling ensured accurate logging while respecting iOS background-task limitations and Android battery-optimisation settings.
- **Motion and Physical Activity:** Accelerometer and gyroscope readings, step counts, and derived activity types (e.g., walking, running, stationary) were collected continuously. These streams supported fine-grained mobility analyses, including temporal patterns of movement and sedentary behaviour.
- **Connectivity and Network Context:** Wi-Fi networks, Bluetooth proximity events, and cellular network status were logged to characterise participants' social and spatial environments. For example, Bluetooth scans indicated nearby devices, supporting inferences about social interactions, while Wi-Fi connectivity provided coarse location cues within the campus and residential areas.

- **Environmental Sensors:** Light intensity, proximity, and ambient pressure readings captured the surrounding physical context. These sensors offered additional information about indoor/outdoor conditions, device placement, and environmental transitions, which could be related to activity types, commuting, or location-based behaviours.
- **App Usage and Digital Behaviour:** Foreground application usage, music playback, notifications, and other engagement patterns were recorded to characterise digital behaviour. These data provided insights into study habits, leisure activities, and social interactions conducted via digital platforms, and could be cross-referenced with EMAs and mood diaries to examine behavioural correlates.

A list of the sensors configured for this study is found in Appendix 13

Integration and Quality-Aware Features

Passive sensor streams in the iLog app provided a rich, multimodal view of participants' behaviours and environments, complementing self-reported data from the diaries. To ensure reliable and high-quality data collection, all data streams were continuously monitored in real time via the calendar dashboard.

The dashboard automatically flagged missing, delayed, or irregular data, enabling proactive interventions. Researchers actively reviewed these flags and, where necessary, prompted participants to re-enable sensors, adjust app permissions, or respond to outstanding question prompts. This real-time feedback loop helped maintain data completeness and minimise gaps caused by sensor dropouts, connectivity issues, or participant oversight.

By integrating passive sensing, self-reports, and active researcher monitoring, the system created a robust, resilient, and quality-aware data-collection ecosystem capable of sustaining participant engagement and producing reliable longitudinal datasets suitable for rigorous behavioural analyses.

6.3.3 Dashboard Design and Monitoring

The dashboard served as the central operational hub for overseeing data inflow, participant engagement, and sensor reliability throughout the study. Designed as a real-time monitoring tool, it provided both high-level summaries and participant-level details, enabling the research team to maintain data quality proactively.

The interface displayed key indicators, including:

- **Question response counts and completion rates**, which were updated continuously as new data arrived.
- **Sensor activity diagnostics**, such as which streams were active, their last upload timestamps, and fluctuations in data sampling rates.

- **Engagement signals**, including inactive or low-contributing participants and deviations from expected participation patterns.

Researchers were able to filter participants, view daily participation summaries, and detect anomalies (e.g., missing data or extended inactivity). This enabled timely communication with participants to resolve technical or motivational issues, improving both participation and data completeness. The dashboard also visualised key behavioural indicators, including daily screen usage, activity levels, estimated sleep duration (derived from prolonged device inactivity), and app interaction diversity, which helped contextualise variations in participation and detect when technical or behavioural factors affected data quality.

Real-time insights allowed the team to respond swiftly. Participants whose sensors had stopped, or whose permissions had been revoked, or whose diary questions remained incomplete were contacted by email. The closed-loop design, which linked data collection, monitoring, and researcher intervention, was essential for maintaining ongoing engagement and ensuring the longitudinal dataset's completeness and reliability.

6.3.4 Engagement and Incentives

To encourage participant engagement, we employed ongoing communication and flexible strategies aimed at mitigating dropout rates. The dashboard's monitoring features allowed the research team to pinpoint participants exhibiting decreased response rates or inconsistent sensor uploads. When these patterns were observed, researchers promptly intervened—frequently within hours—to promote re-engagement, address technical difficulties, or elucidate study requirements.

Incentives were given to promote sustained adherence while remaining sensitive to participants' workloads and constraints. Initially, rewards depended on completion of at least 85% of both sensor and self-report tasks. However, as network challenges and academic obligations intensified over the course of the experiment, this threshold was reduced to 50%. This adjustment maintained fairness across participants facing varying environmental constraints and ensured continued participation as external demands increased.

The combination of real-time monitoring, quick communication, and flexible incentives created a supportive environment. This successfully balanced the demands on participants with the goals of the research. This approach helped maintain a stable group of participants and ensured that most participants stayed throughout the six-week study.

6.3.5 Ethical and Privacy Considerations

The study adhered to strict ethical and privacy standards throughout data collection, processing, and storage. Ethical approval was obtained from both the Uganda National Council for Science and Technology (UNCST) and the Makerere University Institutional Review Board (MUIRB). We also made a formal data-sharing agreement between Makerere University (data controller) and the University of Trento (data processor), ensuring that all data processing activities complied with the Ugandan Data Protection and Privacy Act (2019) and the EU General Data Protection Regulation (GDPR).

Privacy protection measures were integrated into the system design. Data was encrypted while being sent and then kept on secure servers at the University of Trento, where access was tightly controlled. Only authorised members of the research team had access to the dataset, and all access was logged for compliance and transparency.

The study's goals, the types of data collected, possible risks, and the participants' rights, including the right to withdraw, were all explained to the participants through a detailed information sheet. A dedicated study webpage offered clear documentation on installation procedures, privacy protections, sensor permissions, and expected daily tasks. Informed consent was obtained before enrolment, ensuring that participation was entirely voluntary and based on a clear understanding.

6.3.6 Findings

The real-time monitoring dashboard proved instrumental in maintaining data quality. Researchers could detect dropouts, troubleshoot synchronisation issues, and intervene before data loss occurred.

The study produced a rich multimodal dataset combining passive sensing, self-reported diaries, context questions, and behavioural logs. Overall, participation was strong, with most participants contributing consistent data throughout the study period despite occasional technical and environmental challenges (e.g., limited internet access, permissions issues, or academic workload).

Participation and Response Patterns

During the experiment's longitudinal phase, participant engagement was continuously monitored in real time through the dashboard, which displayed daily response rates. A participant-level heatmap (refer to Figure 6.2), a key feature of the dashboard, was employed to track response frequency for each user over the study's duration. This visual representation enabled the research team to promptly identify both overall reductions in participation and specific individual patterns of disengagement.

The heatmap, with participants on the vertical axis and time on the horizontal axis, used colour intensity to show the number of responses. Lighter or blank segments indicated periods of low activity, which often aligned with overall drops in the total submission count. When such declines were detected, either in aggregate daily totals or across multiple individuals in the heatmap, personalised reminder emails were sent. These emails served as both a reminder and a motivational touchpoint, encouraging participants to resume or increase their engagement. As shown in the line graph (Figure 6.3), each email intervention was consistently followed by a noticeable recovery in daily submissions. For example, following a decline to approximately 2000 responses, submissions rebounded to levels near 2400 after the first round of reminders.

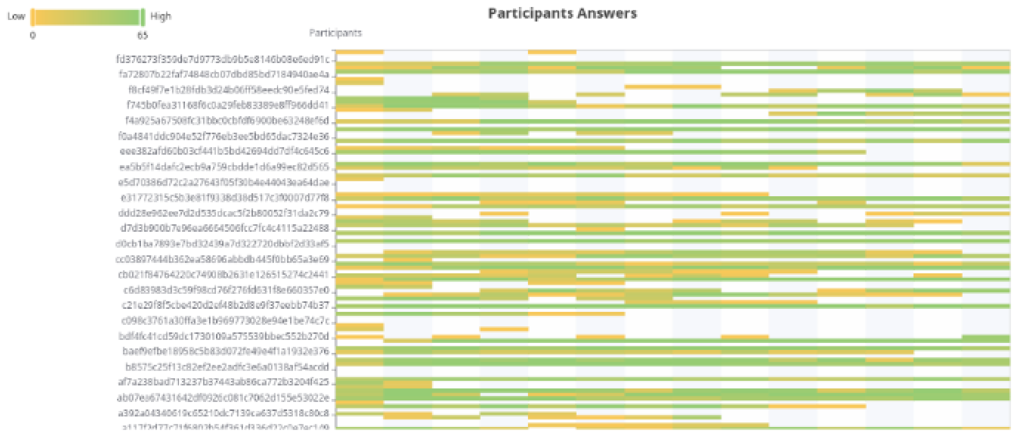


Figure 6.2: A heatmap of the rate of participant answering as shown in the dashboard

This integrated monitoring method, which combined macro-level trend analysis with micro-level participant tracking, worked well to keep people interested, keep dropout rates low, and make sure that data was collected consistently throughout the study period. This made it possible to follow up on specific cases and tell the difference between general trends and individual dropouts.

A minority of participants, however, showed more variable behaviour, including missed days or sudden drops in sensor data, often due to permission revocations, phone battery issues, or reduced motivation. These cases were quickly identified and addressed through researcher follow-up. For instance, in the design of the experiment, questions were scheduled to be sent only during the first two weeks of the study, and for the remaining 4 weeks, only sensor data collection was to continue. Many participants deactivated the iLog app once the question prompts stopped, a pattern which was readily detected in the monitoring dashboard and promptly communicated to participants to reactivate the app. This proactive monitoring ensured that data collection continued without interruption and helped maintain overall adherence to the study's guidelines.

Sensor Data Completeness and Quality

Sensor streams were generally stable across both Android and iOS devices. Motion sensors (steps, accelerometer) and device interaction events (screen unlocks) showed the highest continuity. At the same time, Bluetooth and Wi-Fi data exhibited greater variability due to hardware constraints and the university campus network coverage.

The real-time dashboard enabled researchers to detect anomalies such as:

- sudden gaps in sensor uploads,
- Extended inactivity periods,

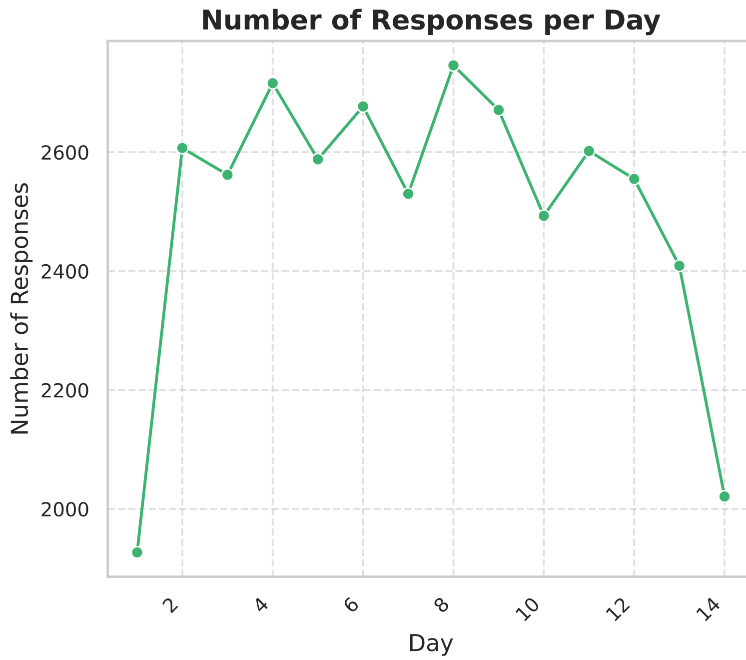


Figure 6.3: A graph showing the answering trend throughout the experiment

- abnormal sampling rates,
- and platform-specific sensor dropouts.

Prompt interventions that reminded participants to update permissions, disable battery optimisation, or reconnect to Wi-Fi significantly improved data recovery and continuity. Participants generally appreciated the research team's reminders and communication, which helped maintain engagement. Despite these challenges, no participants reported major privacy concerns after the study was explained, indicating that the consent and onboarding procedures were effective.

6.3.7 Discussion: Operational Lessons and Trade-offs

The MakOne deployment showed that the researcher dashboard helps sustain data quality and also revealed the trade-offs inherent in real-world mobile sensing studies.

The dashboard's main contribution was to monitor the data collection process in real-time. The heatmap (Figure 6.2) views turned engagement from an aggregate metric into an individually actionable signal. The consistent recovery in response rates following email interventions (Figure 6.3) demonstrates that timely, human-in-the-loop prompting is effective in mitigating attrition. This moves data collection management from a passive, post-hoc activity to an active, preventative one.

The study also elucidated clear platform-specific and context-dependent challenges. Sensor stability varied; motion and device interaction data were robust, while Bluetooth and Wi-Fi streams showed higher volatility. The common participant behaviour of deactivating the app after the question prompts stopped reveals a user mental model where "the study" was equated with active questioning, not passive sensing. This underscores the importance of clear, ongoing communication about the study's full duration and the value of passive data.

Key operational trade-offs were identified. First, the threshold adjustment for incentives, from 85% to 50% completion, was a necessary adaptation to environmental constraints (network issues and academic workloads). This preserved fairness and participation but explicitly traded ideal data completeness for real-world participant continuity. Second, there is a natural conflict between the level of detail in monitoring and the amount of work it creates for researchers. While reminders to participants concerning their contributions are valuable, they require dedication.

In conclusion, the dashboard did not eliminate the challenges of longitudinal sensing but made them visible, manageable, and actionable. It operationalised the principle of "quality-aware" data collection, proving that integrating real-time visualisation with researcher intervention is a viable strategy for improving the reliability and ethical sustainability of intensive longitudinal studies in authentic, resource-constrained settings.

6.4 Limitations of the Evaluation

While the case studies provide robust evidence for the dashboard's utility, certain limitations are acknowledged to contextualise the findings and guide future work.

- **Generalisability:** The participant evaluation (Case Study 1) relied on a synthetic dataset. Although it is with high fidelity and geographically grounded, interacting with one's own personal data may show different emotional responses, privacy concerns, and reflective depth compared to reviewing a simulated dataset. The MakOne study (Case Study 2) was conducted within a specific cultural and institutional context (Makerere University students). Participant tech skills, motivation structures, and daily routines may differ in other populations (e.g., older adults, clinical groups, different cultural settings), potentially affecting engagement patterns and dashboard utility.
- **Sample and Scope:** The Prolific sample in Case Study 1, while diverse, is a self-selected group of online research participants and may not represent the full spectrum of potential end-users. The six-week duration of the MakOne study, while substantial, captures only a segment of an academic semester; longer deployments might have revealed different engagement decay patterns or seasonal effects.
- **Evaluation Depth:** The primary evaluation metric for the researcher dashboard was operational effectiveness (response rates and intervention success). A deeper, qualitative assessment of the researcher experience, such as cognitive load, workflow integration, and

decision-making support, was not formally conducted and remains an area for future investigation.

6.5 Chapter Summary

This chapter presented a comprehensive evaluation of the dashboard system, addressing the distinct needs of its primary user roles. Case Study 1 demonstrated that the participant dashboard successfully supports self-reflection and data awareness, with high marks for usability, trust, and visual clarity. The key insight was the identified tension between the value of reflective depth and the challenge of information overload, offering clear directives for user-centric design iteration.

Case Study 2 validated the researcher dashboard as an operational tool for quality-aware longitudinal research. The real-time monitoring capabilities enabled proactive engagement management and timely interventions, which proved critical in maintaining data completeness and participant adherence in a real-world field study. The deployment showed important practical trade-offs and platform-specific challenges connected to mobile sensing.

The evaluation of these findings shows that the dashboard supports both participant reflection and researcher monitoring through shared data visualisation. While limitations in generalisability and scope are noted, the collective evidence strongly supports the utility, robustness, and design validity of the dual-role dashboard system for facilitating meaningful interaction with rich, multimodal behavioural data.

Chapter 7

Conclusion

7.1 Summary

This thesis addressed a core challenge in personal informatics: how to integrate and present multimodal personal data, dense, rich, and variably reliable, to support trustworthy reflection on daily behaviour. Despite advances in sensing and data collection that have made personal data increasingly abundant, existing systems often neglect the quality and uncertainty of the data they present. They also fail to align behavioural information with the temporal structures through which individuals organise their daily lives.

We introduced a calendar-centred framework that integrates big-thick personal data while explicitly accounting for data quality. The framework presents the digital calendar as a temporal scaffold that unifies planned activities, observed behaviour, contextual signals, and explicit indicators of data quality within a single coherent representation. By embedding data quality awareness directly into the collection, processing, and presentation of personal data, the proposed approach shifts reflective systems away from implicit assumptions of data reliability toward more transparent and interpretable insight generation.

A key contribution of this work is the formulation and application of practical data quality metrics for personal informatics systems, including question delivery success, response latency, answer duration, and sensor data completeness (see section 3.2.1). These metrics enable continuous assessment of data reliability and provide an explicit basis for interpreting behavioural patterns in light of their underlying quality. Rather than treating missing data or delay as incidental artefacts, the framework incorporates them as first-class analytical signals.

Building on this conceptual foundation, the thesis presented a scalable system architecture that supports the synchronisation and querying of heterogeneous data streams. The architecture demonstrates how quality can be preserved across data ingestion, processing, and visualisation layers, enabling consistent quality-aware interpretation throughout the system lifecycle. A calendar-dashboard hybrid interface was designed to operationalise this architecture, allowing users to explore behavioural patterns, contextual information, and data quality within a unified temporal view.

The proposed calendar dashboard system was evaluated through two empirical case studies: a controlled synthetic dataset and a real-world longitudinal deployment using the iLog application. The evaluation examined perceived behavioural understanding, insight reliability, and reflective awareness. The results indicate that integrating contextual data within a calendar-centred interface improves users' understanding of the presented insights and thus supports understanding of everyday behaviour. While the evaluation does not claim causal behaviour change, it provides evidence that quality-aware presentation enhances interpretability and trust, key prerequisites for meaningful personal behaviour reflection.

7.2 Discussion and Limitations

While the proposed system enhances transparency and reliability in the monitoring of personal data, it also exhibits some limitations, which consequently point to avenues for future investigation.

First and foremost, the system's data comes from a combination of sensor readings and self-reported information, which can lead to differences in how complete and accurate the data is. Users may omit entries, misreport events, or encounter sensor noise, all of which affect the reliability of insights. Although the system includes data quality indicators, they may not capture the full complexity of real-world uncertainty.

Moreover, the assessment was done in a controlled setting with a limited number of users. This could mean the results might not fully represent the variety of behaviours seen in real-world situations. Broader longitudinal studies are necessary to assess how users interact with the dashboard over time and whether it influences their data-entry habits or self-awareness.

A key methodological challenge is the potential for misleading correlations (*spurious correlations*), which can arise from the complexity and variety present in personal data. The relationships observed among diverse sensors, activities, or user self-reports might occasionally be due to chance rather than indicating a direct cause-and-effect relationship. Without careful statistical checks, users and researchers might misinterpret the results, which could lead to incorrect conclusions or ineffective actions. To reduce this risk, future research should use strong methods to confirm relationships, determine cause-and-effect, and measure uncertainty.

In summary, while the system successfully demonstrates the possibility and potential of quality-aware personal informatics, it also highlights the need for continued research in adaptive data quality modelling, multimodal integration, and sustained user engagement. To create personal data systems that are context-aware, self-reflective, reliable, and easy to understand, we must address challenges like misleading correlations, how we measure uncertainty, and the combination of different types of data.

7.3 Future Work

Building on the foundations laid in this thesis, several promising directions for future research and system development emerge.

Adaptive Quality Modelling One key area is the refinement of the data quality assessment mechanisms. Future work could explore adaptive models that learn from user behaviour and contextual cues to dynamically estimate reliability. For example, combining statistical uncertainty with behavioural consistency patterns could help detect when self-reported or sensed data become unreliable, prompting users with context-aware feedback or suggestions for correction.

Longitudinal and behavioural Evaluation Long-term evaluation, especially in behavioural studies, will be essential for understanding how users interact with the dashboard over time. Such studies could explore how understanding data quality affects user engagement, trust, and reflection. It could also investigate whether this understanding encourages more careful data entry and interpretation.

Mitigating Spurious Correlations

Due to the high-dimensional and varied nature of personal data, there's a risk that random patterns could be mistakenly seen as meaningful connections. Future research should focus on methods to find and reduce spurious correlations. This includes using strong statistical validation, causal inference techniques, and ways to measure uncertainty. Adding these safeguards would improve the reliability of the insights gained, ensuring that predictive or reflective feedback is based on genuinely informative relationships, rather than random patterns.

Integration of Machine Learning and Predictive Insights

The integration of machine learning and predictive insights presents a promising avenue for leveraging the growing availability of mobile and wearable data. Predictive models could be employed to furnish anticipatory insights, including the forecasting of routine disruptions and the identification of potential risks to well-being, all while upholding transparency concerning data reliability. This shift would transform the system from a reactive tool into a proactive personal assistant, built on reliable data.

Cross-Platform and Sensor Expansion

Future developments could broaden the integration scope to encompass supplementary sensors, such as those monitoring environmental conditions, physiological parameters, or smart-home functionalities, alongside various platforms like calendar systems and communication tools. This expansion would serve to enrich the dashboard's contextual understanding, thereby facilitating a more nuanced interpretation of individual routines and lifestyle trends.

Ethical and Privacy-Aware Design

Finally, continued exploration into privacy-preserving data management is critical. Research could focus on decentralised data storage, federated learning approaches, or user-controlled data quality auditing, ensuring that users retain autonomy over their personal information without compromising analytical depth.

7.4 Closing Remarks

Overall, this thesis demonstrates that integrating explicit data quality awareness into calendar-based personal informatics systems enhances the interpretability and practical utility of multimodal personal data. By aligning behavioural data with the temporal structures through which people organise their lives, and by making data reliability visible, the proposed calendar dashboard system contributes a novel methodological and system-level approach to reflective computing. This work advances the design of personal informatics systems toward more transparent, trustworthy, and contextually grounded support for understanding everyday behaviour.

Appendix A

Likert-scale survey

.1 Questionnaires for the Crowdsourcing Study

RQ1: Spatial Context Awareness

1. "It was clear which locations were associated with each activity/event." *(clarity)*
   [5-point Likert scale]

2. "The visualization of locations helped me understand the broader context of activities." *(usefulness for sensemaking)*
   [5-point Likert scale]

3. "It was easy to navigate the dashboard to find information about locations." *(ease of access/interaction)*
   [5-point Likert scale]

4. "In what ways did the visualizations help you (or fail to help you) interpret the locations of activities or events?"
   [Open-ended]

RQ2 & RQ3: Activity Context Awareness & Event Sequencing

5. "I could distinguish between different types of activities." *(discrimination of categories)*
   [5-point Likert scale]

6. "The dashboard helped me see when activities occurred simultaneously." *(parallel activities)*
   [5-point Likert scale]

7. "The dashboard made it clear how one activity led to or followed another." *(temporal sequencing)*
   [**5-point Likert scale**]

8. "Describe how you used the visualizations to understand the timing or sequence of activities."
   [**Open-ended**]

RQ4: Social Context Interpretation

9. "It was easy to see whether other people were present during activities." *(clarity of representation)*
   [**5-point Likert scale**]

10. "The visualizations helped me reflect on the role of others in the activities." *(support for reflection)*
    [**5-point Likert scale**]

11. "I could quickly identify differences between activities done alone and with others." *(contrast detection)*
    [**5-point Likert scale**]

RQ6: Interpersonal Reflection

12. "The social context visualization prompted me to think about relationships with others." *(relational reflection)*
    [**5-point Likert scale**]

13. "I noticed patterns in who was present across different activities." *(pattern detection)*
    [**5-point Likert scale**]

14. "The dashboard made me consider how social context influenced the activities shown." *(cause-effect awareness)*
    [**5-point Likert scale**]

15. "What reflections did you have about social interactions from the visualizations?"
    [**Open-ended**]

RQ5: Internal States Representation

16. "The dashboard made emotional states easy to understand." *(emotions clarity)*
    [**5-point Likert scale**]

17. "I could connect internal states (e.g., mood, energy) to specific activities." *(activity-state linkage)*
    [**5-point Likert scale**]

18. "The internal states visualization encouraged me to think about their significance." *(depth of reflection)*
    [**5-point Likert scale**]

19. "How did the visualizations influence your reflection on internal states?"
    [**Open-ended**]

RQ7: Material Context Comprehension

20. "The dashboard made it clear which tools or objects were used in each activity." *(clarity)*
    [**5-point Likert scale**]

21. "The material context helped me better understand how activities were carried out." *(sensemaking utility)*
    [**5-point Likert scale**]

22. "I could easily see connections between tools/objects and specific activities." *(association detection)*
    [**5-point Likert scale**]

23. "How did material context (tools/objects) affect your understanding of activities in the dashboard?"
    [**Open-ended**]

RQ8: Perceived Accuracy, Relevance, and Trust

24. "The data in the dashboard seemed accurate." *(accuracy perception)*
    [**5-point Likert scale**]

25. "The dashboard presented the most relevant aspects of everyday activities." *(relevance)*
    [**5-point Likert scale**]

26. "I trusted the dashboard to represent the data in a meaningful way." *(trust)*
    [**5-point Likert scale**]

27. "I felt confident using the dashboard to reflect on the data." *(confidence in use)*
    [**5-point Likert scale**]

28. "What aspects of the dashboard increased or decreased your trust in it?"
    [**Open-ended**]

RQ9: Adoption and Future Engagement

29. "I would like to use a dashboard like this to reflect on my everyday life." *(intention to adopt)*
    [**5-point Likert scale**]

30. “This dashboard would be valuable for long-term self-reflection.” *(perceived value over time)*
[**5-point Likert scale**]

31. “The dashboard would motivate me to regularly review my activities.” *(engagement motivation)*
[**5-point Likert scale**]

32. “I would recommend a dashboard like this to others.” *(social endorsement)*
[**5-point Likert scale**]

33. “In what ways could you imagine using such a dashboard for reflecting on your own activities or contexts?”
[**Open-ended**]

General Open-Ended Questions

1. How would you compare this dashboard to other tracking or reflection tools you may use (e.g., Apple Health, Google Fit)?

2. What did you find most enjoyable or helpful about using the dashboard?

3. What aspects of the interface, if any, did you find confusing or difficult to use?

4. Were there any features you expected to see but did not? What would you add or improve?

5. Describe any challenges you had in understanding how to access or interpret the data.

.2 Full Likert-Scale Survey Items and Responses

This appendix presents all Likert-scale items administered in the post-study survey, along with descriptive statistics. Items are grouped thematically for clarity. All items used a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Responses were collected from N = 130 participants.

Table 1: Responses to location and navigation items.

Item	Mean	SD	Median	Min–Max
It was clear which locations were associated with each activity/event.	4.23	0.92	4	1–5
The visualization of locations helped me understand the broader context of activities.	4.35	0.82	4	2–5
It was easy to navigate the dashboard to find information about locations.	4.30	0.84	5	1–5

.2.1 Location and Navigation

.2.2 Activity Timing and Sequencing

Table 2: Responses to timing and sequencing items.

Item	Mean	SD	Median	Min–Max
I could distinguish between different types of activities.	4.54	0.74	5	1–5
The dashboard helped me see when activities occurred simultaneously.	3.94	1.07	4	1–5
The dashboard made it clear how one activity led to or followed another.	3.77	1.19	4	1–5

.2.3 Social Context Reflection

Table 3: Responses to social context items.

Item	Mean	SD	Median	Min–Max
It was easy to see whether other people were present during activities.	4.23	1.01	5	1–5
The visualizations helped me reflect on the role of others in the activities.	3.78	1.19	4	1–5
I could quickly identify differences between activities done alone and with others.	4.12	1.07	5	1–5
The social context visualization prompted me to think about relationships with others.	3.68	1.25	4	1–5
I noticed patterns in who was present across different activities.	3.85	1.17	4	1–5
The dashboard made me consider how social context influenced the activities shown.	3.92	1.14	4	1–5

.2.4 Internal States Reflection

Table 4: Responses to internal states items.

Item	Mean	SD	Median	Min–Max
The dashboard made emotional states easy to understand.	4.08	1.10	4	1–5
I could connect internal states (e.g., mood, energy) to specific activities.	4.00	1.12	4	1–5
The internal states visualization encouraged me to think about their significance.	3.98	1.10	4	1–5

.2.5 Material Context Understanding

Table 5: Responses to material context items.

Item	Mean	SD	Median	Min–Max
The dashboard made it clear which tools or objects were used in each activity.	3.62	1.24	4	1–5
The material context helped me better understand how activities were carried out.	3.68	1.21	4	1–5
I could easily see connections between tools/objects and specific activities.	3.62	1.23	4	1–5

.2.6 Trust, Confidence, and Recommendation

Table 6: Responses to trust, confidence, and recommendation items.

Item	Mean	SD	Median	Min–Max
The dashboard presented the most relevant aspects of everyday activities.	4.15	0.99	4	1–5
I trusted the dashboard to represent the data in a meaningful way.	4.03	1.01	4	1–5
I felt confident using the dashboard to reflect on the data.	4.00	1.06	4	1–5
I would like to use a dashboard like this to reflect on my everyday life.	3.31	1.39	3	1–5
This dashboard would be valuable for long-term self-reflection.	3.69	1.25	4	1–5
The dashboard would motivate me to regularly review my activities.	3.48	1.31	4	1–5
I would recommend a dashboard like this to others.	3.46	1.36	4	1–5

.3 Time diaries

.4 Sensors Details

Table 7: Time diaries collecting contextual information about the participant every half hour.

A3. What are you doing?	**A4. Where are you?**
1. Sleeping	1. Home apartment, room
2. Personal care	2. Home garden, patio, courtyard
3. Eating (go to A3c, Table 11)	3. Relatives Home
4. Cooking, Food preparation & management	4. House (friends others)
5. Study/work group	5. Classroom, Laboratory
6. Lecture, seminar, conference, university meeting	6. Classroom, Study hall
7. Did not do anything special	7. University Library
8. Rest/nap	8. Other university places
9. Break	9. Canteen
10. Walking	10. Other Library
11. Travelling (go to A3a1, A3a2, Table 11)	11. Gym, swimming pool, Sports centre...
12. Social life	12. Grocery Shop
13. Happy Hour, Drinking, Party	13. Supermarket...
14. Phone/Video calling	14. Street markets
15. In chat on Internet or reading, sending e-mail	15. Shops, shopping centers, indoor markets, other shops
16. Surfed or seeking, reading information via Internet	16. Café, pub, bar
17. Social media (Facebook, Instagram, etc.)	17. Restaurant, pizzeria, Street food vendor
18. Watching TV, video, YouTube, etc.	18. Movie Theatre Museum...
19. Listening to music	19. In the street
20. Reading a book, periodicals, news, etc.	20. Public Park/Garden
21. Movie Theatre Concert...	21. Countryside, mountain, hill, beach
22. Entertainment Exhibit, and Culture	22. Workplace, office
23. Others Entertainment and Culture	23. Weekend home or holiday apartment
24. Arts	24. Hotel, guesthouse, camping site
25. Hobbies	25. Another indoor place
26. Games	26. Another outdoor place
27. Freetime study	
28. Sport (go to A3b, Table 11)	
29. Voluntary work and participatory activities	
30. Household and family care	
31. Grocery Shopping	
32. Other Shopping	
33. Work	
34. Other	

A5. With whom are you?	**A6a. What is your mood?**
1. Alone	1. 😀
2. Friend(s)	2. 🙂
3. Relative(s)	3. 😑
4. Classmate(s)	4. 😕
5. Roommate(s)	5. 🙁
6. Colleague(s)	
7. Partner	
8. Other	

Table 8: Morning questions sent at 8:00 AM.

A1. How would you rate your sleep quality last night?	**A2. How do you expect your day to be?**
1. 😀 very good	1. 😀
2. 🙂 fairly good	2. 🙂
3. 😑	3. 😑
4. 😕 fairly bad	4. 😕
5. 🙁 very bad	5. 🙁

Table 9: Evening questions sent at 10:00 PM.

A7. How was your day?	**A8. Did you have any problem at [college (weekdays)] today?**	**A9. What was the problem you had?**
1. 😀	1. Yes	open-ended question
2. 🙂	2. No	
3. 😑		
4. 😕		
5. 🙁		

A10.Were you able to solve the problem (alone or with help of someone)?	**A11. Is there anything that you would have liked to do today that was not possible because of the Covid-19 virus?**
open-ended question	open-ended question

Table 10: List of motivations to suspend question notifications for a fixed number of hours.

Break options
36. Others
37. I will participate in sports activities
38. I have a work/study meeting
39. I am at the cinema/theatre/hospital/church
40. I am starting classes/lessons/lab
41. I will go to sleep

Table 11: In-depth questions that appear when certain options are selected in the question "What are you doing?"

A3a1. And you travel to/from or related to:	**A3a2. How are you moving?**	**A3b. What kind of sports activity?**	**A3c. Select the main food & drink you ate** (Multiple choices)
1. study 2. social life 3. shopping and services 4. other leisure 5. work 6. changing locality 7. other or unspecified travel purpose	1. on foot 2. by bike 3. by bus/tram 4. by metro/subway/underground 5. by train 6. by e-scooter 7. by car 8. by car as passenger 9. by car sharing 10. by moped, motorbike 11. by moped, motorbike as passenger 12. by motorboat 13. by airplane 14. by taxi/Uber 15. other private transport modes 16. other public transport modes	1. Walking, Trekking, and hiking 2. Jogging and running 3. Cycling, skiing, and skating 4. Ball games 5. Gymnastics and Fitness 6. Water sports 7. Other or unspecified sports or indoor activities 8. Other or unspecified sports or outdoor activities 9. Productive exercise (e.g., hunting, fishing, picking berries, mushrooms, or herbs)	1. Bread, steamed buns and/or breakfast cereals 2. Rice, potatoes, beans, pasta, noodles, dumplings, etc. 3. Vegetables 4. Fruits 5. Meat 6. Fish 7. Processed meat (ham, bacon, sausages) 8. Dairy products (Plain or low-fat milk, yogurt, cheese) 9. Soya-based food (milk, yogurt, tofu) 10. Pastries and sweets 11. Snack/sandwiches (chips...) 12. Water 13. Soda 14. Coffee/tea or similar 15. Others non-alcoholic drink 16. Beer 17. Wine 18. Liquor 19. Other alcoholic drink 20. Other food

Table 12: Additional questions related to food and drinks.

A6b. In the last two hours did you have any snacks or drinks (except breakfast, lunch, and dinner). (administered at hours 02, 04, 06, 10, 12, 15, 17, 19, 22, 24) (Multiple choices)	**A6c. Select the food & drink taken as snack. If you had more than one snack in the last two hours, only focus on the most recent one.** (Multiple choices)
1. No 2. Yes, between now and 30 minutes ago **(go to A6c)** 3. Yes, between 0.5 and 1 hour ago **(go to A6c)** 4. Yes, between 1 and 1.5 hours ago **(go to A6c)** 5. Yes, between 1.5 and 2 hours ago **(go to A6c)**	1. Confectionery (Candy, Chocolate, etc) 2. Cookies, cakes, and pastries 3. Bars (Energy bar, etc.) 4. Crackers/biscuits 5. Seeds, nuts, grains, legumes 6. Savory snacks (Chips, Tapas, Pizza, Nachos, Snack mix, deep frying) 7. Sandwiches (Sandwich, Hamburgers, Hot dogs, Bagel) 8. Frozen (Ice cream, Milkshake, etc.) 9. Bread, steamed buns and/or breakfast cereals 10. Rice, potatoes, beans, pasta, noodles, dumplings, etc. 11. Vegetables 12. Fruits 13. Dairy products (milk, yogurt, cheese) 14. Soya-based food (milk, yogurt, tofu) 15. Meat 16. Fish 17. Processed meat (ham, bacon, sausages) 18. Water 19. Soda 20. Coffee/tea or similar 21. Others non-alcoholic drink 22. Beer 23. Wine 24. Spirit 25. Others alcoholic drink 26. Other food

Table 13: List of sensors. The *Type* column reports HW for hardware-based sensors and SW for software-based sensors.

No.	Type	Name	Sampling Frequency
1	SW	Bluetooth devices (classic)	Once every minute
2	SW	Bluetooth devices (BLE)	Once every minute
3	SW	Cellular network information	Once every minute
4	SW	Connected Wi-Fi network	On change
5	SW	Available Wi-Fi networks	Once every minute
6	HW	Accelerometer	Up to 10 samples/s
7	HW	Accelerometer (uncalibrated)	Up to 10 samples/s
8	HW	Gyroscope	Up to 10 samples/s
9	HW	Gyroscope (uncalibrated)	Up to 10 samples/s
10	HW	Linear acceleration	Up to 10 samples/s
11	HW	Gravity	Up to 10 samples/s
12	HW	Rotation vector	Up to 10 samples/s
13	HW	Geomagnetic rotation vector	Up to 10 samples/s
14	HW	Magnetic field	Up to 10 samples/s
15	HW	Magnetic field (uncalibrated)	Up to 10 samples/s
16	HW	Orientation	Up to 10 samples/s
17	HW	Light	Up to 10 samples/s
18	HW	Proximity	Up to 10 samples/s
19	HW	Ambient temperature	Up to 10 samples/s
20	HW	Pressure (barometer)	Up to 10 samples/s
21	HW	Relative humidity	Up to 10 samples/s
22	HW	Location (GPS/network-based)	Once every minute
23	SW	Movement activity label	Once every 30 seconds
24	HW	Step counter	Up to 10 samples/s
25	HW	Step detector	On change
26	SW	Headset status [ON/OFF]	On change
27	SW	Music playback state	On change
28	SW	Notification listener	On change
29	SW	Running applications	Once every 5 seconds
30	SW	Airplane mode [ON/OFF]	On change
31	SW	Battery charging state [ON/OFF]	On change
32	SW	Battery level	On change
33	SW	Doze mode [ON/OFF]	On change
34	SW	Ring mode [Silent/Normal]	On change
35	SW	Touch events	On change
36	SW	Screen status [ON/OFF]	On change
37	SW	User presence	On change

Bibliography

[1] Gregory D Abowd et al. "Towards a better understanding of context and context-awareness". In: *International symposium on handheld and ubiquitous computing*. Springer, 1999, pp. 304–307.

[2] Bon Adriel Aseniero et al. "Activity river: Visualizing planned and logged personal activities for reflection". In: *Proceedings of the 2020 International Conference on Advanced Visual Interfaces*. 2020, pp. 1–9.

[3] Jorge Luis Victoria Barbosa. "Ubiquitous computing: Applications and research opportunities". In: *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE. 2015, pp. 1–8.

[4] Jakob E Bardram. "The personal awareness assistant: Designing for human awareness of context". In: *International Conference on Pervasive Computing*. Springer. 2005, pp. 276–291.

[5] Carlo Batini and Monica Scannapieco. *Data and Information Quality: Dimensions, Principles and Techniques*. Springer, 2009.

[6] Frank Bentley et al. "Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 20.5 (2013), pp. 1–27.

[7] Niels van Berkel, Denzil Ferreira, and Vassilis Kostakos. "Context-Informed Scheduling and Analysis: Improving Accuracy of Mobile Self-Reports". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 2019, pp. 1–12. DOI: 10.1145/3290605.3300280.

[8] Ivano Bison and Haonan Zhao. "Factors impacting the quality of user answers on smartphones". In: *Proceedings of the Second International Conference on Hybrid Human-Machine Intelligence*. Vol. 3456. CEUR Workshop Proceedings. 2023, pp. 208–213.

[9] Ivano Bison, Haonan Zhao, and Fausto Giunchiglia. "What impacts the quality of the user answers when asked about the current context?" In: *arXiv preprint* arXiv:2405.04054 (2024). URL: https://arxiv.org/abs/2405.04054.

[10] Tom Boellstorff and Bill Maurer. "Making big data, in theory". In: *First Monday* (2013).

[11] Andrea Bontempelli et al. "Lifelong personal context recognition". In: *arXiv preprint arXiv:2205.10123* (2022).

[12] Tobias Bornakke and Brian Lystgaard Due. "Big–thick blending: A method for mixing analytical insights from big and thick data sources". In: *Big Data & Society* 5.1 (2018), p. 2053951718765026.

[13] Danah Boyd and Kate Crawford. "Critical questions for big data". In: *Information, Communication & Society* 15.5 (2012), pp. 662–679.

[14] Peter J Brown, John D Bovey, and Xian Chen. "Context-aware applications: from the laboratory to the marketplace". In: *IEEE personal communications* 4.5 (1997), pp. 58–64.

[15] Eun Kyoung Cho et al. "Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps". In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. ACM. 2022, pp. 1–14. DOI: 10.1145/3491102.3517622.

[16] Eun Kyoung Choe et al. "Understanding quantified-selfers' practices in collecting and exploring personal data". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2014, pp. 1143–1152.

[17] Joan Dallinga et al. "Analysis of the features important for the effectiveness of physical activity–related apps for recreational sports: Expert panel approach". In: *JMIR mHealth and uHealth* 6.6 (2018), e9459.

[18] Tushar Kant Das and P Mohan Kumar. "Big data analytics: A framework for unstructured data analysis". In: *International Journal of Engineering Science & Technology* 5.1 (2013), p. 153.

[19] *Daylio*. https://daylio.net. Accessed: 2025-09-26. 2022.

[20] Anind K Dey, Gregory D Abowd, and Andrew Wood. "CyberDesk: A framework for providing self-integrating context-aware services". In: *Knowledge-based systems* 11.1 (1998), pp. 3–13.

[21] Chris Elsden, David Kirk, and Abigail Durrant. "Know Thyself: A Theory of the Self for Personal Informatics". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, pp. 3808–3820. DOI: 10.1145/2858036.2858498.

[22] Daniel A Epstein et al. "A lived informatics model of personal informatics". In: *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 2015, pp. 731–742.

[23] *Exist.io*. https://exist.io. Accessed: 2025-09-26. 2023.

[24] Denzil Ferreira, Vassilis Kostakos, and Anind K Dey. "AWARE: mobile context instrumentation framework". In: *Frontiers in ICT* 2 (2015), p. 6.

[25] Adrian Furnham. "Response bias, social desirability and dissimulation". In: *Personality and Individual Differences* 7.3 (1986), pp. 385–400. DOI: 10.1016/0191-8869(86)90014-0.

[26] Clifford Geertz. "Thick description: Toward an interpretive theory of culture". In: *The cultural geography reader*. Routledge, 2008, pp. 41–51.

[27] Fausto Giunchiglia. "Contextual reasoning". In: *Epistemologia, special issue on I Linguaggi e le Macchine* 16 (1993), pp. 345–364.

[28] Fausto Giunchiglia, Enrico Bignotti, and Mattia Zeni. "Personal context modelling and annotation". In: *2017 IEEE international conference on pervasive computing and communications workshops (PerCom workshops)*. IEEE. 2017, pp. 117–122.

[29] Fausto Giunchiglia et al. “A Context Model for Personal Data Streams”. In: *Web and Big Data: 6th International Joint Conference, APWeb-WAIM 2022, Nanjing, China, November 25–27, 2022, Proceedings, Part I*. Springer. 2023, pp. 37–44.

[30] Marie-Catherine Harter et al. “Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement”. In: *Organizational Research Methods* 22.4 (2019), pp. 969–1006. DOI: 10.1177/1094428118802626.

[31] Alastair van Heerden et al. “Integration of passive sensing technology to enhance delivery of psychological interventions for mothers with depression: the StandStrong study”. In: *Scientific Reports* 14.1 (2024), p. 13535.

[32] Alexis Hiniker, Sarita Y Schoenebeck, and Julie A Kientz. “Not at the dinner table: Parents' and children's perspectives on family technology rules”. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 2016, pp. 1376–1389.

[33] Fadime Kaya, Nihal Bostanci Daştan, and Erkan Durar. “Smart phone usage, sleep quality and depression in university students”. In: *International Journal of Social Psychiatry* 67.5 (2021), pp. 407–414.

[34] Ivan Kayongo et al. “A methodology and a platform for high-quality rich personal data collection”. In: *The European Journal on Artificial Intelligence* (2025). Advanced online publication. DOI: 10.1177/30504554251333615. URL: https://arxiv.org/abs/2501.16864.

[35] Ivan Kayongo et al. “A Methodology and System For Big-Thick Data Collection”. In: *Jahrestagung der Gesellschaft für Informatik*. 2024.

[36] Michael Kizito et al. “MakOne: Behavioural Data of University Students' Smart Devices in Uganda”. In: *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2025, pp. 815–819.

[37] Dilber Korkmaz et al. “Ambulatory physiological measures obtained under naturalistic urban mobility conditions have acceptable reliability”. In: *Scientific Reports* 15.1 (2025), p. 28940.

[38] Dheeraj Kumar et al. “Fast and scalable big data trajectory clustering for understanding urban mobility”. In: *IEEE Transactions on Intelligent Transportation Systems* 19.11 (2018), pp. 3709–3722.

[39] Nicholas D Lane et al. “A survey of mobile phone sensing”. In: *IEEE Communications Magazine* 48.9 (2010), pp. 140–150.

[40] Ian Li, Anind Dey, and Jodi Forlizzi. “A Stage-Based Model of Personal Informatics Systems”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010.

[41] Ian Li, Anind K. Dey, and Jodi Forlizzi. “Understanding My Data, Myself: Supporting Self-Reflection with Ubicomp Technologies”. In: *Proceedings of the International Conference on Ubiquitous Computing (UbiComp)*. ACM, 2011, pp. 405–414.

[42] Xiaoyue Li et al. “Representing habits as streams of situational contexts”. In: *International Conference on Advanced Information Systems Engineering*. Springer. 2022, pp. 86–92.

[43] Deborah Lupton. "'I just want it to be done, done, done!'food tracking apps, affects, and agential capacities". In: *Multimodal Technologies and Interaction* 2.2 (2018), p. 29.

[44] Wendy E. Mackay. "Diversity in the Use of Electronic and Paper Calendars". In: *Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work*. ACM, 1988.

[45] John McCarthy. "Generality in artificial intelligence". In: *Communications of the ACM* 30.12 (1987), pp. 1030–1035.

[46] Sarah Mennicken, David Kim, and Elaine May Huang. "Integrating the smart home into the digital calendar". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016, pp. 5958–5969.

[47] Rajesh Kumar Mishra and Rekha Agarwal. "Impact of digital evolution on various facets of computer science and information technology". In: *Digital Evolution: Advances in Computer Science and Information Technology* 17 (2024).

[48] Jukka-Pekka Onnela et al. "Beiwe: A data collection platform for high-throughput digital phenotyping". In: *Journal of Open Source Software* 6.68 (2021), p. 3417. DOI: 10.21105/joss.03417. URL: https://doi.org/10.21105/joss.03417.

[49] François Osiurak, Jordan Navarro, and Emanuelle Reynaud. "How our cognition shapes and is shaped by technology: A common framework for understanding human tool-use interactions in the past, present, and future". In: *Frontiers in psychology* 9 (2018), p. 293.

[50] Stephen Payne and Adeline Maximilien. "Personal Ecologies of Calendar Artifacts". In: *Journal of Interaction Science* 3.1 (2015), p. 6.

[51] Miquel Porta. *A dictionary of epidemiology*. Oxford university press, 2014.

[52] James O. Prochaska and Wayne F. Velicer. "The Transtheoretical Model of Health Behavior Change". In: *American Journal of Health Promotion* 12.1 (1997), pp. 38–48. DOI: 10.4278/0890-1171-12.1.38.

[53] Yatharth Ranjan et al. "RADAR-base: open source mobile health platform for collecting, monitoring, and analyzing data using sensors, wearables, and mobile devices". In: *JMIR mHealth and uHealth* 7.8 (2019), e11734.

[54] Bill N Schilit and Marvin M Theimer. "Disseminating active map information to mobile hosts". In: *IEEE network* 8.5 (1994), pp. 22–32.

[55] Thanos G Stavropoulos, Georgios Meditskos, and Ioannis Kompatsiaris. "DemaWare2: Integrating sensors, multimedia and semantic analysis for the ambient care of dementia". In: *Pervasive and Mobile Computing* 34 (2017), pp. 126–145.

[56] Tahia Tazin et al. "Understanding mental health using ubiquitous sensors and machine learning: Challenges ahead". In: *Human Activity and Behavior Analysis* (2024), pp. 222–257.

[57] Manas Tungare, Manuel A Pérez-Quinones, and Brandon Sams. "An Exploratory Study of Calendar Use". In: *Proceedings of the 2008 CHI Conference*. ACM, 2008.

[58] Mattia Zeni, Ilya Zaihrayeu, and Fausto Giunchiglia. "Multi-device activity logging". In: *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: adjunct publication*. 2014, pp. 299–302.

List of Figures

List of Tables