



# Introducing openESM: A database of openly available experience sampling datasets

---

**Björn Siepe<sup>1</sup>**

October 18th, 2025

Talk at the Emotional Disorder Genesis and Evolution (EDGE) Lab Meeting

<sup>1</sup>Psychological Methods Lab, Department of Psychology, Philipps-Universität Marburg

# ESM Research: Scattered Landscape

# ESM Research: Scattered Landscape

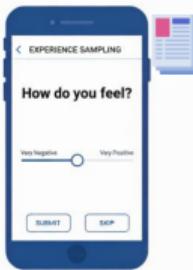


**N = 150**  
**Affect**

# ESM Research: Scattered Landscape



**N = 150**  
**Affect**



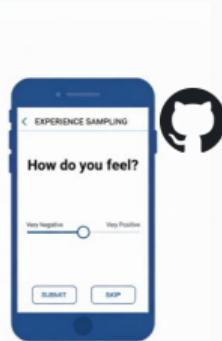
**N = 104**  
**Life Events**

Logos from (Freepik, 2025a,b; Google, 2025; GitHub, 2018)

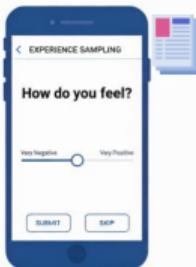
# ESM Research: Scattered Landscape



**N = 150**  
**Affect**



**N = 79**  
**Personality**



**N = 104**  
**Life Events**

Logos from (Freepik, 2025a,b; Google, 2025; GitHub, 2018)

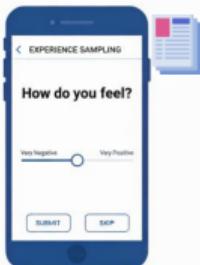
# ESM Research: Scattered Landscape



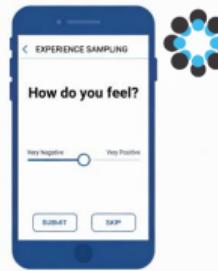
**N = 150**  
**Affect**



**N = 79**  
**Personality**



**N = 104**  
**Life Events**



**N = 243**  
**Depression**



Logos from (Freepik, 2025a,b; Google, 2025; GitHub, 2018)

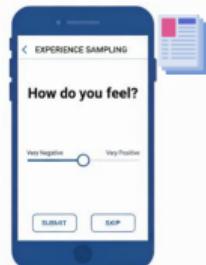
# ESM Research: Scattered Landscape



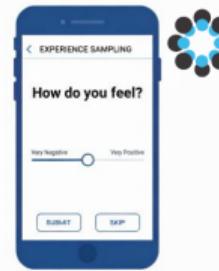
**N = 150**  
**Affect**



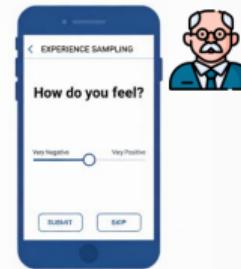
**N = 79**  
**Personality**



**N = 104**  
**Life Events**



**N = 243**  
**Depression**



**N = 150**  
**Work Stress**



Logos from (Freepik, 2025a,b; Google, 2025; GitHub, 2018)

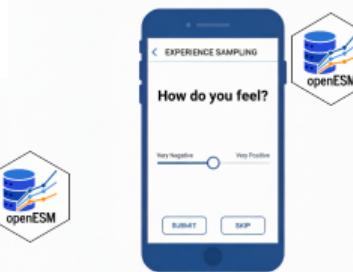
# The Solution: openESM



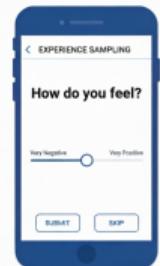
N = 150  
Affect



N = 79  
Personality



N = 243  
Depression



N = 104  
Life Events



N = 150  
Work Stress

# Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

---

<sup>1</sup>R logo: <https://www.r-project.org/logo/>

<sup>2</sup>Python logo: <https://www.python.org/community/logos/>

# Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

 <sup>1</sup> and  <sup>2</sup> packages for easy access

---

<sup>1</sup> R logo: <https://www.r-project.org/logo/>

<sup>2</sup> Python logo: <https://www.python.org/community/logos/>

# Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

 <sup>1</sup> and  <sup>2</sup> packages for easy access



## Substantive Research

Test theories across contexts

---

<sup>1</sup> R logo: <https://www.r-project.org/logo/>

<sup>2</sup> Python logo: <https://www.python.org/community/logos/>

# Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations



<sup>1</sup>



<sup>2</sup>

packages for easy access



## Substantive Research

Test theories across contexts

## Design Research

Optimize measurement  
choices

---

<sup>1</sup>R logo: <https://www.r-project.org/logo/>

<sup>2</sup>Python logo: <https://www.python.org/community/logos/>

# Introducing openESM

60 harmonized datasets • 16k+ participants • 740k+ observations

<sup>1</sup><sup>2</sup>

packages for easy access



## Substantive Research

Test theories across contexts

## Design Research

Optimize measurement  
choices

## Statistical Methods

Benchmark and develop  
methods

---

<sup>1</sup>R logo: <https://www.r-project.org/logo/>

<sup>2</sup>Python logo: <https://www.python.org/community/logos/>

# Live Demo



Home About Datasets Search Documentation

## A database of open experience sampling datasets

openESM is a database of harmonized openly available experience sampling datasets. The platform enables easy reuse of datasets with consistent and detailed metadata standards to advance reproducible research into daily life.

[Start Exploring Datasets](#)



**60**

**Datasets**

Explore our growing collection of datasets.



**16.000+**

**Individuals**

Obtain insights from thousands of participants.



**100+**

**Constructs**

Investigate a diverse range of psychological variables.

# Design Process

# Design Process

## 1. Cleaning

Harmonized NAs,  
removed redundant  
columns

*E.g., standardized ID,  
beep, day*

# Design Process

## 1. Cleaning

Harmonized NAs,  
removed redundant  
columns

*E.g., standardized ID,  
beep, day*

## 2. Harmonization

Same variable names  
for same questions

*E.g., “depressed”  
across datasets*

# Design Process

## 1. Cleaning

Harmonized NAs,  
removed redundant  
columns

*E.g., standardized ID,  
beep, day*

## 2. Harmonization

Same variable names  
for same questions

*E.g., “depressed”  
across datasets*

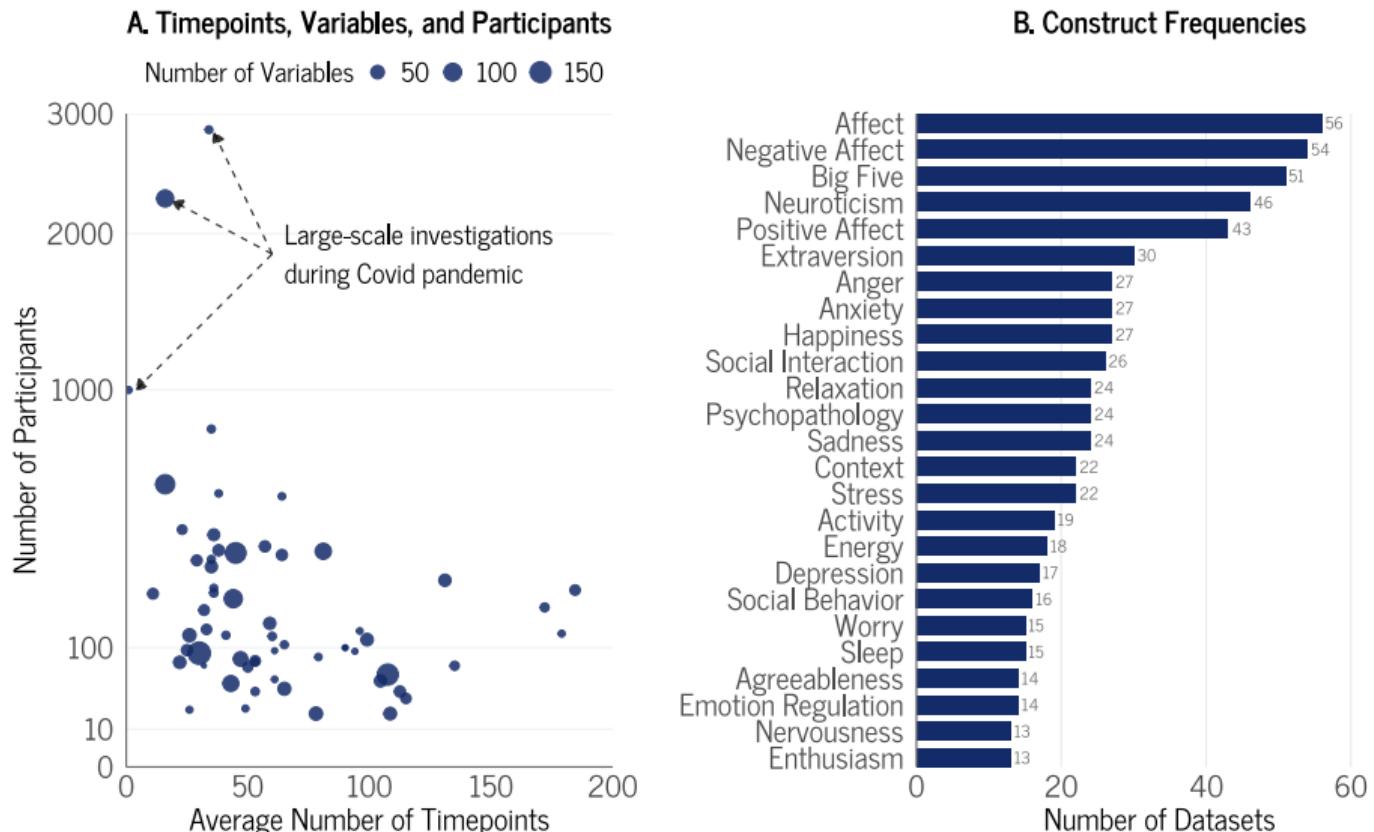
## 3. Annotation

Hierarchical construct  
tags

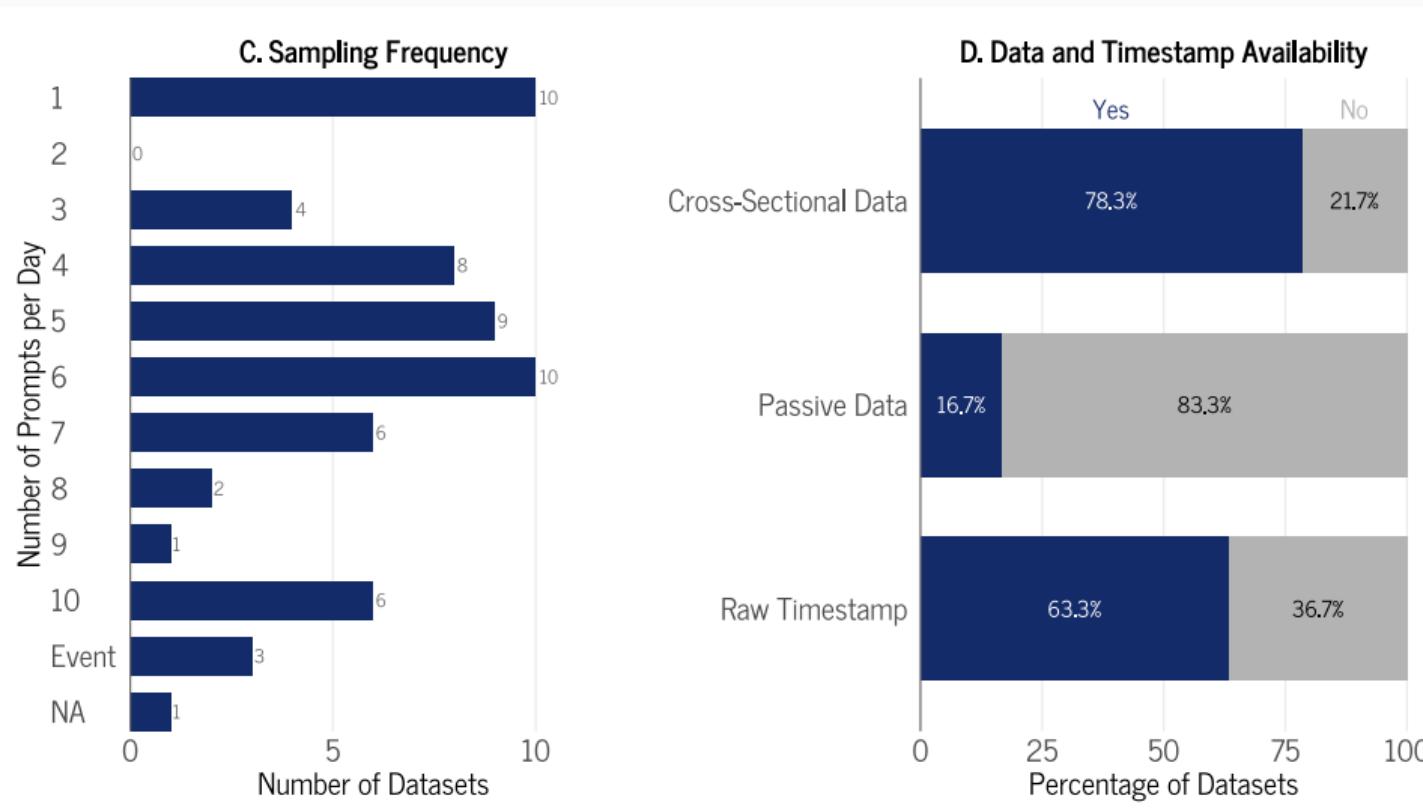
*E.g., “angry” → anger,  
negative affect, affect*

# What's inside

# What's inside

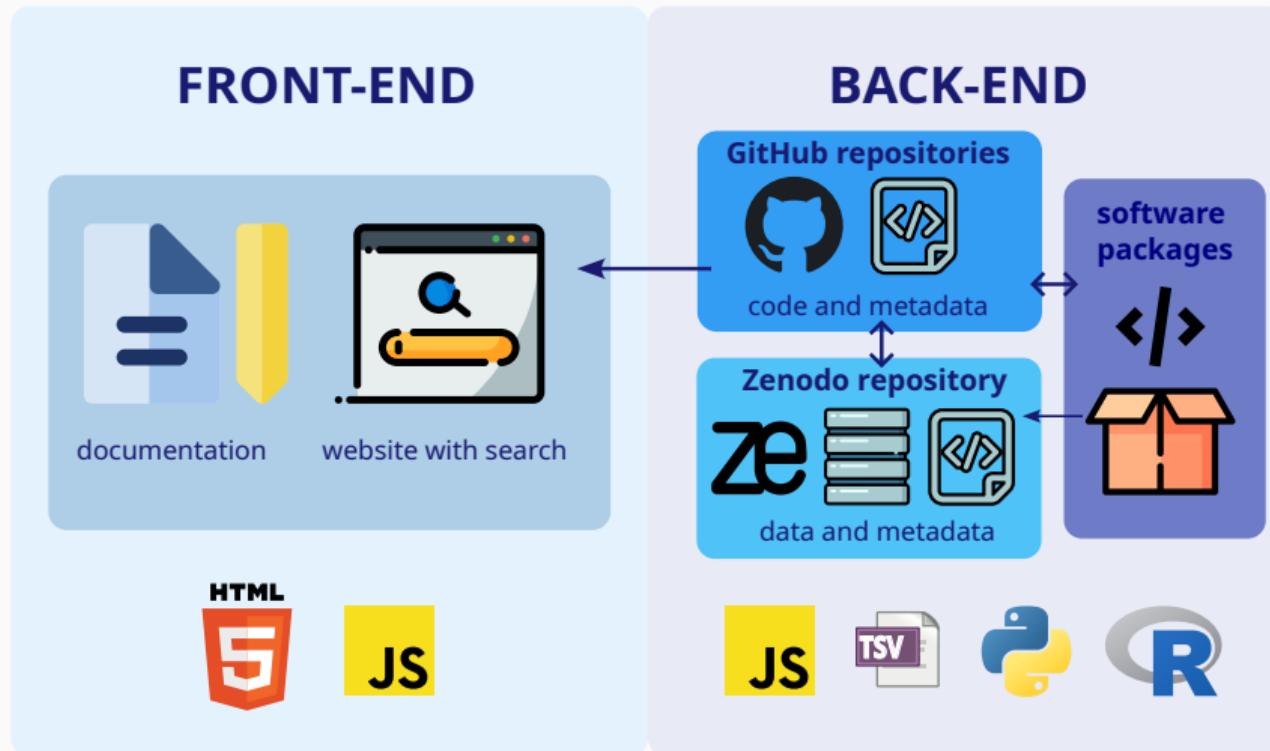


# What's inside



# Architecture Behind It

# Architecture Behind It



## Example Analysis

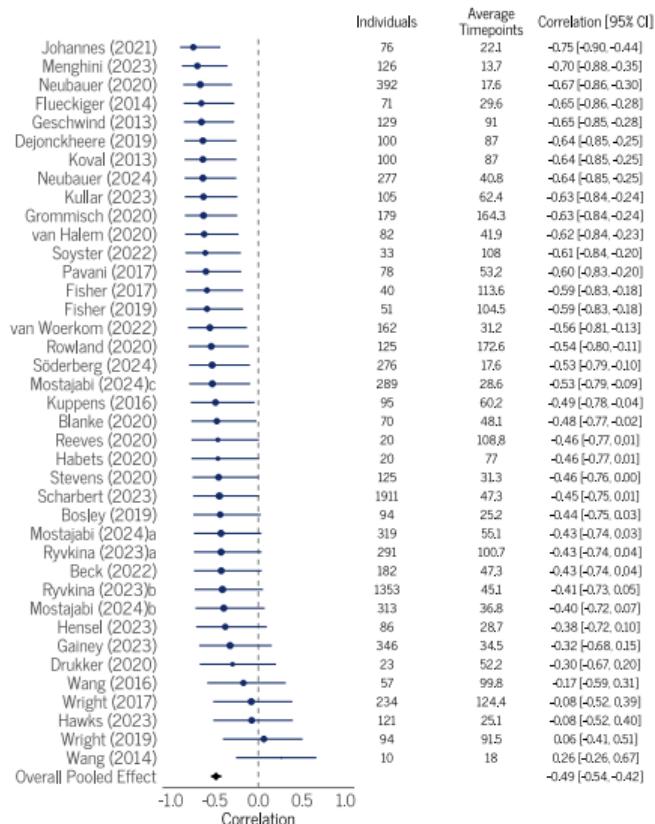
- **Question:** Within-person correlation of positive and negative affect

## Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations

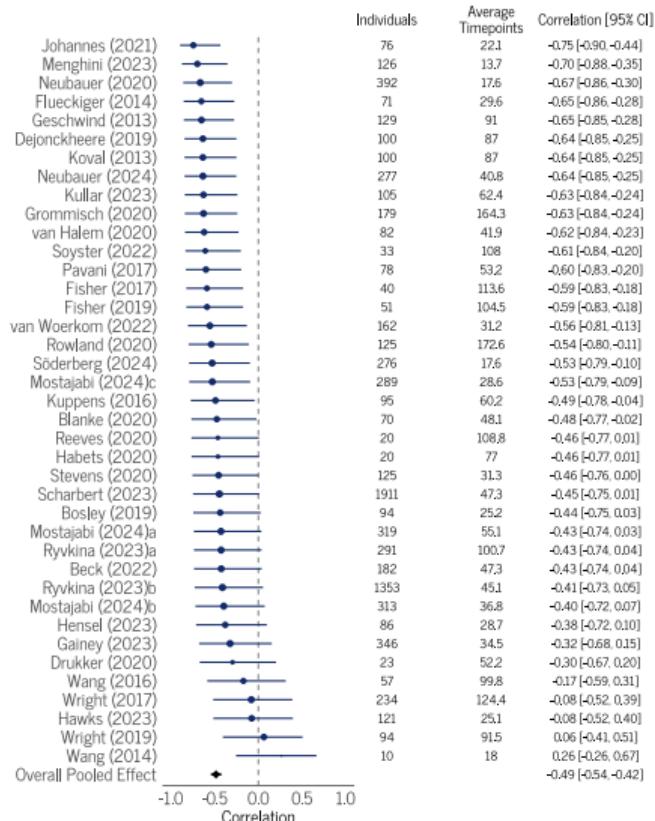
# Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled  $r = -0.49$  [95% CI:  $-0.54, -0.42$ ]



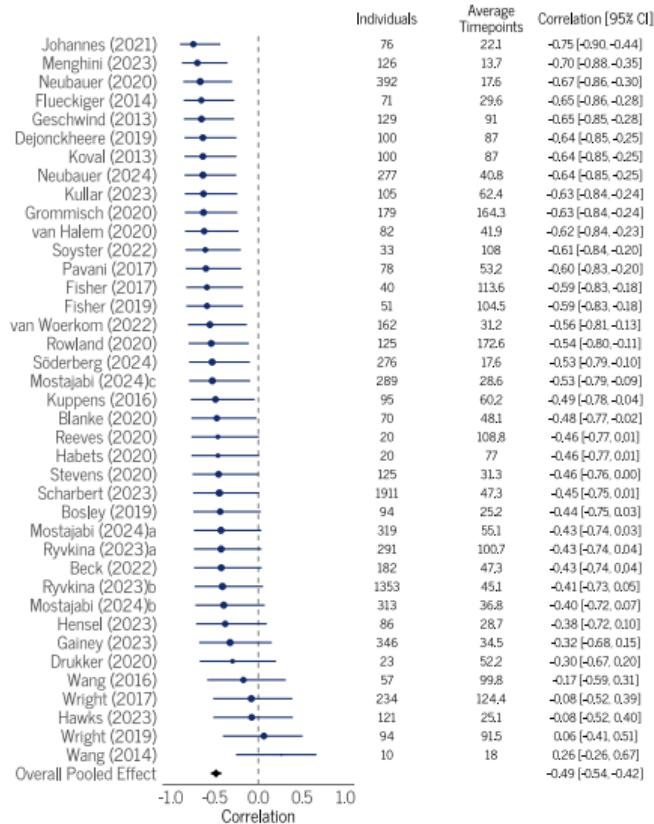
# Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled  $r = -0.49$  [95% CI:  $-0.54, -0.42$ ]
- **Design insight:** More prompts/day → stronger negative correlation



# Example Analysis

- **Question:** Within-person correlation of positive and negative affect
- **Sample:** 39 datasets, 8,456 individuals, 529K observations
- **Result:** Pooled  $r = -0.49$  [95% CI:  $-0.54, -0.42$ ]
- **Design insight:** More prompts/day → stronger negative correlation
- **Robustness:** Stable across 12 preprocessing/modeling choices



# Next Steps

# Next Steps



## Community

Involve researchers, expand  
non-WEIRD data

# Next Steps



## Community

Involve researchers, expand  
non-WEIRD data



## Data & Metadata

Extend coverage and  
documentation

# Next Steps



## Community

Involve researchers, expand  
non-WEIRD data



## Data & Metadata

Extend coverage and  
documentation



## Software

Add benchmarking and  
filtering tools

# Next Steps



## Community

Involve researchers, expand  
non-WEIRD data



## Data & Metadata

Extend coverage and  
documentation



## Software

Add benchmarking and  
filtering tools

**Goal:** Turn openESM into a continuously evolving community resource for  
cumulative ESM research

# Thanks to the team



Jonas Haslbeck



Matthias Kloft



Anabel Büchner



Yong Zhang



Eiko Fried

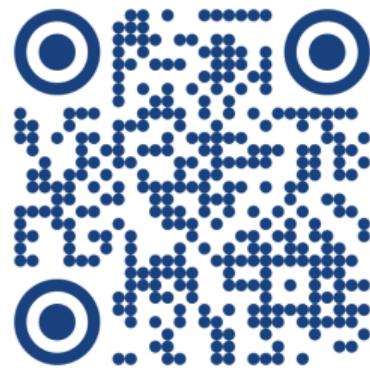


Daniel Heck

## Get In Touch

-  bjoern.siepe@uni-marburg.de
-  <https://bsiepe.github.io/>
-  bsiepe

Paper & Slides



# References i

- Freepik (2025a). Article Icon. URL <https://www.flaticon.com/free-icons/article>. Accessed: November 17, 2025.
- Freepik (2025b). Old Icon. URL <https://www.flaticon.com/free-icons/old>. Accessed: November 17, 2025.
- GitHub (2018). GitHub Octicon Logo. URL <https://commons.wikimedia.org/wiki/File:Octicons-mark-github.svg>. Licensed under MIT License <http://opensource.org/licenses/mit-license.php>; Accessed via Wikimedia Commons.
- Google (2025). Experience sampling visualization on smartphone. Image generated by Gemini 2.5 based on user prompt. Accessed: November 17, 2025.