MobileWell100+: A Multivariate Longitudinal Mobile Dataset for Investigating Individual and Collective Well-Being

Technical Report

Oresti Banos[*]
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada
oresti@ugr.es

Carlos Bailon
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada

Miguel Damas
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada

Carmen Goicoechea
Mind, Brain and Behavior Research Center (CIMCYC)
University of Granada

Pandelis Perakakis
Department of Social, Work, and Differential Psychology
Complutense University of Madrid

Hector Pomares
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada

Ciro Rodriguez-Leon
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada

Daniel Sanabria
Mind, Brain and Behavior Research Center (CIMCYC)
University of Granada

Claudia Villalonga
Research Centre for Information and
Communication Technologies (CITIC-UGR)
University of Granada

Abstract

This study engaged 103 participants over a period spanning from November 14 to December 16, 2021, ensuring representation across various demographic factors: 51 females, 52 males, aged 18-70, with varied annual incomes and from 17 Spanish regions. The MobileWell100+ dataset, openly accessible, encompasses a wide array of data, including demographic details, COVID-19-related inquiries, emotional, behavioral, and well-being data. Complementing this, social welfare data from external sources offers contextual insight. Methodologically, the project presents a promising avenue for uncovering new social, behavioral, and emotional indicators, supplementing existing literature. Notably, artificial intelligence is considered to be instrumental in analysing these data, discerning patterns, and forecasting trends, thereby advancing our comprehension of individual and population well-being. Ethical standards were upheld, with participants providing informed consent. Data collection involved smartphone app installation, passive sensor data gathering, and regular surveys, facilitating a comprehensive understanding of participants' well-being dynamics.

---

[*]Corresponding author

1 Introduction

The focus on well-being and mental health has intensified, highlighting the need for reliable methods to capture and analyze daily life intricacies. Traditional approaches are limited by retrospective methods and static measures, leaving gaps in understanding. Smartphone technology offers a solution, enabling real-time monitoring of social, behavioral, and emotional data. Objective sensor data and digital questionnaires collected via users' mobile phones can provide comprehensive insights into well-being dynamics, facilitating pattern detection and future trajectory forecasting.

To validate this methodology, a study was conducted over nearly a month, from November 14 to December 16, 2021, involving 103 participants representative of the Spanish population. The MobileWell100+ open-access dataset encompasses demographic characteristics, COVID-19-related inquiries, and emotional, social, behavioral, and well-being data. Additionally, social welfare data is sourced from external databases, including Education, Income, Health, Employment, and Social cohesion and protection, to provide context to the collected mobile data. The methodology and resulting data from this project have the potential to unveil novel social, behavioral, and emotional indicators for examining group and population dynamics, supplementing those proposed in existing literature. Notably, artificial intelligence plays a pivotal role in analyzing these data, identifying, and forecasting relevant patterns, thereby enhancing our understanding of well-being in contemporary society.

2 Study description

The participant sample for this study was sourced through a market research company that established a panel of individuals meeting the specific requirements of the study. The selection process employed a quota stratified sampling methodology to ensure representation of the Spanish population in terms of gender, age, location, and annual income.

Before the study began, candidates were provided with a detailed information sheet that outlined the study's parameters, including its duration and start date, the number of surveys to be completed, and the frequency of their completion, as well as data privacy considerations. The rewards for participating were also outlined, with a minimum of 80% completion rate of the daily surveys required to receive the reward.

All study procedures were conducted in accordance with relevant ethical guidelines and regulations. The study was approved by the Ethical Committee of the University of Granada under reference number 2214/CEIH/2021. Prior to participating, all individuals provided informed consent and confirmed that they were at least 18 years old. Participation in the study was strictly voluntary and all data collected was anonymous and confidential. The study adhered to the ethical standards outlined in the Declaration of Helsinki.

A total of 103 individuals completed the minimum required registration period (i.e., November 15, 2021 to December 15, 2021). During the study, 7 participants withdrew, but they were promptly replaced with other individuals with similar characteristics. The participants were composed of 51 (49.5%) females and 52 (51.5%) males, with ages ranging from 18 to 70 years (mean±std age=44.3±16.1). Annual net income was classified according to the criteria established by the Spanish Statistical Institute (INE). To ensure representation of the diverse Spanish population, the Nielsen Geographic Zones criteria were used. Of the 110 participants, 77 (70%) completed 80% or more of the surveys.

For the data collection, the study participants were required to install a self-developed mobile app (`https://github.com/orestibl/postcovid-ai/tree/main/mobile-app/v1`) on their smartphones, enter the identification number provided by the recruitment company, grant the necessary permissions for the app's proper functioning, and provide their digital informed consent to participate in the study. As a part of the enrollment process, participants were asked to complete an initial survey, which included demographic and COVID-19 related information, as well as questionnaires aimed at measuring their well-being.

Once enrolled, the app initiated the passive data collection through the smartphone's sensors, including physical activity recognition, and indicators of social activity such as connection type, screen usage, WiFi networks, ambient light, and noise. Moreover, the self-reported emotional data was collected using the ESMs implemented through the app. The app pushed notifications to participants at six designated times per day, randomly distributed between 7:00-8:00, 10:00-11:00, 13:00-14:00, 16:00-17:00, 19:00-20:00, and 22:00-23:00. The notification persisted for one hour before disappearing. Upon opening the notification, the app prompted participants to complete the corresponding survey, which was then transmitted to the data storage server. In addition to the daily surveys, the app prompted participants to complete weekly questionnaires on their socio-economic, health, and well-being status, to monitor any changes over time. Table 1 outlines the different data types, variables, and indicators collected via the phone.

Table 1: Overview of the data collected in the population study.

Data type	Variable	Instruments or Indicators
Sensor	Activity recognition	Detected physical activities
	Wifi	Connections to WiFi networks
	Connectivity	Type of connections with the network
	Light	Ambient light measurements
	Noise	Ambient noise measurements
	Screen	Smartphone screen status
Initial Survey	Participant's characteristics	Questions on socio-demographic and COVID-19 related data
	Psychological measures	International PANAS Short Form (I-PANAS-SF)
		General life satisfaction and seven domain of life
		Flourishing Scale (FS)
		Patient Health Questionnaire – 9 (PHQ-9)
		Generalized Anxiety Disorder Scale (GAD-7)
		Brief Resilience Scale (BRS)
		Acceptance And Action Questionnaire – II (AAQ-II)
Daily Survey	Affect	Valence
		Energetic Arousal
		Tense Arousal
	Emotional event	Report on any remarkable situations at the emotional level
Weekly Survey	Follow up variables	Questions on socio-demographic and COVID-19 related data
		General life satisfaction and seven domain of life