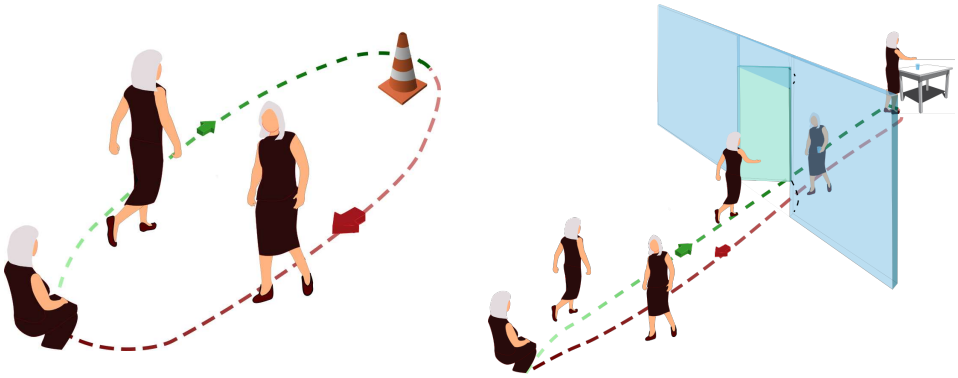
COTIDIANA DATASET – SUPPLEMENTARY MATERIALS

I. ILLUSTRATION OF ACTIVITIES

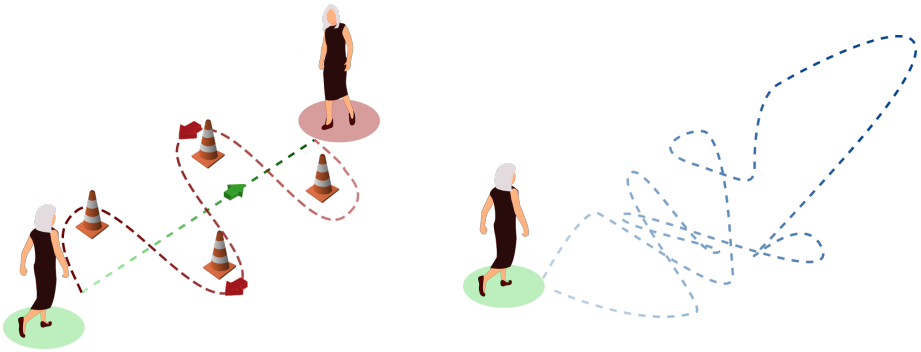
A. Mobility and Physical Activity

Fig. 1: Smartphone placement during mobility activities. Participants were asked to wear a shoulder bag with a smartphone inside, and to place the other smartphone inside the trouser pocket, whenever possible.



(a) Timed Up and Go. (b) Daily Living Activity.

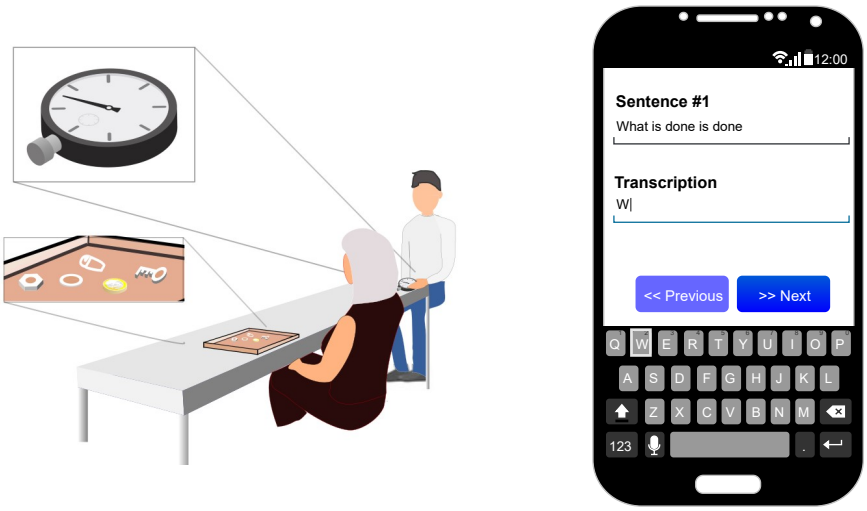
Fig. 2: Illustration of the Timed Up and Go (TUG) and the Daily Living Acitivity (DLA) activities. In TUG, the participant walks a few meters, at normal pace, standing up from a chair, bypassing a cone, and sitting down back again. In DLA, the participant executes two multitasking laps, the first (1) with hands free and the second (2) carrying a glass of water.



(a) Straight and Curvy lap. (b) Random lap.

Fig. 3: Illustration of the Simple Walk (SW) activity. This is a variant from TUG activity, where the participant is asked to execute three distinct laps, namely, (a1) a straight, (a2) a curvey, and (b) a random, free walk.

B. Finger Dexterity



(a) Moberg Pick-Up test.

(b) Keyboard Sentence Transcription activity.

Fig. 4: Illustration of the Moberg Pick-Up (MPU) test and the Keyboard Sentence Transcription (KST) activity. The MPU is a validated timed test, in which participants grab uneven-shaped objects and move them into a small box. In KST, the participants were transcribe ten sentences, at the time, using a smartphone keyboard.

TABLE I: Transcription sentences defined for each idiom (Portuguese and German). Sentences were chosen according to complexity metrics to match the spatial displacement of the characters on both portuguese and german smartphone keyboards (e.g. length, inter-character distance).

Portuguese	**German**
vamos por partes	Eile mit Weile
o seu a seu dono	Ende gut, alles gut
quem cala consente	Gelernt ist gelernt
a cada boca uma sopa	Alles zu seiner Zeit
muita parra pouca uva	Eulen nach Athen tragen
muitos cozinheiros estragam a sopa	Geklagtes Leid ist halbes Leid
bem dizer faz rir, bem fazer faz calar	Der Neid frisst seinen eigenen Herrn
boa romaria faz, quem em casa fica em paz	Die Ratten verlassen das sinkende Schiff
a bom bocado, bom grito ou bom suspiro	Geschenk vom Feind ist nicht gut gemeint
no aperto e no perigo se conhece o amigo	Eine Schwalbe macht noch keinen Sommer

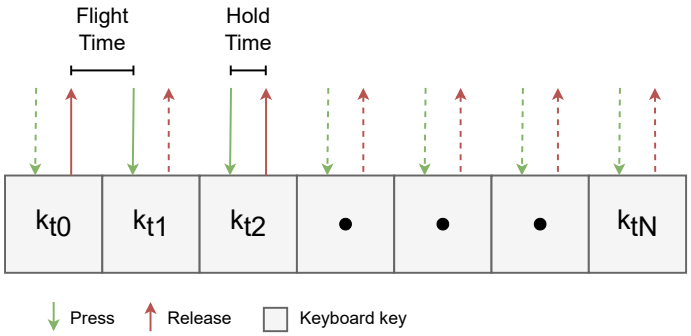


Fig. 5: Illustration of the rationale behind two common dynamic-derived keyboard metrics: flight and hold times. Such metrics help track the people's inner writing patterns over time and so their fine motor skills. Building temporal aggregates from these metrics usually offer relevant insights.

II. Data Sources

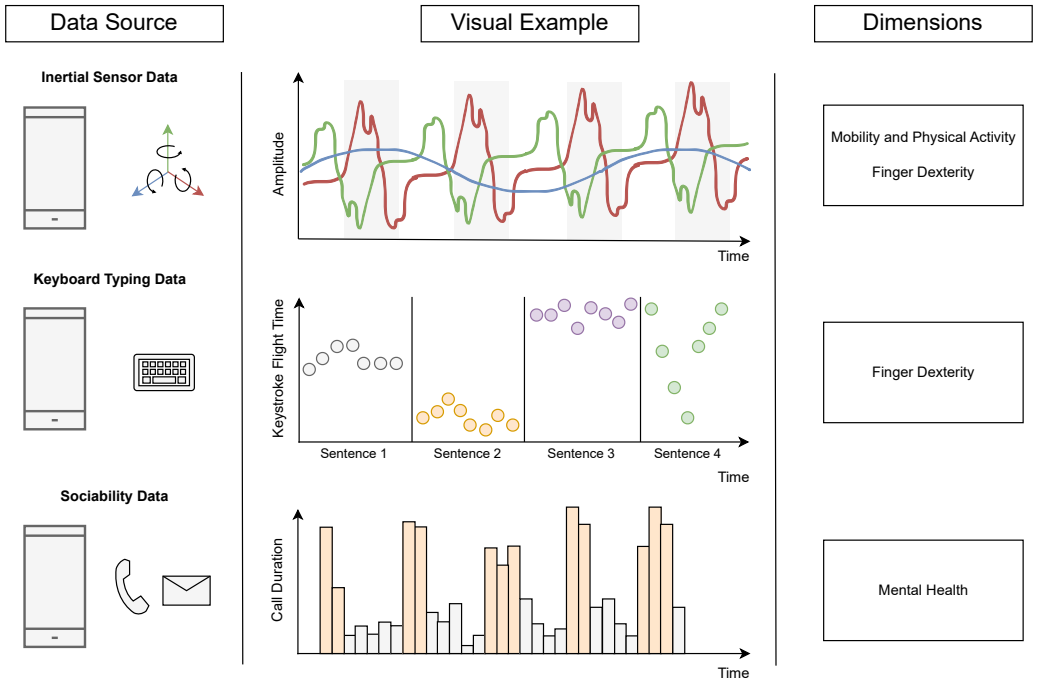


Fig. 6: Illustrative pipeline of the three main sensing data sources of the COTIDIANA Dataset, alongside the dimensions in the scope of which each was collected. Inertial sensor (accelerometer, gyroscope) and magnetometer signals were collected during mobility and dexterity activities. Data from keyboard dynamics were gathered in transcription activities. Sociability log data were retrieved from subjects' personal smartphones.

III. Data Cleaning and Annotation

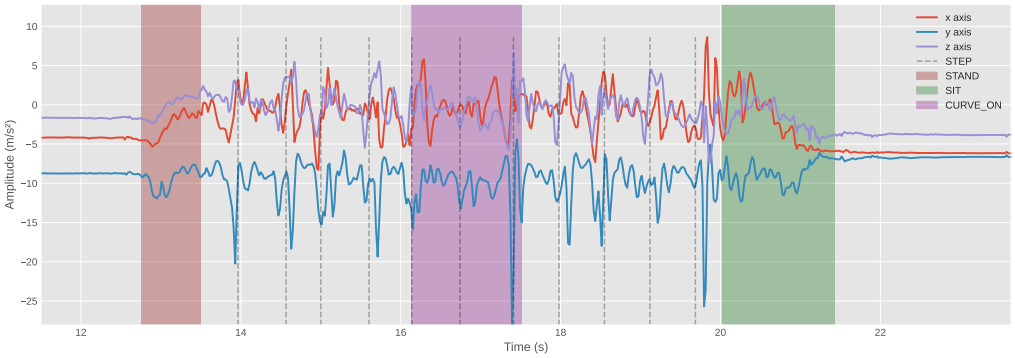


Fig. 7: Graphical representation of a Timed-Up and Go acquisition. Accelerometer signals are depicted under each of its three (x,y,z) axes, over time, alongside regions of interest manually annotated during the video labelling process. Actions of interest, namely, standing (red), curve (purple), sitting (green) and step (dashed grey) actions were manually labelled from videos being highlighted with different colours in the signal background.

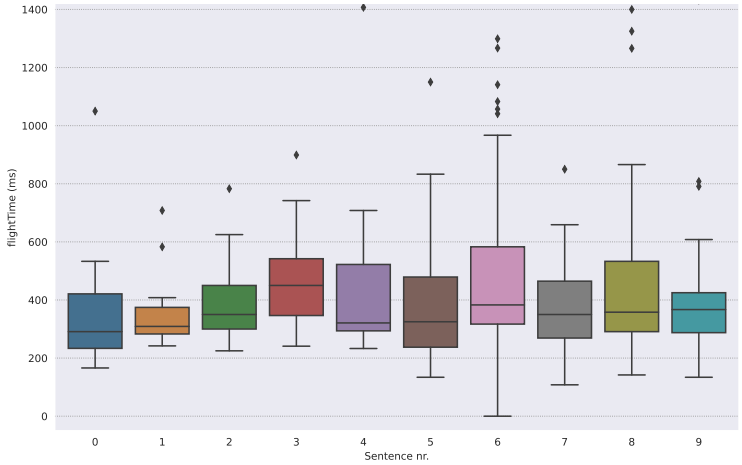


Fig. 8: Graphical representation of keyboard typing data during the sentence transcription activity. Keystroke flight times are displayed over each transcribed sentence (from 0 to 9), where the multiple flight times of each are represented in a single boxplot.

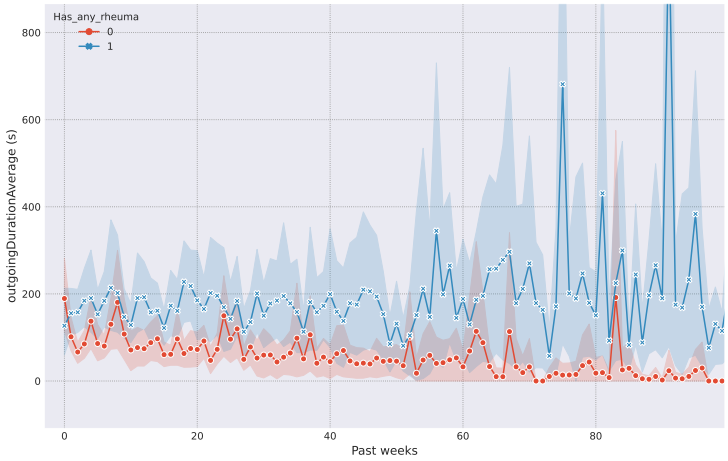


Fig. 9: Graphical representation of retrospective communications data retrieved from call logs. Outgoing Calls Duration is displayed over time (in weeks) with two parallel trends, (1) for rheumatic patients (blue) and (2) healthy controls (red).

IV. Metric Nomenclature and Formulas

TABLE II: Description of the adopted nomenclature for variables of each dimension. A standardised nomenclature has been defined according to some generic activity-related fields that were common between multiple variables.

Type of Data	Field	Options	Description	Variable nomenclature
Mobility & Physical Activity	activity	{TUG, DLA, SW}	Walking activity	MOB_{activity}_{position}_{metric}
	position	{ann, pocket, bag}	Smartphone position	
Finger Dexterity	activity	{KST, MPU}	Fine motricity activity	HD_{activity}_s{sentence number}_{source}_{metric} (KST)
	number	[0, 9]	Transcribed sentence number	
	source	{kb, sensor}	Smartphone source (keyboard/IMU)	HD_{activity}_{hand side}_{metric} (MPU)
	hand side	{dominant, non_dominant}	Hand performing the test	
Communications	type	{CALLS, SMS}	Communication type	COM_{type}_{week}_{metric}
	week	{w1, w2}	Past week number	
Questionnaires	type	{EQ, HAQ, HADS, SDQ}	Questionnaire type	QUEST_{type}_{domain}_{number}
	domain	{questionnaire domain}	Domain assessed	
	number	{0, 1, ..., N, total}	Question number	

$$X_{t_{50Hz}} = X_1 + (X_2 - X_1) \times \frac{t_{50Hz} - t_1}{t_2 - t_1} \quad (1)$$

$$MCR(X) = \frac{1}{N-1} \sum_{n=1}^{N} |sign(X_\mu[n]) - sign(X_\mu[n-1])| \quad (2)$$

$$IQR(X) = P_{75}(X) - P_{25}(X) \quad (3)$$

$$Energy(X) = \sum_{n=1}^{N} X[n]^2 \quad (4)$$

$$AUC(X) = \frac{N-1}{2} \sum_{n=1}^{N} |X[n] + X[n-1]| \times |t[n] - t[n-1]| \quad (5)$$

$$Asym_{step} = \frac{N_{steps} - 1}{2} \sum_{n=2}^{N_{steps}} \frac{|t_{step[n+1]} - t_{step[n]}|}{t_{step[n]} + t_{step[n-1]}} \quad (6)$$

$$HT_{keys} = {t_k}^{\mathrm{r}} - {t_k}^{\mathrm{p}} \quad (7)$$

$$FT_{keys} = {t_{k+1}}^{\mathrm{p}} - {t_k}^{\mathrm{r}} \quad (8)$$

Eq. 1 formulates the interpolation procedure, where t_{50Hz} are the timestamps from the 50Hz timeline, $\{t_2, t_1\}$ the adjacent timestamps from the original timeline, and X the respective signal amplitude. Further examples presented include the steps' mean crossing rate, which may differentiate signal oscillations between soft and firm steps (eq. 2), interquartile range, which measures the signal amplitude range (eq. 3), energy, which quantifies the signal energy generated by steps (eq. 4), and area under the curve, which provides insights on the step duration and smoothness (eq. 5), where X represents the signal window spanning the step range, N its length, X_μ the signal with offset (mean) removed, t the signal timeline, and P the percentile operator. Step asymmetry is also formulated by eq. 6, capturing discrepancies between the duration of consecutive pairs of steps. With respect to Finger Dexterity, keystroke hold and flight times are introduced by eqs. 7 and 8, respectively. While hold time looks for the time during a keystroke press (p) and release (r) of the same key, the flight time captures the time between a release and press of two consecutive keys (see Figure 5).

V. Auxiliary Analysis

TABLE III: Association between TUG, DLA, and SW activities. Multiple SWs correspond to the three distinct laps of the SW activity. Strong correlations were found between the three activities when evaluating general indicators e.g., test duration, sitting/standing time), indicating a fair support on the validity of the non-validated activities (DLA, SW).

Variable	Activities		*r*	*p*
Test duration (s)	TUG	DLA	0.702	<.001
		SW1	0.710	<.001
		SW2	0.770	<.001
		SW3	-	0.821
	DLA	SW1	0.790	<.001
		SW2	0.767	<.001
		SW3	-	0.266
Speed (m/s)	TUG	DLA	0.630	<.001
		SW1	0.740	<.001
		SW2	0.820	<.001
		SW3	-	-
	DLA	SW1	0.860	<.001
		SW2	0.880	<.001
		SW3	-	-
Standing time (s)	TUG	DLA	0.517	<.001
Sitting time (s)	TUG	DLA	0.611	<.001

r - pearson's coefficient; *p* - significance value at 0.05