UNIVERSITÀ DI TRENTO
Department of Information Engineering and Computer Science

Polo Ferrari 2 - Via Sommarive, 9 - 38123 Trento (Italy)
http://disi.unitn.it

A worldwide diversity pilot on daily routines and social practices (2020)

Fausto Giunchiglia, Ivano Bison, Matteo Busso, Ronald Chenu-Abente, Marcelo Rodas, Mattia Zeni, Can Gunel, Giuseppe Veltri, Amalia de Götzen, Peter Kun, Amarsanaa Ganbold, Altangerel Chagnaa, George Gaskell, Sally Stares, Miriam Bidoglia, Luca Cernuzzi, Alethia Hume, Jose Luis Zarza

April 23, 2021

Technical Report # DISI-2001-DS-01

TO BE CITED AS: Fausto Giunchiglia, Iano Bison, Matteo Busso, Ronald Chenu-Abente, Marcelo Rodas, Mattia Zeni, Can Gunel, Giuseppe Veltri, Amalia De Götzen, Peter Kun, Amarsanaa Ganbold, Altangerel Chagnaa, George Gaskell, Sally Stares, Miriam Bidoglia, Luca Cernuzzi, Alethia Hume, Jose Luis Zarza. *A worlwide diversity pilot on daily routines and social practices (2020)*. University of Trento Technical Report. No. #DISI-2001-DS-01, (2021).

A worldwide diversity pilot on daily routines and social practices (2020)

Fausto Giunchiglia[1], Ivano Bison[1], Matteo Busso[1], Ronald Chenu-Abente[1], Marcelo Rodas[1], Mattia Zeni[1], Can Gunel[1], Giuseppe Veltri[1], Amalia de Götzen[2], Peter Kun[2], Amarsanaa Ganbold[3], Altangerel Chagnaa[3], George Gaskell[4], Sally Stares[4], Miriam Bidoglia[4], Luca Cernuzzi[5], Alethia Hume[5], and Jose Luis Zarza[5]

[1] University of Trento
[2] Aalborg University
[3] National University of Mongolia
[4] University in London
[5] Universidad Católica "Nuestra Señora de la Asunción"

Abstract. This paper describes a dataset collected at the end of 2020 as part of the WeNet project, a Horizon 2020 funded project that aims at developing a diversity-aware, machine mediated paradigm for social interactions. The aim of the survey was to measure aspects of diversity based on social practices and related daily behaviours. The data collection was organized in two phases. The first involved a large sample of university students from five universities, located in Denmark, Italy, Mongolia, Paraguay and the United Kingdom. The respondents had to fill a survey aimed at investigating their social practices and specific socio-demographic, cultural and psychological elements. In the second phase, a sub-sample of the respondents participated to a 4 weeks data collection in which they were asked to fill in a self-reported time diary. This was done via a smartphone application, called iLog, which was also collecting data from thirty-four smartphone sensors, twenty-four hours a day. This dataset allows to investigate the diversity and daily routines of university students in a multi-layered perspective, both within and across countries, in a synchronic and diachronic way.

1 Background & Summary

The large-scale survey described in this document has been produced as part of a large international and multidisciplinary project called *WeNet - The Internet of us*, which aims at harnessing the diversity of people. The WeNet main aim is to bootstrap an online virtual community where the diversity of its members is leveraged and exploited to improve their "well-being" and quality of interactions. In this framework diversity is assumed to be a key distinguishing feature of life, and it is defined as the variability that exists across humans and social relations, e.g., in terms of geographical locations or mobility constraints; personal or interpersonal skills; cultural, religious, economic, or social statuses; beliefs, attitudes, desires, or intentions).

This survey was conducted involving students from the following universities: Aalborg University (AAU) (DK), London School of Economics (LSE) (UK), the National University of Mongolia (NUM) (MN), the Universidad Católica "Nuestra Señora de la Asunción" (UC) (PY), and the University of Trento (UNITN) (IT). The main features of this data collection can be summarised as follows:

- It is both cross-cultural and multi-country;
- It covers different aspects of the students' social life. The data collected allow for the study of similarities and differences in how students experience their daily life. They allow to study phenomena such as lifestyle, routine behaviours (e.g. mobility, cooking, eating) but also to develop sophisticated spatial/time use/predictive models of analysis;
- It includes two forms of collected data: self-reported data collected from an online survey, and behavioural data collected by an app that gathers data from smartphone's sensors and also collects user answers from a self-reported time diary;
- It includes both synchronic data (via questionnaires) and diachronic data (via time diary plus sensor data).

This survey was conducted int the middle of the COVID-19 pandemic. This was not planned nor expected. In this context, the criterion of ecological validity became problematic as these exceptional circumstances invalidated some of initial design choices. We managed the situation by adjusting the survey as needed. We plan to run a second similar survey at the end of 2021. The goal would be to study and compare the consequences of the pandemic in relation to the students' life.

2 Descriptor organization

It is important to note that the aim of this document is to introduce the available datasets, without providing an exhaustive description. More information can be acquired from the many citations and in particular from the datasets' documentation as described in Section 11.2.

This data descriptor is organized as follows. Section 3 provides a high-level view of how diversity has been modeled and measured; Section 4 briefly mentions the tools used; Section 5 describes how the tools have been used to perform the data collection; Section 6 provides information about the participants to this data collection; Section 7 describes how the survey has been conducted and managed; Section 8 describes the incentives; section 9 describes how we have handled privacy and ethics; Section 10 provides a brief summary of how the quality of the process and of the data has been controlled and validated; Section 11 is the core of this descriptor and provides the details of the dataset produced and of how they have been organized; Section 12 describes how the data can be obtained and the limitations on their use; Section 13 describes how some of the code used to produce or manipulate this data can be obtained, and the limitations under which this can be done; Section 14 provides info aimed at

making the use of these data easier. Finally; the document ends with some notes of the contributions by the authors, acknowledgement, and competing interests.

3 Modeling and measuring Diversity

The notion of diversity has been extensively used, and this term has been given a variety of meanings, often quite different from each other. In this work we use this term leveraging on a distinction between what we call *objectual* and *functional* diversity. The chosen terminology follows the distinction, made in [15], where anything in the real world is called *Object* when it is described in terms how it appears, i.e., of how it is perceived through senses, and is called a *Function* when it is described in term of how it behaves, i.e., of how it impacts the world. By objectual diversity, called "observable diversity" or "surface level diversity" in the Social Sciences [7, 27, 28, 34, 45, 35, 22, 23, 26], we mean here the kind of diversity that applies to "observable" demographic characteristics such as sex, culture (race, ethnicity, national origin), age, membership in formal organisations (religious or political), or physical features. On the other hand, by functional diversity, called "less observable diversity" or "deep level diversity" in the Social Sciences [7, 27, 28, 34, 45, 35, 22, 23, 26], we mean here diversity as it applies to less "observable" characteristics. Examples of deep level diversity involve technical abilities, tenure in the organisation, socio-economic and cultural background, personality traits, cognitive abilities, and values.

These two types of diversity, in the Web, but not only, play a fundamentally independent and in many ways antithetic role. Thus, while objectual diversity can be in many ways associated with *bias* [16], and is taken as a negative phenomenon which is becoming more and more pervasive in the Web [1], essential diversity can be taken as a positive phenomenon which should be exploited as much as possible. Functional diversity is actually the main focus of WeNet, whose main goal is to exploit it towards more inclusive, more extensive and better social relations. [6] In the dataset described in this paper, diversity is modeled based on earlier work in the social sciences [7, 27, 28, 34, 45, 35, 22, 23, 26]. Along this line of thought, *surface* (i.e., objectual) and *deep* (i.e., functional) diversity fall under the broader area of theory of *social practices*, when studied at the group level, and of *behavioural routines* when studied at the individual level [39, 49, 10–12, 3–5, 20, 31, 36].

In turn, following again the social sciences, in this study the social practices are modeled in terms of three components, namely:

- *Material*, namely the material objects, such as a car or a membership, which allow to exploit a certain practice,
- *Competence*, namely the knowledge, skills, and abilities which enable a certain practice, and

[6] https://www.internetofus.eu/2020/02/10/a-diversity-aware-internet-when-technology-works-for-people/.

– *Meaning*, namely a set of elements that give meaning to one or more social practices. It primarily captures the cultural component present in a specific society, as expressed by a subject while, at the same time, motivating this subject to perform that particular practice.

Thus for instance, travelling by public transport may be motivated by a personal attention to the environment. In the same way, being careful about the environment justifies (and gives meaning to) the fact that one uses a bicycle, sorts garbage, becomes vegetarian, and so on. Material, competence and meaning are the three key components used to define behavioural routines, as studied in the theory of social practices, while, at the same time, they can also be recombined to give rise to new and different behavioural routines. In turn, behavioural routines, no matter how generated, if recognised at the community level, become social practices. [42, 37, 25].

This survey has developed indicators and tools which allow to measure diversity across the three dimensions underlying social practices (competence, material and meaning) and, also, to measure how they are organised and performed in daily routines while, at the same time, maintaining a level of comparability across university students from different country and cultural communities. The process was articulated as a two-stage data collection, as follows:

– The first synchronic data collection, administered through a set of three standard close-ended questionnaire, allowed to collect self-reported general data on materialon social practices;
– The second diachronic data collection, administered via a smartphone app, allowed us to observe the students' daily routines.

4 The data collection tools

The questionnaires were managed with the LimeSurvey [32] platform. An invitation to participate in the online survey was sent through LimeSurvey to the email address of students enrolled at the various universities.

Time diaries and sensors' data collection was performed via the iLog app. [50, 19, 14, 13, 17, 17] is a list of publications which describe the use of iLog and of iLog collected data in various experiments. Currently, iLog runs on Android devices; the iOS version is under development. The possibility of collecting both user answers and sensor data makes iLog quite unique (see [38, 46, 30] for a list of other tools currently available). This double facility is quite important in that it allows to improve the state of the art in time diaries [18, 43], especially if structured [24].

5 The data collection

We had three types of data collection, namely (i) closed-ended questionnaires (synchronic), (ii) time diaries, and (iii) sensor data (the last two being diachronic).

5.1 Questionnaires

Diversity is a complex, multidimensional, and multi-layered phenomenon. In other words, it is a latent concept that cannot be captured as a whole with a single measuring instrument. Its analysis requires the decomposition of diversity into elementary parts that can be measured and reconstructed. We decided to focus only on few specific subsets of diversity. Furthermore, in order to increase the amount of information collected and to reduce the burden on respondents, the entire questionnaire was divided into three sub-questionnaires, as follows.

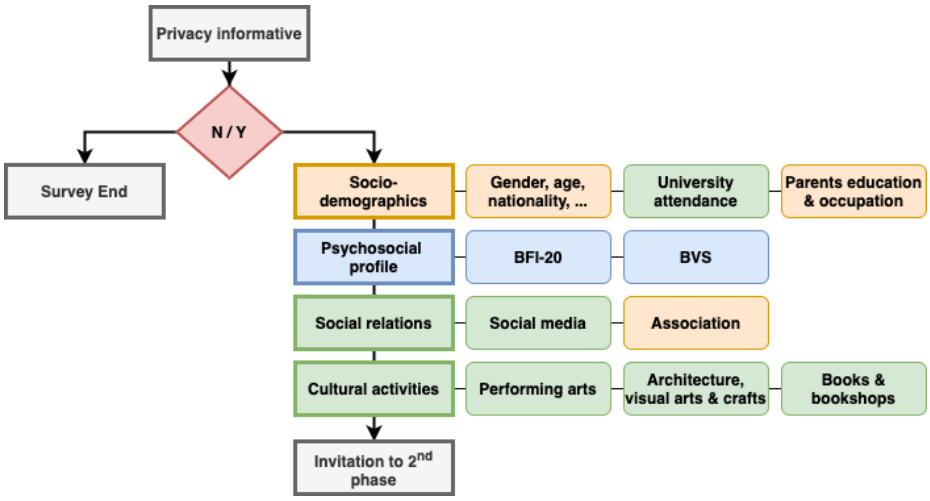


Fig. 1. Structure of the first questionnaire (Main), administered to the whole population. Material components are represented in orange, meaning components in blue, competence components in green, and other information in grey.

- The first (see Figure 1) was administered to the whole population with the aim of collecting broad general information related mainly to surface diversity and, secondly, to cultural consumption and leisure (deep diversity), and finally, to some dimensions pertaining to social relations (virtual and real).
- The second (see Figure 2) was administered only to iLog participants and was mainly devoted to finding deep diversity information. This questionnaire was mainly focused on exploring specific social practices, such as moving, cooking and grocery shopping, and physical activities.
- The third and last questionnaire (Figure 3) was also administered only to iLog participants and was, again, mainly devoted to finding deep diversity information. This questionnaire explored the user's experience with the app and testing a multiple intelligence scale.

The structure of the questionnaires and their contents are self-evident from the three figures mentioned above. All three questionnaires gathered information

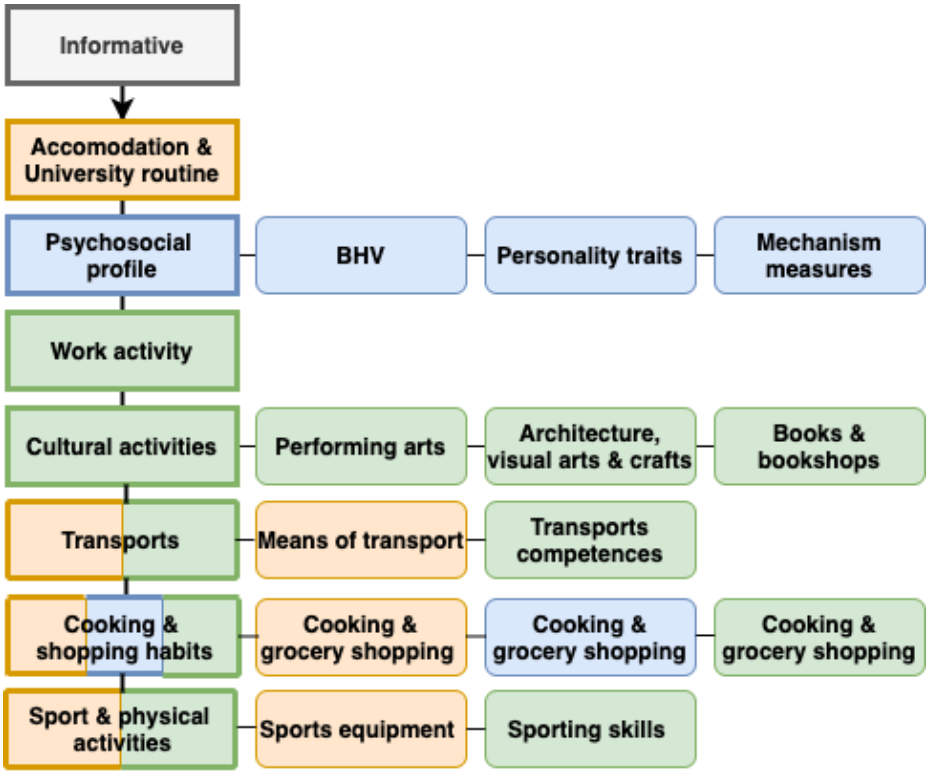


Fig. 2. Structure of the second questionnaire (iLog.1), administered exclusively to the iLog participants. Material components are represented in orange, meaning components in blue, competence components in green, and other information in grey.

related to material, competence, and meaning. In particular four standard scales were used as proxy for meaning and one as proxy for competence. Regarding meaning, the following scales were used:

- two scales about personality, namely the Big Five Inventory [9] in the main questionnaire (BFI-20 in Figure 1) and a Jungian scale on personality types [29, 6, 33, 47, 48] in the second questionnaire (Personality traits in Figure 2);
- two scales about values, namely the Basic Values Survey [21] in the main questionnaire (BVS in Figure 1) and the human values survey [40, 41] in the second questionnaire (BHV Figure 2).

Concerning competence, the multiple intelligence scale [44] was administered in the third questionnaire (Multiple intelligences in Figure 3). Each question and scale can be used as a single elementary piece of information on specific diversity characteristics. Their combination, in turn, can be used as a complex measure of diversity on specific social practices.

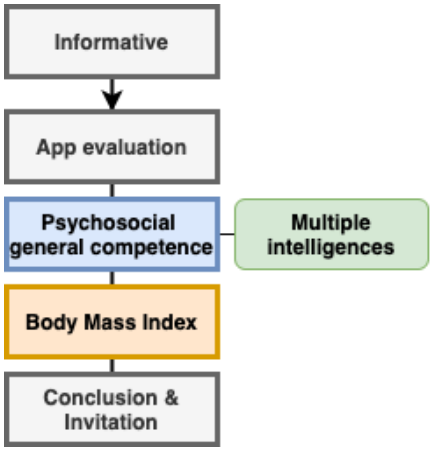


Fig. 3. Structure of the third questionnaire (iLog.2) administered exclusively to the iLog participants. Material components are represented in orange, meaning components in blue, competence components in green, and other information in grey.

5.2 Time diaries

Time Use Diaries (TUDs) are meant for how individuals spend their time. TUDs allow to measure the frequency and duration of human activities, behaviours and experiences offering a detailed view of social behaviour. In a diary study, data are self-reported activity sequences in time episodes that can range from a few days to even a month or longer with a regular time interval. This type of data is usually collected via a self-completed time diary [43] that allows registering (at fixed time intervals) the sequence of an individual's activities. For each main

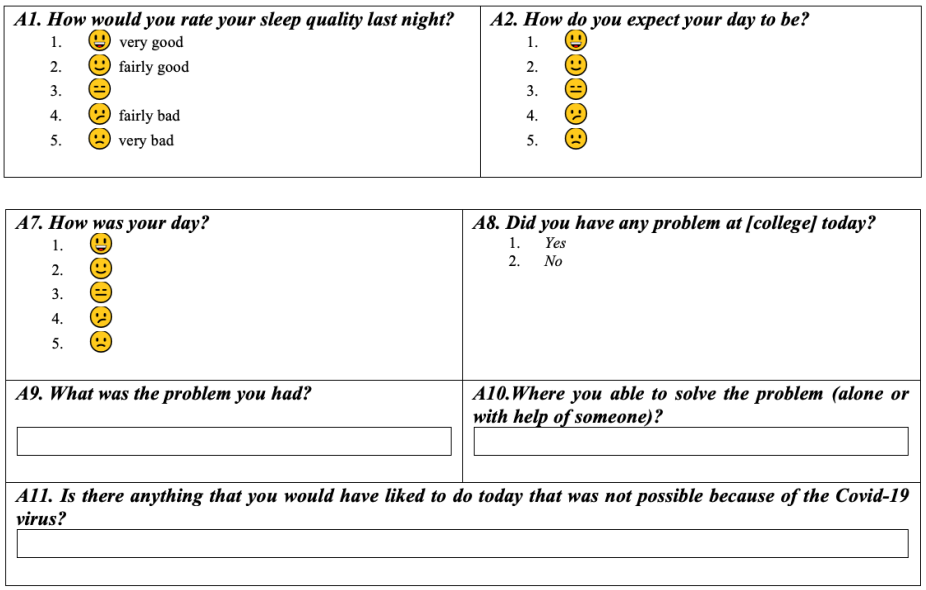


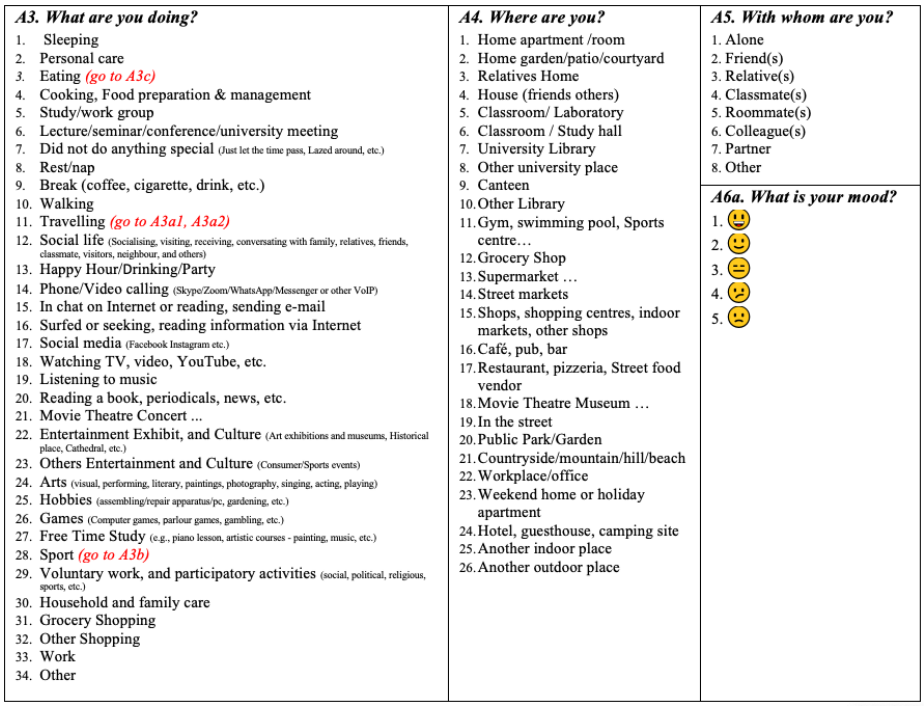
Fig. 5. Evening questions sent using iLog.

activity in each interval, additional information is usually recorded, for instance about "where" and "with whom" this activity was done.

In this survey, we have developed three different time diaries with different timings and different objectives, as follows:

- The first diary collects information about the beginning and the end of the day. Every time, at 08:00 AM (Figure 4), the subject received two qualitative questions about the sleep quality and the expectations on the day. At the end of the day, namely at 10:00 PM, (Figure 5) the subjects were asked (a) to rate their day; (b) if they had any problems during the day and (c) how did they solve them; and, finally, they received a (d) question about the COVID-19 pandemic.
- The second is a standard time diary (Figure 6) with special sections on three main activities. Every half hour for the first two weeks and every hour for the second two weeks, the participants received a notification on their smartphone with four questions as follows:
  - their activity "What are you doing?" providing the participant with 34 answer categories such as sleeping, eating, working, etc.;
  - the current location "Where are you?" providing the participant with 26 categories such as home, workplace, university, restaurant, etc.;
  - the persons being with the participants at the time of the question "Who is with you?" providing the participant with 8 categories such as nobody; partner, friends, etc; and
  - their mood "What is your mood?" providing the participant with a scale of 5 levels ranging from happy to sad.

If the subject claimed to be "eating", "traveling" or "doing sport", four different in-depth questions were asked for further information (Figure 7). Specifically



A3. What are you doing?

1. Sleeping
2. Personal care
3. Eating *(go to A3c)*
4. Cooking, Food preparation & management
5. Study/work group
6. Lecture/seminar/conference/university meeting
7. Did not do anything special (Just let the time pass, Lazed around, etc.)
8. Rest/nap
9. Break (coffee, cigarette, drink, etc.)
10. Walking
11. Travelling *(go to A3a1, A3a2)*
12. Social life (Socialising, visiting, receiving, conversating with family, relatives, friends, classmate, visitors, neighbour, and others)
13. Happy Hour/Drinking/Party
14. Phone/Video calling (Skype/Zoom/WhatsApp/Messenger or other VoIP)
15. In chat on Internet or reading, sending e-mail
16. Surfed or seeking, reading information via Internet
17. Social media (Facebook Instagram etc.)
18. Watching TV, video, YouTube, etc.
19. Listening to music
20. Reading a book, periodicals, news, etc.
21. Movie Theatre Concert ...
22. Entertainment Exhibit, and Culture (Art exhibitions and museums, Historical place, Cathedral, etc.)
23. Others Entertainment and Culture (Consumer/Sports events)
24. Arts (visual, performing, literary, paintings, photography, singing, acting, playing)
25. Hobbies (assembling/repair apparatus/pc, gardening, etc.)
26. Games (Computer games, parlour games, gambling, etc.)
27. Free Time Study (e.g., piano lesson, artistic courses - painting, music, etc.)
28. Sport *(go to A3b)*
29. Voluntary work, and participatory activities (social, political, religious, sports, etc.)
30. Household and family care
31. Grocery Shopping
32. Other Shopping
33. Work
34. Other

A4. Where are you?

1. Home apartment /room
2. Home garden/patio/courtyard
3. Relatives Home
4. House (friends others)
5. Classroom/ Laboratory
6. Classroom / Study hall
7. University Library
8. Other university place
9. Canteen
10. Other Library
11. Gym, swimming pool, Sports centre…
12. Grocery Shop
13. Supermarket …
14. Street markets
15. Shops, shopping centres, indoor markets, other shops
16. Café, pub, bar
17. Restaurant, pizzeria, Street food vendor
18. Movie Theatre Museum …
19. In the street
20. Public Park/Garden
21. Countryside/mountain/hill/beach
22. Workplace/office
23. Weekend home or holiday apartment
24. Hotel, guesthouse, camping site
25. Another indoor place
26. Another outdoor place

A5. With whom are you?

1. Alone
2. Friend(s)
3. Relative(s)
4. Classmate(s)
5. Roommate(s)
6. Colleague(s)
7. Partner
8. Other

A6a. What is your mood?

1. 😃
2. 🙂
3. 😑
4. 😕
5. 🙁

Fig. 6. Standard time diary with the questions sent every 30 minutes using iLog.

- when eating, the subject had to report foods and drinks selecting them from 20 categories, such as rice, potatoes, meat, beer, etc.;
- When doing sport, the subject had to state the type of sport selecting them from 9 categories, such as jogging and running, water sports, etc;
- When traveling, the subject had to state (a) the reason for the travel within 7 categories such as study, social life, etc. and (b) the means of transport within 16 categories such as car, bus, etc.

– In the third time diary (Figure 8), the subjects received an additional set of questions about food and drinks. These questions were asked every two hours outside the main meal hours.

5.3 Sensors

iLog collects data in the background, from a pre-selected list of sensors, with no user intervention. The data are generated as time-series, consisting of tuples composed of a timestamp and one or more values. The collected sensors are

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**
a. study *b.* social life *c.* shopping and services *d.* other leisure *e.* work *f.* changing locality *g.* other or unspecified travel purpose	*a.* on foot *b.* by bike *c.* by bus/tram *d.* by metro/ subway/ underground *e.* by train *f.* by e-scooter *g.* by car *h.* by car as passenger *i.* by car sharing *j.* by moped, motorbike *k.* by moped, motorbike as passenger *l.* by motorboat *m.* by airplane *n.* by taxi/ Uber *o.* other private transport mode *p.* other public transport mode	*a.* Walking, Trekking, and hiking *b.* Jogging and running *c.* Cycling, skiing, and skating *d.* Ball games *e.* Gymnastics and fitness *f.* Water sports *g.* Other or unspecified sports or indoor activities *h.* Other or unspecified sports or outdoor activities *i.* Productive exercise (e.g. hunting, fishing, picking berries, mushrooms, or herbs)	[MULTIPLE CHOICES] *a.* Bread, steamed buns and/or breakfast cereals *b.* Rice, potatoes, beans, pasta, noodles, dumplings, etc. *c.* Vegetables *d.* Fruits *e.* Meat *f.* Fish *g.* Processed meat (ham, bacon, sausages) *h.* Dairy products (Plain or low-fat milk, yoghurt, cheese) *i.* Soya-based food (milk, yoghurt, tofu) *j.* Pastries and sweets *k.* Snack/sandwiches (chips…) *l.* Water *m.* Soda *n.* Coffee/tea or similar *o.* Others non-alcoholic drink *p.* Beer *q.* Wine *r.* Spirit *s.* Others alcoholic drink *t.* Other food

Fig. 7. In-depth questions that appear when certain options are selected in the question "What are you doing?".

A6b. In the last two hours did you have any snacks or drinks (except breakfast, lunch, and dinner)***?***	***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)*	[MULTIPLE CHOICES] *a.* Confectionery (Candy, Chocolate, etc) *b.* Cookies, cakes, and pastries *c.* Bars (Energy bar, etc.) *d.* Crackers/biscuits *e.* Seeds, nuts, grains, legumes *f.* Savoury snacks (Chips, Tapas, Pizza, Nachos, Snack mix, deep frying) *g.* Sandwiches (Sandwich, Hamburgers, Hot dogs, Bagel) *h.* Frozen (Ice cream, Milkshake, etc.) *i.* Bread, steamed buns and/or breakfast cereals *j.* Rice, potatoes, beans, pasta, noodles, dumplings, etc. *k.* Vegetables *l.* Fruits *m.* Dairy products (milk, yoghurt, cheese) *n.* Soya-based food (milk, yoghurt, tofu) *o.* Meat *p.* Fish *q.* Processed meat (ham, bacon, sausages) *r.* Water *s.* Soda *t.* Coffee/tea or similar *u.* Others non-alcoholic drink *v.* Beer *w.* Wine *x.* Spirit *y.* Others alcoholic drink *z.* Other food

Fig. 8. Additional questions related to food and drinks.

reported in the Tables 1, 2, 3. In these tables, the value *Small/Big* in the last column (column *Category*) intuitively means that the size of the dataset generated by these sensors is comparatively small (or big) (the implications of this classification will be clear in Section 11). The sensor data collected by iLog are organized in three categories, as follows:

- *Hardware (HW) sensors*, namely the sensors that one can find in a phone, e.g., accelerometer, gyroscope, GPS. The complete list of HW sensors used in this survey is reported in Table 1;
- *Software (SW) sensors*, by which we mean all the SW events that can be collected from the Operating system and SW, for instance the Wifi the HW is connected and so on. The complete list of SW sensors is reported in Table 2;
- *QU sensors* (where QU stands for Questionnaire), by which we mean events which are connected with the compilation of the Time Use Diary, mainly related to the various execution times, e.g., when a question arrived or was answered. The complete list of QU sensors is reported in Table 3.

No	HW Sensor	Estimated Frequency	Category
1	Accelerometer	up to 10 samples per second	Big
2	Gyroscope	up to 10 samples per second	Big
3	Light	up to 10 samples per second	Big
4	Location	Once every minute	Small
5	Magnetic Field	up to 10 samples per second	Big
6	Pressure	up to 10 samples per second	Big

Table 1. HW sensors.

In these three sensor tables, the frequency by which the sensors are captured is reported, according to the following conventions: *on change* means that the value of the sensor is recorded only when the current value is changed (along with a timestamp of when it happened), *up to X samples per second* means that for each second the value of the sensor will be stored up to maximum of X times (these values are estimated), and *once every Y* means that the values of a sensor is recorded once the time Y is passed (these values are estimated). The data collected from the *HW sensors* (Table 1) are as follows:

- Sensor 1 (Accelerometer) measures the acceleration of the phone is subjected and it captures it as a 3D vector;
- Sensor 2 (Gyroscope) measures the rotational forces to which the phone is subjected and it captures it as a 3D vector;
- Sensor 3 (Light) measures the ambient illumination around the phone, measured in illuminance (lux);
- Sensor 4 (Location) returns the geocoordinates of where the phone is located, for more accuracy this sensor combines GPS and WIFI/cellular connections;
- Sensor 5 (Magnetic Field) measures the magnetic field to which the phone is subjected and it captures it as a 3D vector;

– Sensor 6 (Pressure) measures the ambient air pressure to which the phone is subjected.

No	SW Sensor	Estimated Frequency	Category
7	Airplane Mode [ON/OFF]	On change	Small
8	Battery Charge [ON/OFF]	On change	Small
9	Battery Level	On change	Small
10	Bluetooth Devices	Once every minute	Small
11	Bluetooth LE (Low Energy) Devices	Once every minute	Small
12	Cellular network info	Once every minute	Small
13	Doze Mode [ON/OFF]	On change	Small
14	Headset Status [ON/OFF]	On change	Small
15	Movement Activity Label	Once every 30 seconds	Small
16	Movement Activity per Time	Once every 30 seconds	Small
17	Music Playback (no track information)	On change	Small
18	Notifications received	On change	Small
19	Proximity	up to 10 samples per second	Small
20	Ring mode [Silent/Normal]	On change	Small
21	Running Applications	Once every 5 seconds	Small
22	Screen Status [ON/OFF]	On change	Small
23	Step Counter	up to 10 samples per second	Small
24	Step Detection	On change	Small
25	Touch event	On change	Small
26	User Presence	On change	Small
27	WIFI Network Connected to	On change	Small
28	WIFI Networks Available	Once every minute	Small

Table 2. SW sensors.

The data collected from the *SW sensors* (Table 2) are as follows:

– Sensor 7 (Airplane Mode) returns whether the phone's Airplane mode is on or off, Airplane mode turns off all the connectivity features of the phone;
– Sensor 8 (Battery Charge) returns whether the phone is currently charging its battery;
– Sensor 9 (Battery Level) returns the phone's battery level;
– Sensor 10 (Bluetooth Devices) returns all Bluetooth devices detected by the phone;
– Sensor 11 (Bluetooth Low Energy) returns all the low energy Bluetooth devices detected by the phone;
– Sensor 12 (Cellular Network info) returns information related to the cellular network (cellid, dbm, type) to which the phone is connected to;
– Sensor 13 (Doze Mode) returns whether the phone's doze mode is on or off. Doze mode is a low battery consumption state in which the phone enters after some time of not being used;
– Sensor 14 (Headset status) returns whether the headphones of the phone where connected;

- Sensor 15 (Movement activity label) returns a label identifying the activity performed by the user. This value is computed by Android using Google's Activity Recognition API along with low power signals from multiple sensors in the device. Possible activities are: *still, in_vehicle, on_bicycle, on_foot, running, tilting, walking*;
- Sensor 16 (Movement activity per Time) similar to the previous sensor, again computed via the Google API, but data are presented grouped by time instead of being grouped by labels;
- Sensor 17 (Music Playback) returns whether music is being played on the phone (yes or no) using the default music player from the operating system;
- Sensor 18 (Notifications received) measures when the phone receives a notification and when it is dismissed by the user;
- Sensor 19 (Proximity) measures the distance between the user's head and the phone, depending on the phone it may be measured in centimeters (i.e., the absolute distance) or as labels (e.g, 'near', 'far');
- Sensor 20 (Ring Mode) returns the current ring status of the phone (normal/silent/vibrate);
- Sensor 21 (Running Applications) returns the name of the application (or application package) that is currently running in the foreground of the phone;
- Sensor 22 (Screen status) returns whether the phone's screen is on or off;
- Sensor 23 (Step Counter) uses the Android API to measure the number of steps made by the user (while carrying the phone) since the phone was turned on;
- Sensor 24 (Step Detection) similar to the previous, uses the Android API to generate a step value each time the user takes a step;
- Sensor 25 (Touch event) generates a touch value each time the user touches the screen;
- Sensor 26 (User Presence) sensor that detects when the user is present near the phone, for example when the user unlocks the screen;
- Sensor 27 (WIFI Network connected to) returns information related to the WiFi network to which the phone is connected to, if connected will also report the WiFi network id;
- Sensor 28 (WIFI Networks available) returns all WiFi networks detected by the smartphone.

The data collected from the *QU sensors* are reported in Table 3. In this table we use two different concepts: *Time diary* and *Task*, the difference being that *Time diary questions /answers* are administered at fixed time intervals (as it is the case, e.g., with the questions reported in the Figures 4, 5, and 6), while *Task questions /answers* can be administered any time, depending on an event triggering them (as it is the case, e.g., with the questions reported in Figure 8). They are as follows:

- Sensor 29 (Time Diary questions) contains the question, the question id (for traceability purposes) and the timestamp when it was generated on the server and sent to the cloud provider for delivery;

- Sensor 30 (Time Diary confirmation) contains the timestamp at which each question, identified by its unique id, has been delivered to the device of the participant (which may coincide or not with the time the participant sees it);
- Sensor 31 (Time Diary answers) contains the answer, the timestamp when it was saved in the server, and the difference between answer and notification times in milliseconds;
- Sensor 32 (Task questions) contains the question and the timestamp when it was sent from the server;
- Sensor 33 (Task confirmation) contains the timestamp when each specific question was notified to the participant;
- Sensor 34 (Task answers) contains the answer, the timestamp when it was saved in the server, and the difference between answer and notification times in milliseconds.

No	QU Sensor	Estimated Frequency	Category
29	Time Diary questions	On change	Small
30	Time Diary confirmation	On change	Small
31	Time Diary answers	On change	Small
32	Task questions	On change	Small
33	Task confirmation	On change	Small
34	Task answers	On change	Small

Table 3. Questionnaire sensors.

6 The participants

The results of the data collection are described in Table 4 (one column for each of the three questionnaires plus one for iLog). The number of participants is computed after data preparation; in other words, the numbers reported as those of the people for which there are data. Concerning the first questionnaire the number of participants includes both complete and incomplete answers questionnaire. Furthermore, the numbers provided above do not consider the Chinese participants who participated only in the first questionnaire as, then, they did not participate in the followup questionnaires. The next version of the survey should be extended to include the missing data from China. The data distributed, as from Section 11, do include the data of from the Chinese participation to the first questionnaire.

In terms of geographic distribution, the participants were from 3 different continents (see Figure 9) with more than 20.000 accesses to the questionnaire of the 1st phase (see below).

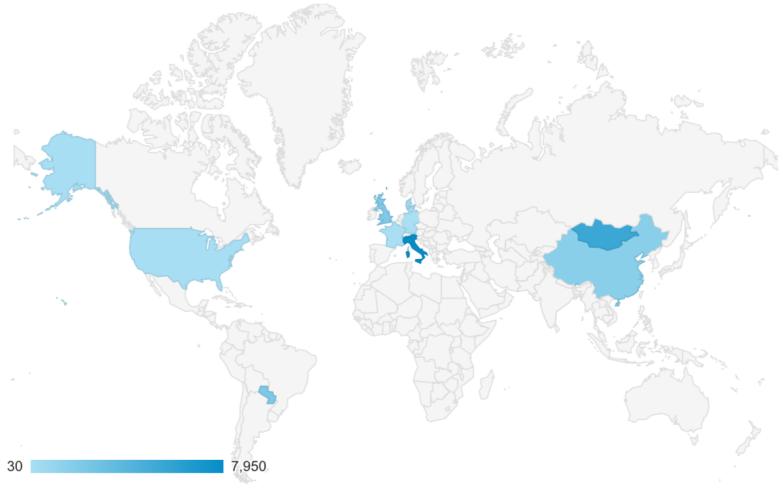


Fig. 9. Participants per country. Only countries with n $\geq$ 30 are visualized.

Site	1st QU	2nd QU	3rd QU	iLog
AAU	412	16	15	27
LSE	1980	143	45	86
NUM	3972	214	152	224
UC	1342	33	25	42
UNITN	5692	287	215	238
Total	13398	693	452	617

Table 4. Participants per site during the three waves and the iLog data collection.

7 The data collection process

The full data collection process was identically applied in all the pilot sites. Various management roles were identified, including that of translators, local survey administrators, field supervisors (in charge of monitoring and helping the participants), experiment owners (in charge of monitoring the technical aspects) among others. These organizational details, as well as the ethical and legal aspects, are described in [26].

The data collection process spanned over a period of time of approximately 13 weeks, involving participants from different countries (see Figure 9 and Table 4). The process was articulated in the following phases (see Figure 10):

1. *Translation and Adaptation.* In this phase, each site received the English version of the questionnaires and the app, including the time diaries and the list of sensors to be collected. These tools were evaluated and adapted, in coordination with all the partners, to the specific context (e.g., invitation letters, type and amount of incentives for the participants of iLog, privacy and

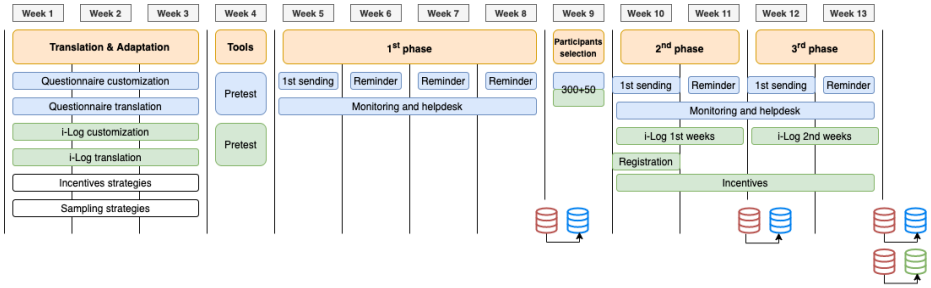


Fig. 10. Steps and phases of the data collection process.

ethics documentation etc.). Very limited changes were done in some countries to better adapt the questionnaire to the local situation or academic organisation. Concerning the standard scales mentioned above, the translations were completed by a forward translator from the original English version, and then validated back through via panel and back-translation processes by independent translators.

2. *Tools.* After translation and adaptation, the tools were tested locally. A first test was conducted in order to check and validate the translations and to evaluate the usability of the tools. A second test was conducted by sending the questionnaires to a small sample of participants, both project partners and students from the various universities. As far as questionnaires were concerned, approximately 30 participants were involved. This test was also used to ascertain the completion times. Concerning iLog, a two-week validation test was carried out.
3. *First phase.* This was the first of the three phases of the data collection. This phase started by sending an email containing the description of the survey, the invitation to the first questionnaire and information on the second part of the survey. This invitation was then reiterated through 4 reminders sent every week to all students who had not yet completed the survey.
4. *Participants selection.* At this stage, a subset of the eligible participants was selected to participate in the second part of the survey. The requirements were two: having consented to the processing of personal data and being in possession of a smartphone compatible with the app.
5. *Second phase.* This phase started by sending the second questionnaire to the selected subset of participants, followed by a reminder after one week. When sending the second questionnaire, an email was sent with the instructions for how to download iLog, accompanied by a short specification manual.
6. *Third phase.* During this last phase, the final questionnaire was sent, followed by a reminder one week after. It is worth noticing that, during this phase, the frequency of administration of time diaries via iLog was reduced.
7. *Closing the survey.* At the end of the survey a last email was sent, with the steps to follow before uninstalling the app as well as a last reminder, where needed, to fill in the second and third questionnaires.

To facilitate the monitoring of the the iLog survey and to identify possible problems, daily reports were produced containing: (1) the number of notifications each participant responded to; and (2) the amount of data collected by the individual sensors. Using this information, the local field supervisors were able to contact the inactive participants every 3 days and support them as needed. A further element of contact was the daily sending of the results of the daily prize (see below the description of incentives).

8 Incentives

No incentives were used for the first phase. Incentives were used in the second phase, covering also the third phase. The availability of questionnaires was communicated in the invitation to participate to the second phase. We had three types of remuneration:

1. Payments for completing at least the 85% of:
   - the 1st two weeks of the survey;
   - the 2nd two weeks of the survey;
2. Daily prizes (random extraction)
3. Final prizes (random extraction), for:
   - the 1st two weeks of the survey;
   - the 2nd two weeks of the survey.

The remuneration was adjusted according to the basket of goods that can be purchased in each country, see Table 5.

Incentives	AAU	LSE	NUM	UC	UNITN
Payments					
1st weeks	150kr	0	10k MNT	25k Gs	20€
2nd weeks	150kr	0	10k MNT	25k Gs	20€
Daily prizes	5 of 40kr	0	5.000MNT	10 vouchers	5 of 5€
Final prizes					
1st weeks	3 of 800kr	£150 (1/50)	100k MNT	1 restaurant voucher	3 of 100€
2nd weeks	3 of 1200kr	£150 (1/50)	150k MNT		3 of 150€

Table 5. Incentives.

9 Ethics and Privacy

All the survey activities and results at each site are compliant with the academic and national ethical privacy-protecting laws and guidelines. Additionally, for the non-European experiments, the activities and results have been developed to be also compliant to that a selected European country, as requested by the European Commission. The Italian legislation was selected as reference. The details are described in [2].

10 Technical Validation

A team of trained assistants monitored the participants to ensure compliance and quality. This team included:

- a local Field Supervisor, that spoke the language of the pilot site and interfaced with the participants;
- a central tool monitoring Technical person;
- a central Experiment Owner, in charge of monitoring the overall experiment.

Furthermore, during the phase of data preparation, the following corrective actions were performed:

- Partial records (e.g., people that filled the survey but did not participate in the app-based data collection, test users) were expunged from the dataset to ensure the completeness of the provided data.
- Data collection irregularities (e.g., data filled in wrong formats or in the wrong fields due to software-related issues) were fixed or removed.

11 Data and documentation

In this section we present: (a) how the survey dataset is organized in component datasets, with their formats; (b) the documentation; and (c) the version and release of the dataset; and (d) dataset packages which can be downloaded.

11.1 Files and formats

The data relative to the participants from Table 4, for the five sites and for each of the four main data sources (the three questionnaires and the iLog data collections) has been organized in the dataset components in Table 6. Notice that some of the sensors from Tables 1, 2 and 3 have been bundled in a single dataset component to facilitate the analysis of the data. The mapping of these datasets to the sensors described in Tables 1, 2, 3 is reported in Table 7. We have the following:[7]

- **Survey answers (limesurveys)**: the responses to the surveys as from Section 5.1, for all sites and for all questionnaires. The resulting four files, all containing the same data, are differentiated by their format (CSV, R, STATA, SPSS);
- **HW/SW/QU Sensor**: one dataset for each of the sensors labeled "small" in the sensor list tables, one dataset per site;
- **Sensor pack (Parquet-big-sensors)**: the sensors labeled "big" in the sensor list tables, one data set per site, in Apache Parquet format;

[7] The size of each dataset, as reported in the last column, is dependent on the number of participants which varies a lot (see Table 4). The number reported is an approximate average which only provides an indication of the order of magnitude.

No	Filename	Category	Num. of Files	Formats	Size (approx)
1	limesurveys	Survey answers	4 (1 per format)	R, STATA CSV, SPSS	9 MB
2	locationeventpertime	HW sensor	5 (1 per site)	CSV	6 MB
3	activitiesperlabel	SW sensor	5 (1 per site)	CSV	9 MB
4	activitiespertime	SW sensor	5 (1 per site)	CSV	9 MB
5	applicationevent	SW sensor	5 (1 per site)	CSV	114 MB
6	bluetoothlowenergyevent	SW sensor	5 (1 per site)	CSV	77 MB
7	bluetoothnormalevent	SW sensor	5 (1 per site)	CSV	112 MB
8	cellularnetwork	SW sensor	5 (1 per site)	CSV	10 MB
9	notificationevent	SW sensor	5 (1 per site)	CSV	21 MB
10	phoneevents	SW sensor	5 (1 per site)	CSV	29 MB
11	proximityevent	SW sensor	5 (1 per site)	CSV	75 MB
12	stepcounterevent	SW sensor	5 (1 per site)	CSV	16 MB
13	stepdetectorevent	SW sensor	5 (1 per site)	CSV	8 MB
14	touchevent	SW sensor	5 (1 per site)	CSV	16 5MB
15	userpresenceevent	SW sensor	5 (1 per site)	CSV	10 MB
16	wifinetworksevent	SW sensor	5 (1 per site)	CSV	88 MB
17	contributionquestions	QU sensor	5 (1 per site)	CSV	500 KB
18	timediariesconfirmation	QU sensor	5 (1 per site)	CSV	3 MB
19	tasksconfirmation	QU sensor	5 (1 per site)	CSV	900 KB
20	contributionanswers	QU sensor	5 (1 per site)	CSV	2.5 MB
21	parquet-big-sensors	Sensor pack	5 (1 per site)	Parquet	20.3 GB
22	parquet-small-sensors	Sensor pack	5 (1 per site)	Parquet	702.5 MB
23	ls-ilog-id-matching	Id matching	5 (1 per site)	CSV	3 KB

Table 6. Dataset components.

- **Sensor pack (Parquet-small-sensors)**: the sensors labeled "big" in the sensor list tables, one dataset per site, in Apache Parquet format;
- **Id matching**: the user ids which allow to merge the data from the questionnaires to the TUD and sensor data.

Special attention was dedicated to the geolocation of participants, which was collected mainly via GPS. Geolocation is considered personal information as it can lead to the re-identification of the individuals with a relatively low effort, in particular if date and time are also available [8]. As such, the raw geolocation from the dataset has been anonymized via two methods, as from below:

- **Location decimal roundown (RD)**: with two anonymization steps
  - the last decimals for latitude and longitude have been truncated from 5 decimals to 4 decimals. this means an average of 50 meter precision for locations; and
  - the year component has been removed from all timestamps, thus the data format changes from `yyyy/mm/dd/ttttttttt` to `mm/dd/ttttttttt`.
- **Point of Interest (PoI)**: all the GPS readings but the Points of Interest were deleted. In a nutshell: if latitude and longitude do not change for a

No	Dataset Component	Sensors contained
1	limesurveys	3 questionnaires
2	locationeventpertime	4
3	activitiesperlabel	15
4	activitiespertime	16
5	applicationevent	21
6	bluetoothlowenergyevent	11
7	bluetoothnormalevent	10
8	cellularnetwork	12
9	notificationevent	18
10	phoneevents	7-9, 13, 14, 17, 20, 22, 27
11	proximityevent	19
12	stepcounterevent	23
13	stepdetectorevent	24
14	touchevent	25
15	userpresenceevent	26
16	wifinetworksevent	28
17	contributionquestions	29, 32
18	timediariesconfirmation	30
19	tasksconfirmation	33
20	contributionanswers	31, 34
21	parquet-big-sensors	1, 2, 3, 5, 6
22	parquet-small-sensors	4, 7-34

Table 7. Matching sensors to dataset components.

period of time then a Point of Interest (PoI) tag is added to the stream with the name of the nearest PoI and a timestamp.

The anonymization has resulted in the duplication of the datasets reported in Table 6. The naming convention followed for naming datasets is as follows:

`WeNet1-Filename-site-anonym-format`

where `Filename` is a value of the first column in Table 6, `site` is the acronym of one of the universities, `anonym` is one of `rd` or `poi`, and `format` is the name of the format.

11.2 Documentation

The main documentation consists of the following:

- **Descriptor**: this document, a general high-level description of the survey data. File name:

  `WeNet1-doc-descriptor`;

- **LimeSurvey**: one document per site (five in total). It contains the complete set of questionnaire questions and answer options (all three questionnaires). File name:

`WeNet1-doc-limesurvey-site`

– **iLog**: one document per site (five in total). It contains the complete set of Time Use Diaries and the sensors used with their configurations. File name:

`WeNet1-doc-ilog-site`

– **Codebook**: one document per site (five in total). It contains extensive and detailed structural and statistical information about the dataset. File name:

`WeNet1-doc-codebook-site`

All the documents above are included in a compressed single file which can be downloaded together with the reference datasets.

11.3 Version

First Version and release.

11.4 Download

The datasets in Table 6 are organized in a set of *packages*, as follows

– **limesurveys**: one per five sites plus one overall, times 2 anonymizations, times 4 formats. Total: 48 packages;
– **csv-small-sensors**: one per five sites plus one overall, times 2 anonymizations. Total: 12 packages;
– **parquet-big-sensors**: one per five sites plus one overall, times 2 anonymizations. Total: 12 packages;
– **parquet-small-sensors**: one per five sites plus one overall, times 2 anonymizations. Total: 12 packages;
– **parquet-all-sensors**: parquet-small-sensors, plus parquet-small-sensors. Total: 12 packages;
– **parquet-all-data**: limesurveys, plus parquet-all-sensors, plus Id Matching. Total: 48 packages;
– **csv-all-data**: limesurveys, plus, csv-small-sensors, plus Id Matching. Total: 48 packages.

Each package contains also the relevant documentation as from Section 11.2 (possibly duplicated or redundant).

12 Survey data availability

The main entry point documentation for this survey is the Research infrastructure Catalogue.[8] This catalogue can be searched via metadata that provide a high level description of the datasets. This catalogue contains one entry point for each of the download packages listed in Section 11.4. The catalogue enables direct to all the documentation, as from Section 11.2. However, because of the type of data, to be fully compliant with GDPR, in order to have access to any of the datasets described in this document, a licence must be signed. The details of how to enable this are provided in the catalogue. The licence is not needed for WeNet partners. The constraints under which these data can be used are detailed in the licence. Some relevant licensing conditions are: (i) the data of this survey may only be used for research purposes; (ii) Redistribution of these data is forbidden; (iii) This data cannot be made public (e.g., on a Web site) or given to a third part; and (iv) the use of this data must be properly credited by citing this report (as University of Trento Technical report, or as journal data publication, see the details in the first page).

13 Survey code availability

No code availability.

14 Usage Notes

No usage notes.

Author contributions

The order of names is by contribution of the Institution and, inside each Institution, by contribution of the individuals. As such, the order of names does not necessarily reflect the importance of the contribution of the single individuals. The roles of the authors, presented by by their initials, is as follows:

- *Experiment management*: F.G., I.B., A.D.G.;
- *Experiment design*: F.G., I.B., A.D.G., M.B., R.C.A., G.V.;
- *Technical support*: M.R., M.Z., C.G., M.B.;
- *Data Collection*: M.B., R.C.A., M.R., A.D.G, P.K.,A.G., A.C., G.G., S.S, M.B., L.C., A.H., J.L.Z.;
- *Data Preparation and correction*: R.C.A., C.G., I.B., M.B..

[8] WeNet Research Infrastructure; https:ri.internetofus.eu. University of Trento, DataScientia: http://livepeople/datascientia.disi.unitn.it.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 FET Proactive project "WeNet – The Internet of us", grant agreement No 823783. This research have deeply benefitted from the huge number of discussions that we have had with all the people working in the WeNet project.

Competing interests

All authors declare no competing interests during the data collection, preparation and analysis of this dataset.

References

1. Baeza-Yates, R.: Bias on the web. Communications of the ACM **61**(6), 54–61 (2018). https://doi.org/10.1145/3209581
2. Bona, R., Chenu-Abente, R., Bison, I., Busso, I., Giunchiglia, F.: Personal data collection: Ethical and privacy compliance procedures. Tech. rep., University of Trento (2020)
3. Bourdieu, P.: Outline of a Theory of Practice. No. 16, Cambridge university press (1977)
4. Bourdieu, P.: Distinction: A social critique of the judgement of taste. Harvard university press (1984)
5. Bourdieu, P.: The logic of practice. Stanford university press (1990)
6. Briggs-Myers, I., Myers, P.B.: Gifts differing: Understanding personality type (1995)
7. Cummings, A., Zhou, J., Oldham, G.R.: Demographic differences and employee work outcomes: Effects on multiple comparison groups. In: annual meeting of the Academy of Management, Atlanta, GA (1993)
8. De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M., Blondel, V.D.: Unique in the crowd: The privacy bounds of human mobility. Scientific reports **3**(1), 1–5 (2013)
9. Donnellan, M.B., Oswald, F.L., Baird, B.M., Lucas, R.E.: The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. Psychological assessment **18**(2), 192 (2006)
10. Dreyfus, H.: Being-in-the-world: A commentary on heidegger's being and time (1991)
11. Giddens, A.: Central problems in social theory: Action, structure, and contradiction in social analysis, vol. 241. Univ of California Press (1979)
12. Giddens, A.: The Constitution of Society. Cambridge: Polity Press (1984)
13. Giunchiglia, F., Bignotti, E., Zeni, M.: Human-like context sensing for robot surveillance. International Journal of Semantic Computing **12**(01), 129–148 (2017)
14. Giunchiglia, F., Bignotti, E., Zeni, M.: Personal context modelling and annotation. In: 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). pp. 117–122. IEEE (2017)
15. Giunchiglia, F., Fumagalli, M.: Teleologies: Objects, actions and functions. In: ER-International Conference on Conceptual Modeling. pp. 520–534. ICCM (2017)

16. Giunchiglia, F., Otterbacher, J., Kleanthous, S., Batsuren, K., Bogina, V., Tal, A.S., Kuflik, T.: Diversity, bias and algorithmic transparency. arXiv preprint arXiv:2104.05658 (2021)
17. Giunchiglia, F., Zeni, M., Big, E.: Personal context recognition via reliable human-machine collaboration. In: 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). pp. 379–384. IEEE (2018)
18. Giunchiglia, F., Zeni, M., Bignotti, E., Zhang, W.: Assessing annotation consistency in the wild. In: 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). pp. 561–566. IEEE (2018)
19. Giunchiglia, F., Zeni, M., Gobbi, E., Bignotti, E., Bison, I.: Mobile social media and academic performance. In: International conference on social informatics. pp. 3–13. Springer, Cham (2017)
20. Goffman, E.: Asylums: Essays on the social situation of mental patients and other inmates. Harmondsworth: Penguin (1975)
21. Gouveia, V.V., Milfont, T.L., Guerra, V.M.: Functional theory of human values: Testing its content and structure hypotheses. Personality and Individual Differences **60**, 41–47 (2014)
22. Harrison, D.A., Price, K.H., Bell, M.P.: Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion. Academy of management journal **41**(1), 96–107 (1998)
23. Harrison, D.A., Price, K.H., Gavin, J.H., Florey, A.T.: Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. Academy of management journal **45**(5), 1029–1045 (2002)
24. Hellgren, M.: Extracting more knowledge from time diaries? Social Indicators Research **119**(3), 1517–1534 (2014)
25. Ivano, B., Giuseppe Alessandro, V., George, G.: D1.1: Early taxonomy of diversity. Tech. rep. (2020), written by the "WeNet - The Internet of US" (WeNet) project's consortium under EC grant agreement 823783
26. Ivano, B., Matteo, B., Ronald, C.A., Martina, C., George, G., Marcelo Dario, R.B., Giuseppe, S., Sally, S., Giuseppe Alessandro, V.: D1.4 final model of diversity: The research protocol of the diversity pilot study. Tech. rep. (2021), written by the "WeNet - The Internet of US" (WeNet) project's consortium under EC grant agreement 823783
27. JACKSON, S.E.: Consequences of group composition for the interpersonal dynamics of strategic issue processing. Advances in Strategic Management **8**, 345–382 (1992), https://ci.nii.ac.jp/naid/20001538270/en/
28. Jackson, S.E., May, K.E., Whitney, K.: Understanding the dynamics of diversity in decision-making teams
29. Jung, C.G.: Psychological types. collected works of cg jung. Tr. HG Baynes. Rev. RFC Hull. Princeton: Princeton UP (1971)
30. Kreuter, F., Haas, G.C., Keusch, F., Bähr, S., Trappmann, M.: Collecting survey and smartphone sensor data with an app: Opportunities and challenges around privacy and informed consent. Social Science Computer Review **38**(5), 533–549 (2020)
31. Latour, B., Woolgar, S.: Laboratory Life: The Construction of Scientific Facts, vol. 80. Princeton University Press (1986)
32. LimeSurvey Project Team / Carsten Schmitz: LimeSurvey: An Open Source survey tool. LimeSurvey Project, Hamburg, Germany (2012), http://www.limesurvey.org
33. Mascarenas, D.D.L.: A jungian based framework for artificial personality synthesis. In: EMPIRE@ RecSys. pp. 48–54 (2016)

34. Maznevski, M.L.: Understanding our differences: performance in decision-making groups with diverse members. Human relations **47**(5), 531–552 (1994)
35. Pelled, L.H.: Demographic diversity, conflict, and work group outcomes: An intervening process theory. Organization science **7**(6), 615–631 (1996)
36. Reckwitz, A.: Toward a theory of social practices: A development in culturalist theorizing. European journal of social theory **5**(2), 243–263 (2002)
37. Røpke, I.: Theories of practice—new inspiration for ecological economic studies on consumption. Ecological economics **68**(10), 2490–2497 (2009)
38. Runyan, J.D., Steenbergh, T.A., Bainbridge, C., Daugherty, D.A., Oke, L., Fry, B.N.: A smartphone ecological momentary assessment/intervention "app" for collecting real-time data and promoting self-awareness. PloS one **8**(8), e71325 (2013)
39. Schatzki, T., Cetina, K.K.: The practice turn in contemporary theory (2001)
40. Schwartz, S.H.: Are there universal aspects in the structure and contents of human values? Journal of social issues **50**(4), 19–45 (1994)
41. Schwartz, S.H., Melech, G., Lehmann, A., Burgess, S., Harris, M., Owens, V.: Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. Journal of cross-cultural psychology **32**(5), 519–542 (2001)
42. Shove, E., Pantzar, M.: Consumers, producers and practices: Understanding the invention and reinvention of nordic walking. Journal of consumer culture **5**(1), 43–64 (2005)
43. Sorokin, P.A., Berger, C.Q.: Time-budgets of human behavior, vol. 2. Harvard University Press (1939)
44. Tirri, K., Nokelainen, P.: Identification of multiple intelligences with the multiple intelligence profiling questionnaire iii. Psychology Science **50**(2), 206 (2008)
45. Tsui, A.S., Egan, T.D., O'Reilly III, C.A.: Being different: Relational demography and organizational attachment. Administrative science quarterly pp. 549–579 (1992)
46. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A.T.: Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing. pp. 3–14 (2014)
47. Wilde, D.J.: Teamology: the construction and organization of effective teams. Springer Science & Business Media (2008)
48. Wilde, D.J.: Jung's personality theory quantified. Springer Science & Business Media (2011)
49. Wittgenstein, L., Anscombe, G.: Philosophical investigations (1953)
50. Zeni, M., Zaihrayeu, I., Giunchiglia, F.: Multi-device activity logging. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. pp. 299–302 (2014)