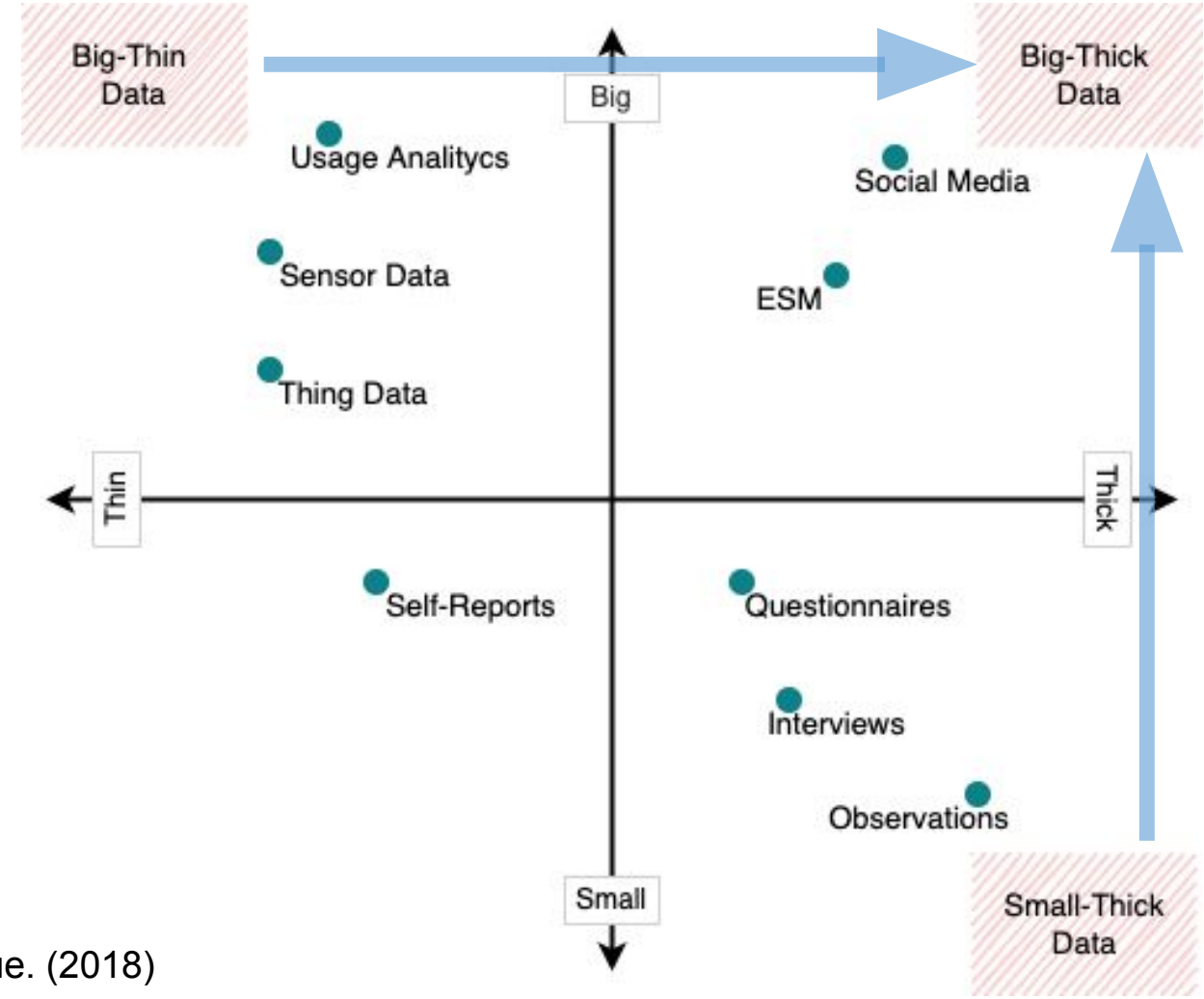
# Representing behavioral data in context: Behavioral data as Knowledge Graphs (KG)

Knowledge Graph Engineering

# Big-Thick Data




Adapted from:

- Gomez Ortega, A., et al. (2022)
- Tobias Bornakke and Brian L. Due. (2018)

# What is Personal Context?

Context is "a theory of the world which encodes an individual's subjective perspective about it ." [1]


[1] F. Giunchiglia. Contextual reasoning. Epistemologia, special issue on I Linguaggi e le Macchine, 16:345–364, 1993.

# Existing Personal Context Models

## CONON model [2]

- Focus on modeling locations.
- Upper ontology and lower domain-specific ontology
- Lower domain specific ontology depends on the use case scenario
- They applied this model in a SmartHome environment.

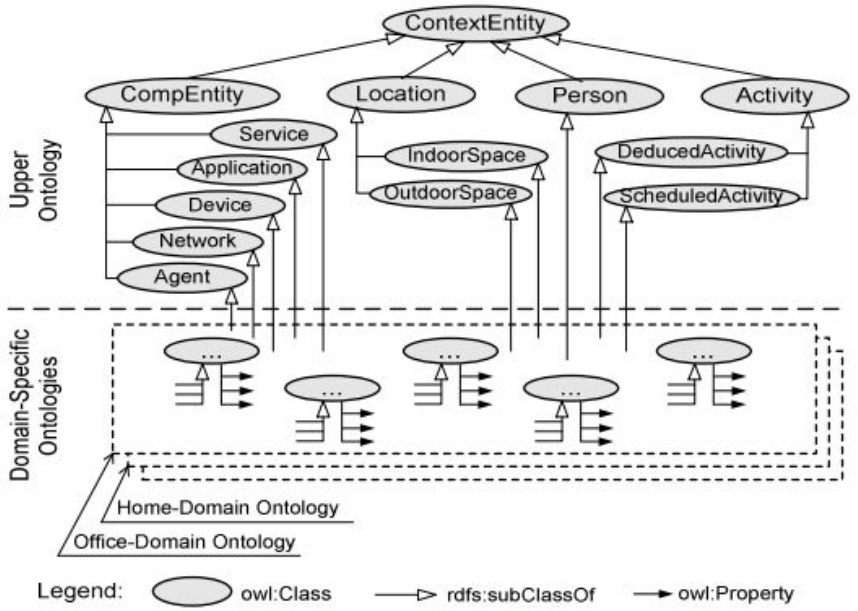



Figure 1. Partial Definition of CONON upper ontology

[2] Wang, Xiao Hang, et al. "Ontology based context modeling and reasoning using OWL." IEEE annual conference on pervasive computing and communications workshops, 2004. Proceedings of the second. Ieee, 2004.

# Existing Personal Context Models

## CaCONT model[3]

- Fucose on five elements:
  - User, space, environment, device, service
- One model for each element
- Use rule-based system to do recognition

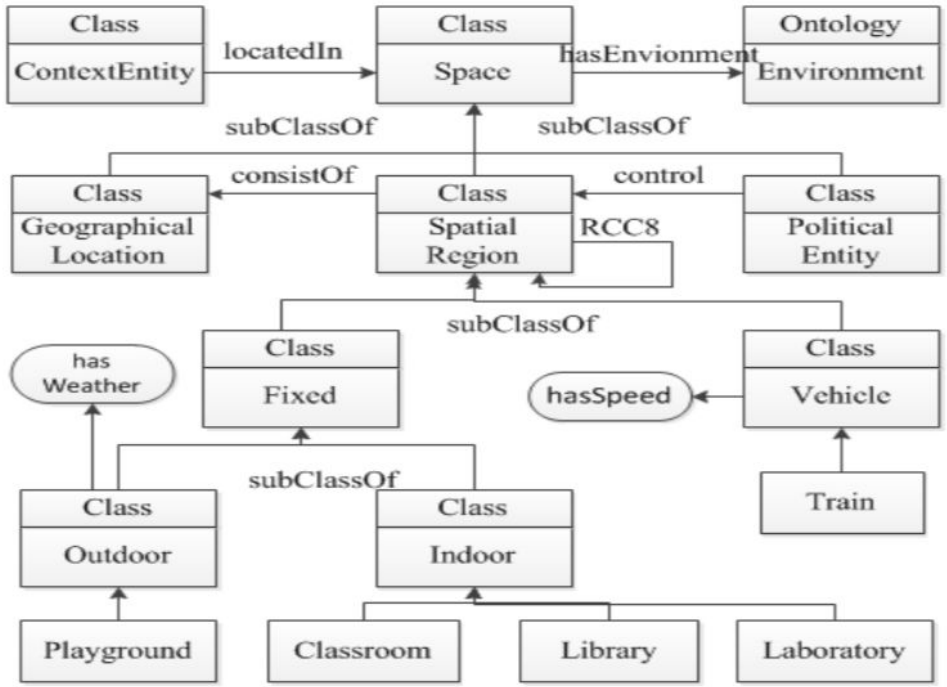



Figure 4. Space Model

[3] Xu, Nan, et al. "CACOnt: A ontology-based model for context modeling and reasoning." Applied Mechanics and Materials. Vol. 347. Trans Tech Publications Ltd, 2013.

# Existing Personal Context Models

## Two main limitations:

- They do not account for subjectivity of context descriptions
  - "Supermarket" is "a place for shopping" for normal customer,
  - but it is "a working place" for staffs who work there.

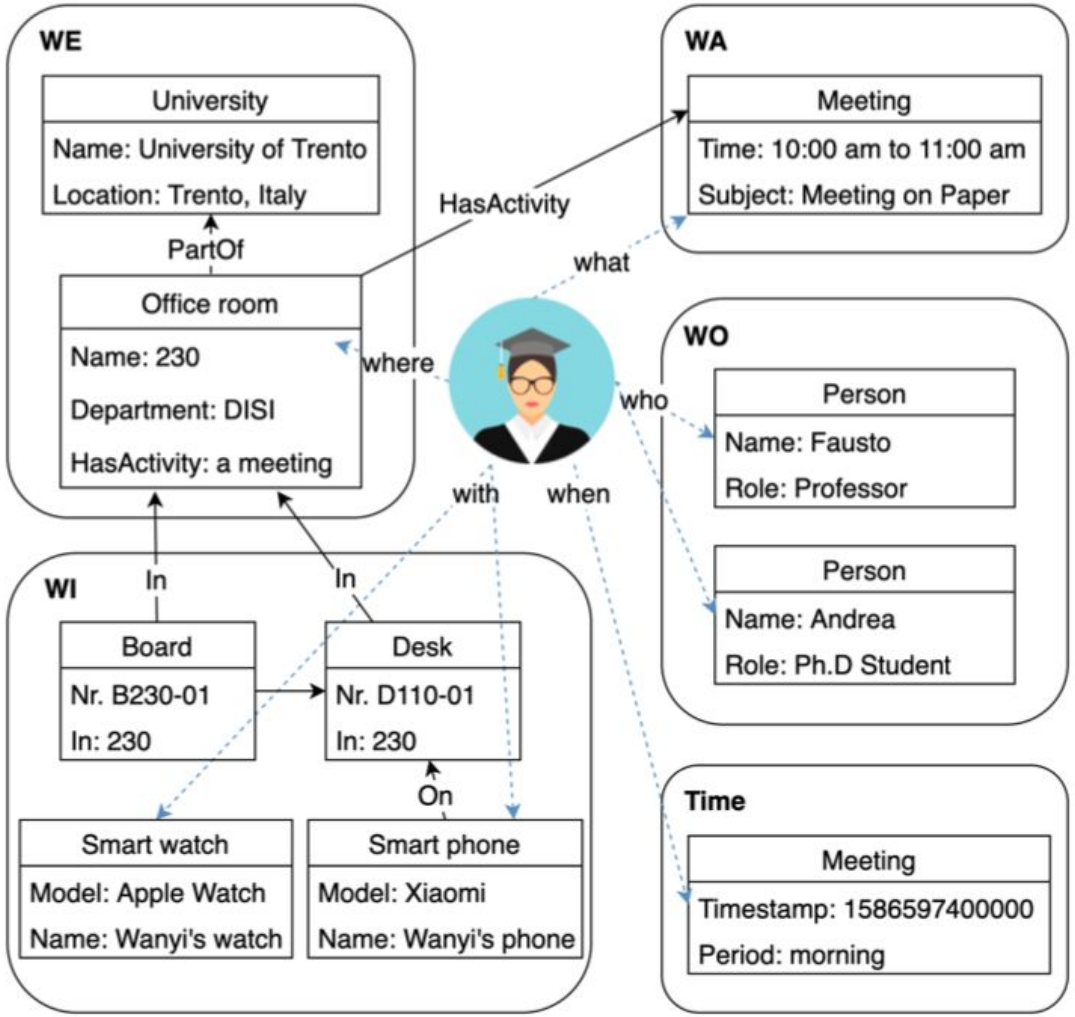
Subjective context

- The correlations between different aspects are important, while most model are restrict to one of few of the aspects.

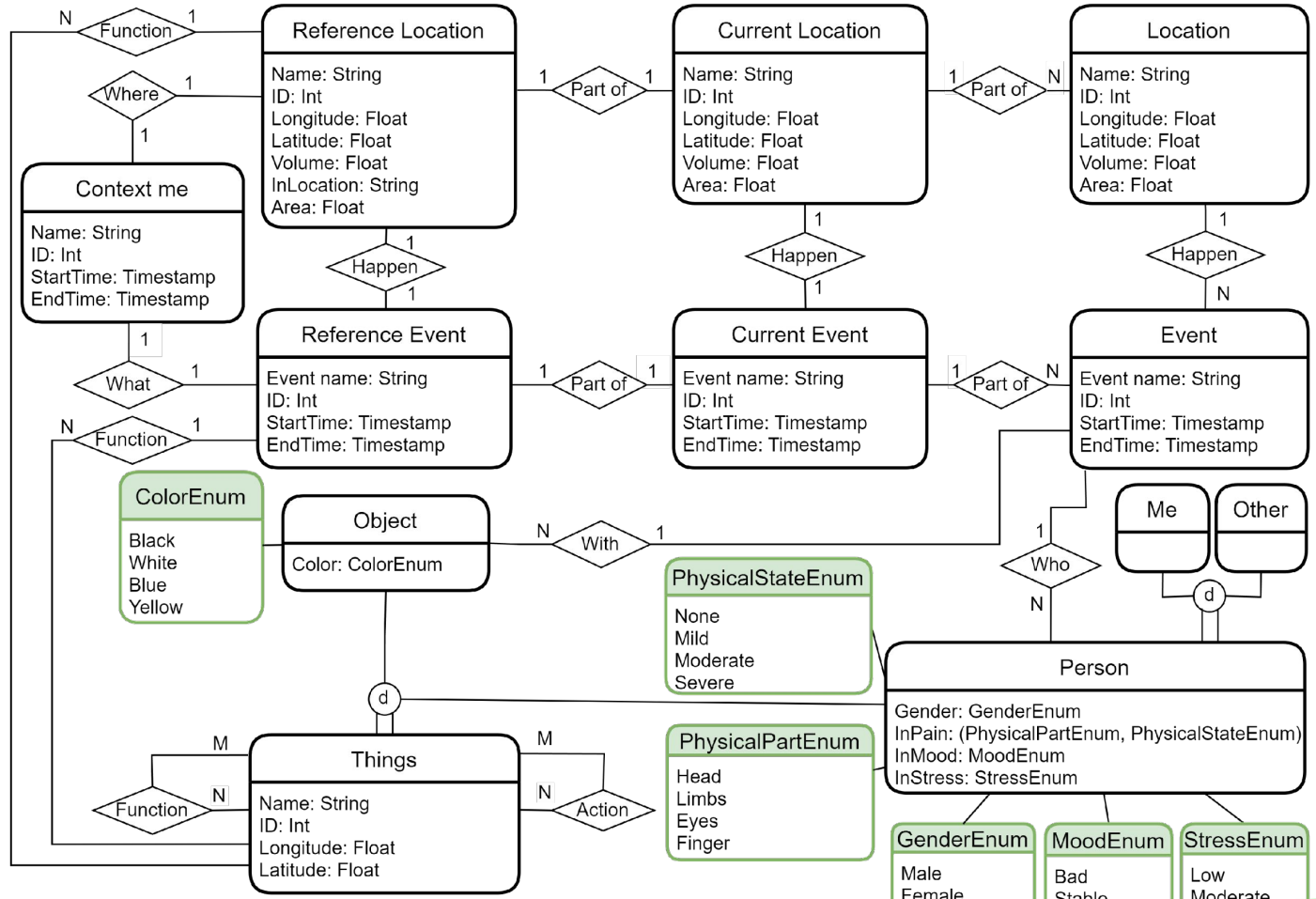
Multi-modality context model

# Personal Context Model

- MyWorld = <me, Context>
- Context = <TIME, WE, WA, WO, WI>
  - TIME: "When does this context occur?"
  - WE: "Where are you?"
  - WA: "What are you doing?"
  - WO: "Who is with you?"
  - WI: "What are you with?"



# Entity Type Graph (ETG) Representation

# Endurant & Perdurant Context

- Context aggregates different dimensions of a person's environment basing on points of view, i.e., humans fundamentally use two elements to drive their representation: location and activity.
- The context can provide different representation of the same state of affairs depending on which element is more important.
- **Endurant context**: "I'm in the office".
- **Perdurant context**: "I'm having a meeting"

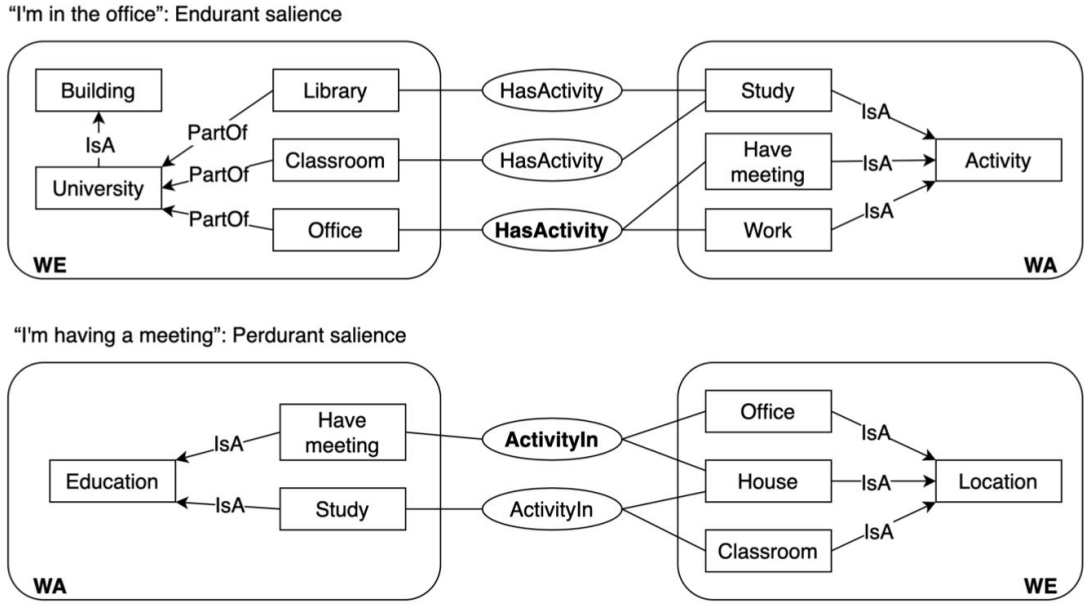



Figure 4.2: The difference between the notions of endurant and perdurant context.

# Objective & Subjective Context

| Level | TIME | WE | WA | WO |
|---|---|---|---|---|
| Objective Context | 2020-04-11 11:30am | Via Sommarive, 9, 38123 Povo TN | Talk | Prof. Fausto |
| Subjective Context | Morning | Office | Have meeting | Supervisor |
| Machine Context | 1586597400000 | 46°04'01.9"N 11°09'02.4"E | Accelerometer: 0g,0g,0g | "Fausto" is in contact list |

Table 4.1: An example of our three-partitioned context model.

# Personal Context Annotation - Activities

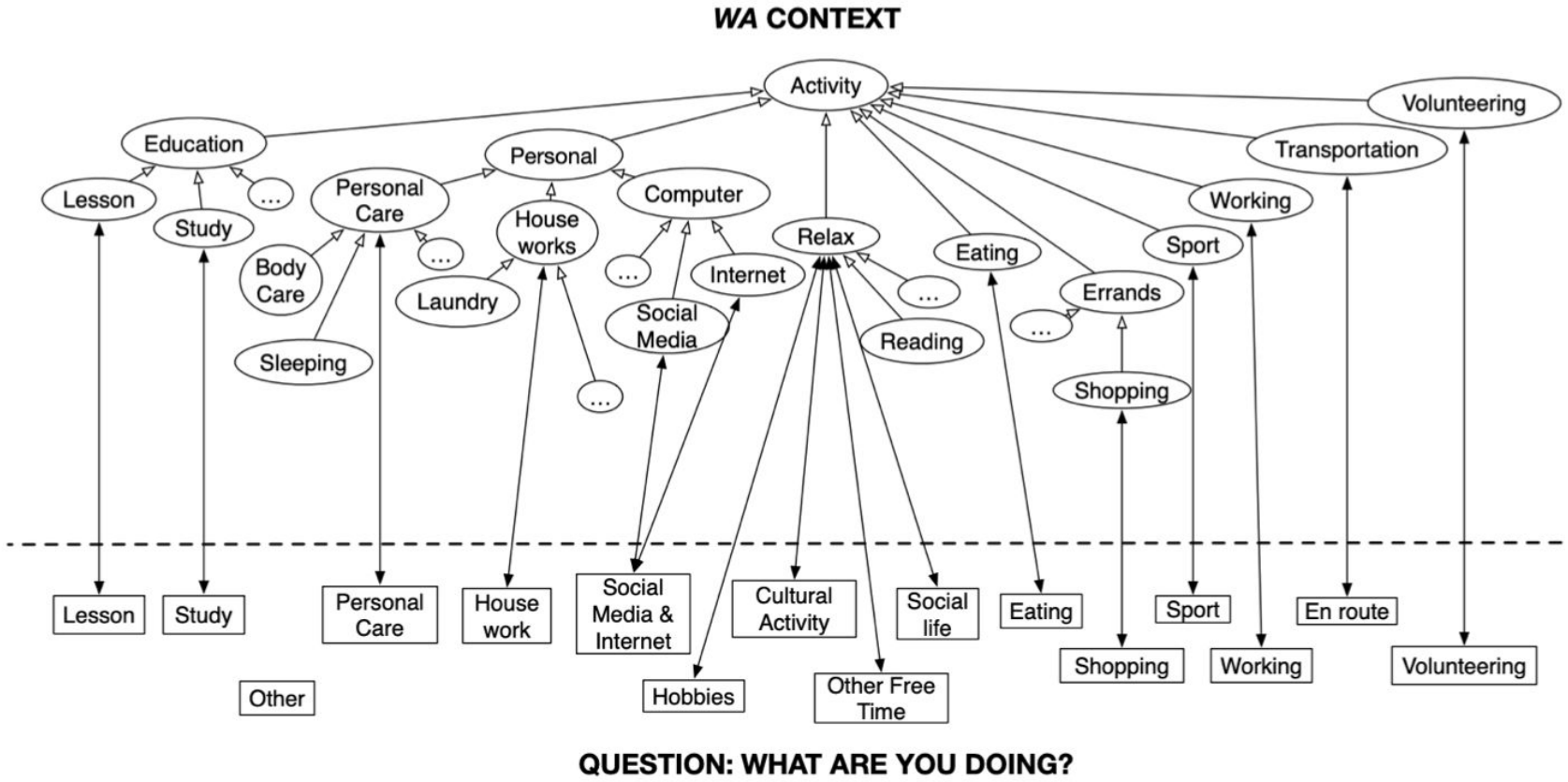



Figure 4.3: The mapping from the WA context to the activities annotation list.

# Personal Context Annotation - Locations

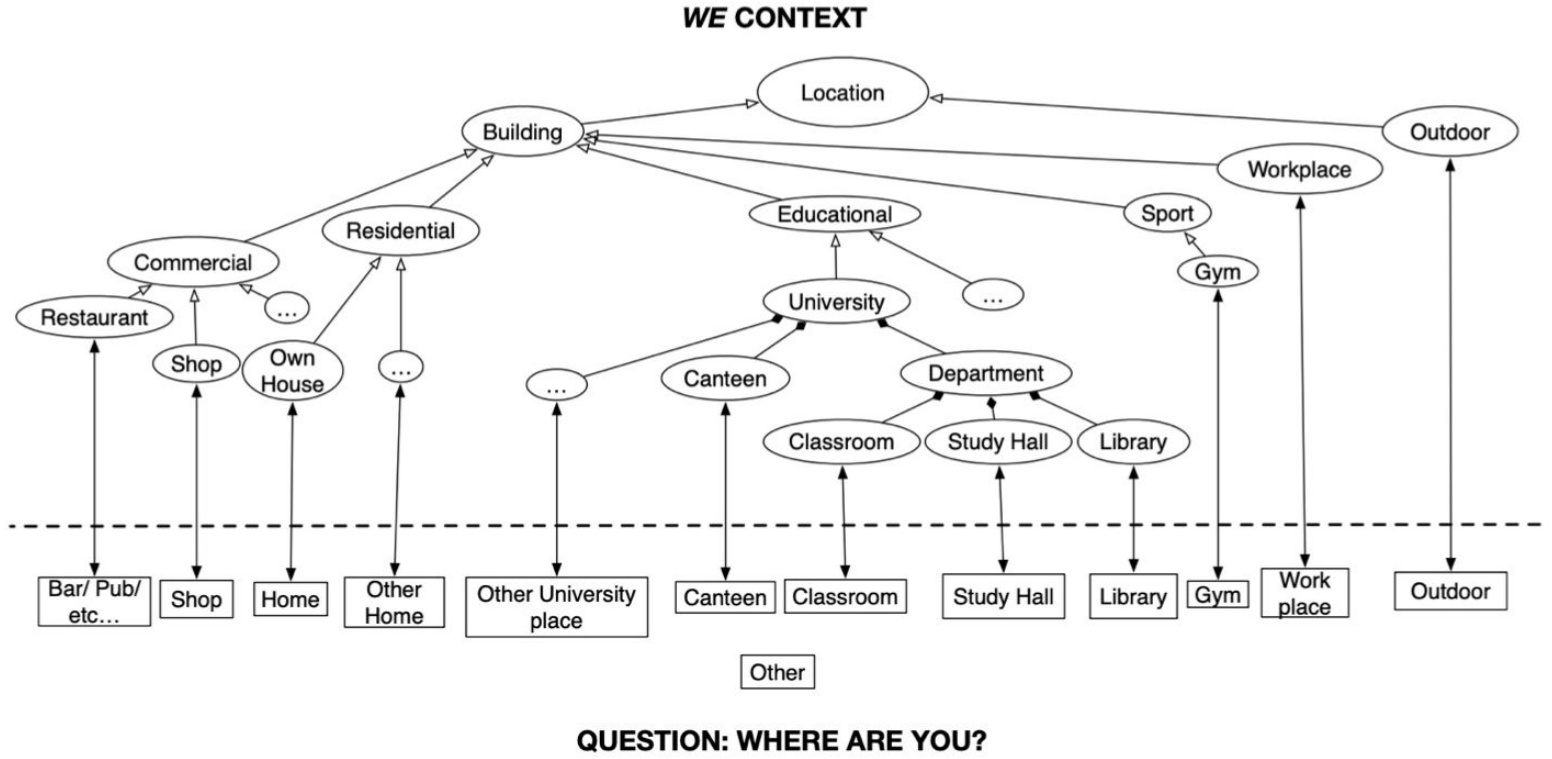



Figure 4.4: The mapping from the WE context to the locations annotation list.

# Personal Context Annotation – Social Relations

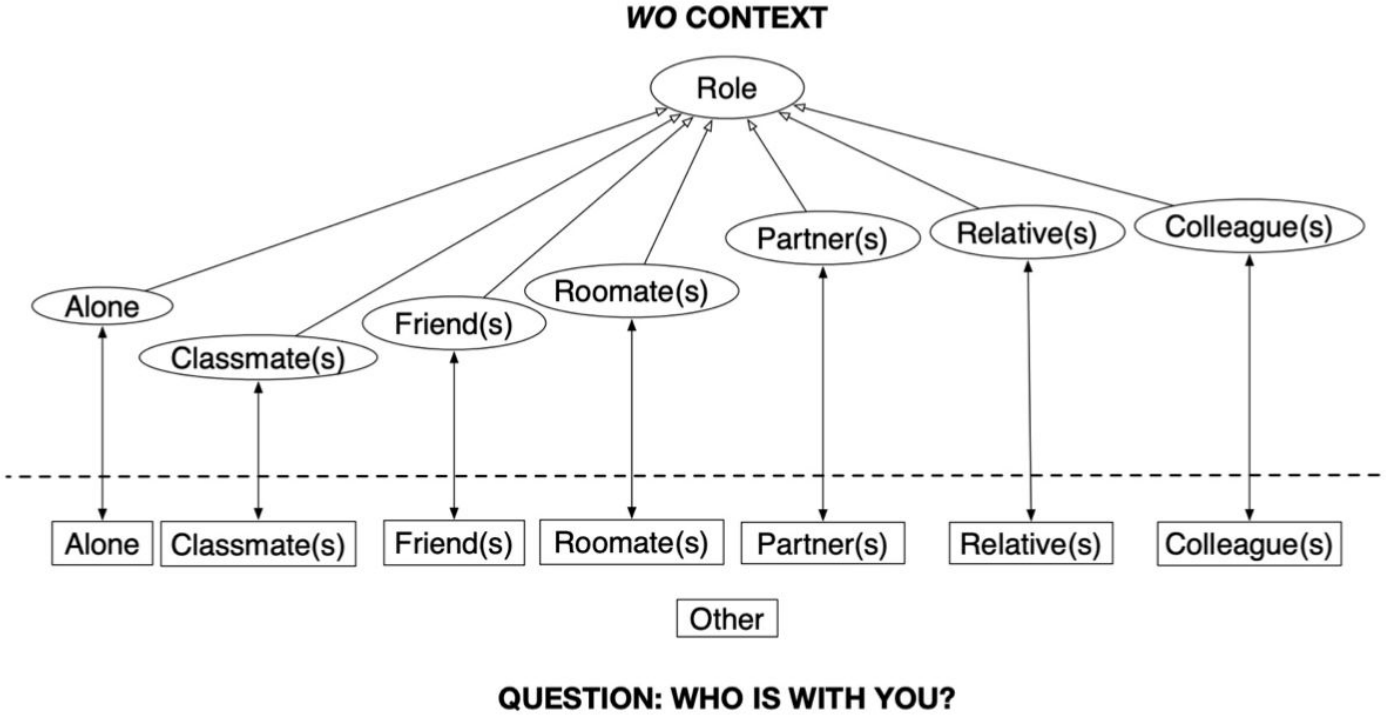



Figure 4.5: The mapping from the WO context to the social relations annotation list.

# 1. Where to collect (sensor) data from?

# Which devices can generate sensor data?

- IoT
- Dedicated devices
- Smartphones
- Wearables

- **Each of them can be used to collect a subset of data related to the user.**
- **The device convenience depends on the specific conditions of the study.**

# Dedicated devices - Movement

- Custom made devices
- They are very accurate in the measurements
- Most of the times are obtrusive and expensive to deploy at scale

# Dedicated devices - Health

- Devices built for a specific purpose, e.g., glucose monitoring
- They are perfect for their specific task and very accurate
- Expensive
- Hard to scale

# Wearables Examples

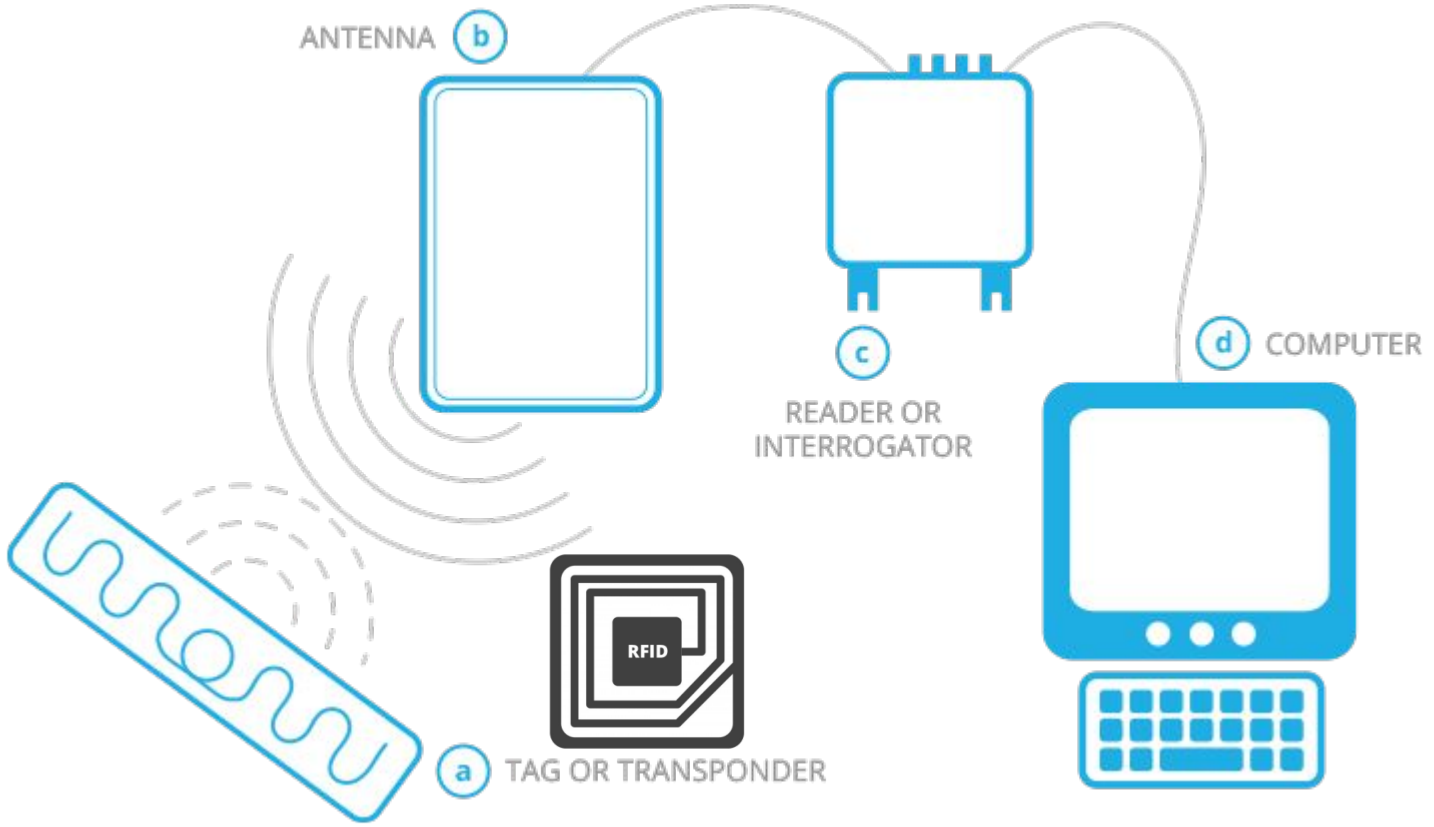

Suunto (Wear Os)


Fitbit

Apple Watch

Google Glasses

# Radio Frequency Identification (RFID)



# 2. Sensor data by Types

- Hardware
- Software
- (Externals)

# iLog Hardware Sensors List (incomplete)

| 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 |

# Hardware Sensors Description

1 (**Accelerometer**) measures the acceleration to which the phone is subjected and it captures it as a 3D vector;

2 (**Gyroscope**) measures the rotational forces to which the phone is subjected and it captures it as a 3D vector;

3 (**Light**) measures the ambient illumination around the phone, measured in illuminance (lux);

4 (**Location**) returns the geocoordinates of where the phone is located, for more accuracy this sensor combines GPS and WIFI/cellular connections;

5 (**Magnetic Field**) measures the magnetic field to which the phone is subjected and it captures it as a 3D vector;

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

# iLog Software Sensors List (incomplete)

| 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 |

# Software Sensors Description (1)

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;

8 (**Battery Charge**) returns whether the phone is currently charging its battery;

9 (**Battery Level**) returns the phone's battery level;

10 (**Bluetooth Devices**) returns all Bluetooth devices detected by the phone;

11 (**Bluetooth Low Energy**) returns all the low energy Bluetooth devices detected by the phone;

12 (**Cellular Network info**) returns information related to the cellular network (cellid, dbm, type) to which the phone is connected to;

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;

# Software Sensors Description (2)

14 (**Headset status**) returns whether the headphones of the phone where connected;

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*;

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;

17 (**Music Playback**) returns whether music is being played on the phone (yes or no) using the default music player from the operating system;

18 (**Notifications received**) measures when the phone receives a notification and when it is dismissed by the user;

# Software Sensors Description (3)

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');

20 (**Ring Mode**) returns the current ring status of the phone (normal/silent/vibrate);

21 (**Running Applications**) returns the name of the application (or application package) that is currently running in the foreground of the phone;

22 (**Screen status**) returns whether the phone's screen is on or off;

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;

# Software Sensors Description (4)

24 (**Step Detection**) similar to the previous, uses the Android API to generate a step value each time the user takes a step;

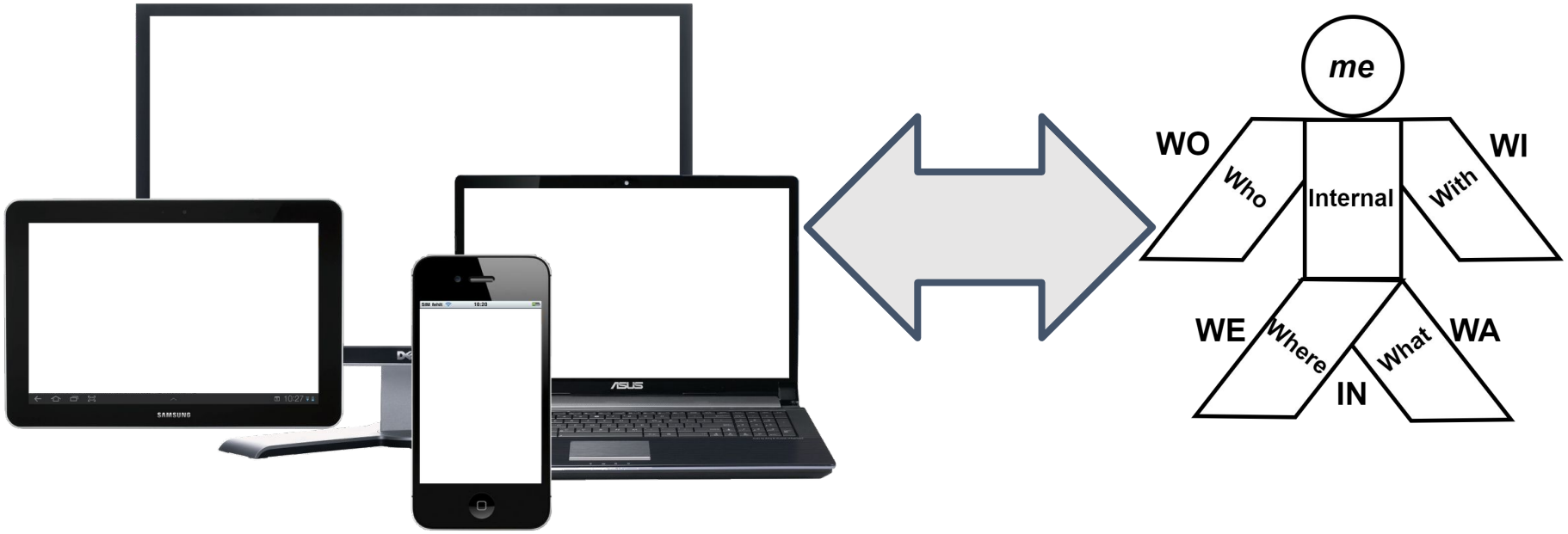
25 (**Touch event**) generates a touch value each time the user touches the screen;

26 (**User Presence**) sensor that detects when the user is present near the phone, for example when the user unlocks the screen;

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;

28 (**WIFI Networks available**) returns all WiFi networks detected by the smartphone.

# 3. Sensor data by Context



# Location (WE) in Context

- A location can be recognized via:
  - Spatial location
  - Characteristics of the environment
  - Artificial elements
  - Sound
  - etc
- Example of sensors for Location recognition are:
  - GPS (Hardware)
  - Temperature (Hardware)
  - Wi-Fi networks available (Software)
  - Cell tower (Software)
  - Microphone (Hardware)

# Event (WA) in Context

- An Event can be recognized via:
  - Movement
  - Sound
  - Smartphone usage
  - Location
  - etc
- Example of sensors for Event recognition are:
  - Accelerometer (Hardware)
  - Gyroscope (Hardware)
  - Microphone (Hardware)
  - GPS (Hardware)
  - Screen status (Software)
  - Application running on the device (Software)

# Persons (WO) in Context

- The presence of other people can be recognized via:
  - Location
  - Presence of other devices
  - Social media
  - Sound
- Important to highlights that nowadays both **hardware** and **software** presence are important

- Example of sensors for Persons recognition are:
  - GPS (Hardware)
  - Bluetooth (Hardware)
  - Wi-Fi networks (Software)
  - Cell tower (Software)
  - Tags
  - Microphone (Hardware)
  - Running application (Software)
  - Notifications (Software)

# Objects (WI) in Context

- The presence of other objects can be recognized via:
  - Wi-Fi or Bluetooth signals
  - Sound
  - etc
- Example of sensors for Objects recognition are:
  - Bluetooth (Hardware)
  - Wi-Fi networks (Software)
  - Microphone (Hardware)
  - Tags

# Internal State (IN) in Context

- The presence of the internal state is related to:
  - Heart Status.
  - Biological Signals.
  - Self Assessment.
- Example of sensors for Objects recognition are:
  - Heart Rate (Hardware).
  - Heart Rate Variability (Hardware, Software).
  - Electrodermal Activity (EDA) Sensor (Hardware).

# 4. External Sensors

1. Social Media
2. Images
3. Personal (Calendar and Contact)

# External Sensors

- There is no single solution, each type of data allows to detect different aspects
- Moreover, each field of study has its own means: sociologists, engineers, data scientists, psychologists, doctors, etc.
- Acquisition is automated

- Sociologists use time diaries
- CS (web) web data
- CS (web) social media
- CS (multimedia) use video streams
- CS (embedded systems) use sensor streams
- Doctors use health records

# Social media data

- **Social media** is any digital tool that allows users to quickly create and share content with the public.
- **Social media** encompasses a wide range of websites and apps. Some, like X (former Twitter), specialize in sharing links and short written messages.

# Social media data – Pro and cons

**Pro**:

- Global coverage
- Easy to collect
- Free
- Human contribution
- Personalization

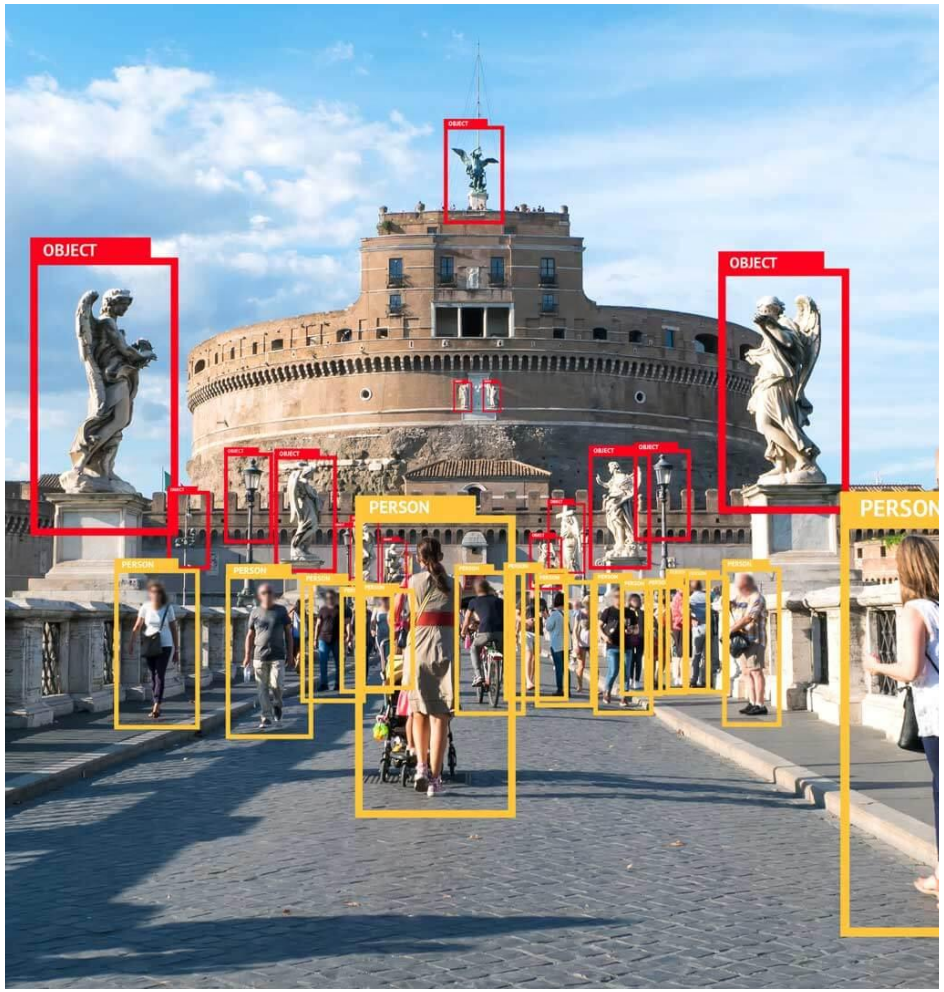
**Cons**:

- Privacy?
- Only for certain types of analysis (e.g., social relations)

* depending on the device

# Images

- Image recognition is an active field of study
- Its main purpose is object detection (e.g., faces) but it is not limited to it
- It includes Analyzing images in video streams for activity detection



# Images – Pro and cons

**Pro**:

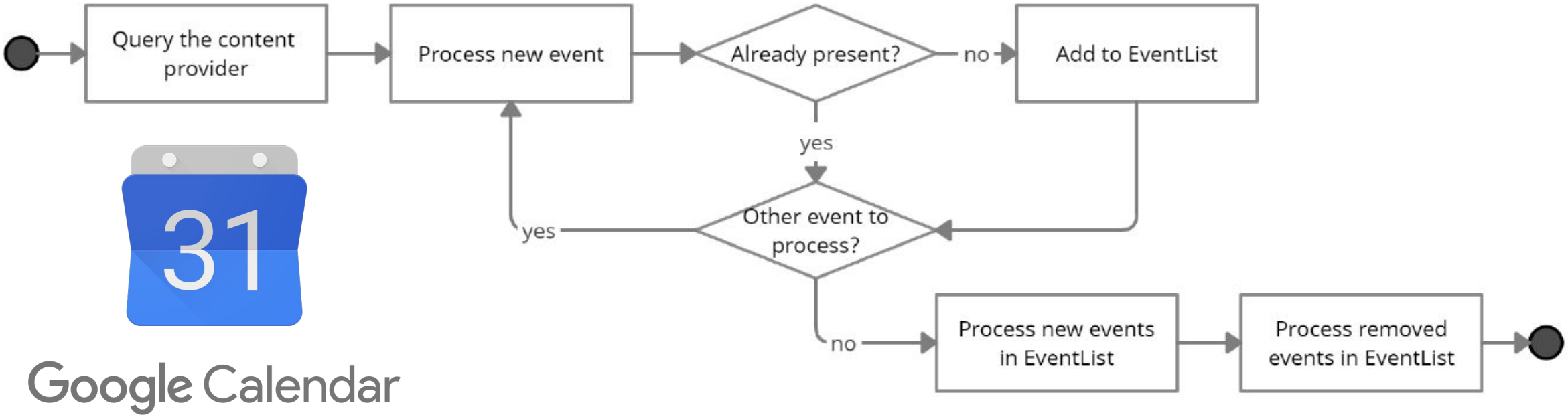
- Open world
- Easier to deploy (less devices)

**Cons**:

- Expensive (devices and computation)
- Personalization (not empowering)
- Privacy?
- Is it accurate enough?
- Where is the human contribution?

# Personal - Calendar

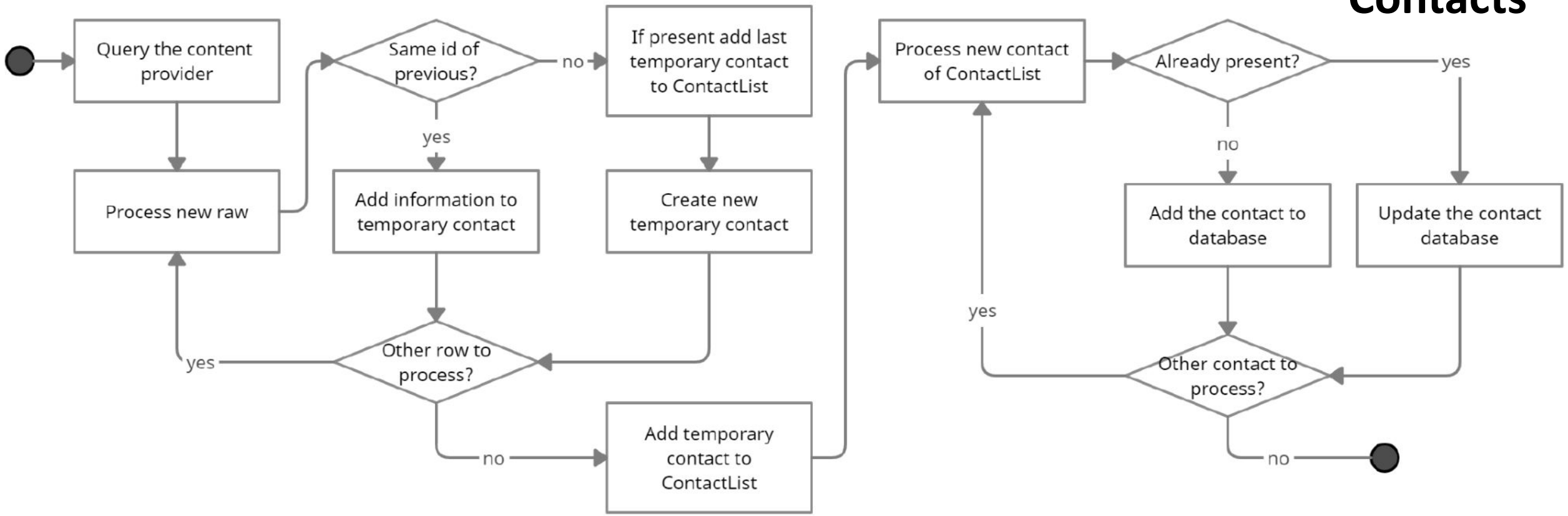
**Calendar** is the extraction of events from the Calendar.




Google Calendar

# Personal - Contacts

**Contacts** is the extraction of contact information list from the cellphone's Contacts application.

# Use Case: Student's University Life

We ran the SmartUnitn One and Two experiments in 2016 and 2018 respectively to collect time use data from students in the University of Trento. **The final goal of these experiments is to understand the how the students spend their time and how their time management ability affects their academic performance.** Instead of using traditional survey, we designed and developed an App i-Log installed in students' mobile phone to collect students' answers and collect data of sensors embedded in the phone.

Smart Unitn One — *2016*

Smart Unitn Two — *2018*

# Tools and Data Types

- Tools: i-Log
  - Data collection
  - Time diaries
- Data Types:
  - Smartphone-based data
  - Survey-based data
  - Third-party data


| What are you doing? | Where are you? | Who is with you? |
|---|---|---|
| Lesson | Class | Alone |
| Study | Study Hall | Classmate(s) |
| Eating | Library | Friend(s) |
| Personal Care | Other University place | Roomate(s) |
| En route (*) | Canteen | Partner(s) |
| Social life | Bar/ Pub/etc... | Relative(s) |
| Social media & internet | Home | Colleague(s) |
| Cultural Activity | Other Home | Other |
| Sport | Workplace | |
| Shopping | Outdoors | |
| Hobbies | Gym | |
| Other Free Time | Shop | |
| Work | Other Place | |
| Housework | (*) How are you travelling? | |
| Volunteering | By Foot | |
| Other | By Bus | |
| | By Train | |
| | By Car | |
| | By Motorbike | |
| | By Bike | |
| | Other | |

## Smart Unitn One

2016

| What are you doing? | Where are you? | Who is with you? |
|---|---|---|
| Lesson | Class | Alone |
| Study | Study Hall | Classmate(s) |
| Eating | Library | Friend(s) |
| Personal Care | Other University place | Roomate(s) |
| En route (*) | Canteen | Partner(s) |
| Social life | Bar/ Pub/etc... | Relative(s) |
| Social media & internet | Home | Colleague(s) |
| Cultural Activity | Other Home | Other |
| Sport | Workplace | |
| Shopping | Outdoors | |
| Hobbies | Gym | |
| Other Free Time | Shop | |
| Work | Other Place | |
| Housework | (*) How are you travelling? | |
| Volunteering | By Foot | |
| Other | By Bus | |
| | By Train | |
| | By Car | |
| | By Motorbike | |
| | By Bike | |
| | Other | |

Figure 3.1: The questionnaire in SmartUnitn One.

## Smart Unitn Two

2018

| What are you doing? | Where are you? | With whom are you? | What is your mood? |
|---|---|---|---|
| Sleeping | Home Apartment Room | Alone | 1. 😃 |
| Self-care | Relatives Home | Friend(s) | 2. 🙂 |
| Eating | House (friends others) | Relative(s) | 3. 😑 |
| Study | Classroom / Laboratory | Classmate(s) | 4. 😕 |
| Lesson | Classroom / Study hall | Roommate(s) | 5. 🙁 |
| Social life | University Library | Colleague(s) | |
| Watching YouTube Tv-shows etc. | Other university place | Partner | |
| Social media (Facebook Instagram etc.) | Canteen | Other | |
| Travelling (*) | Other Library | | |
| Coffee break cigarette beer etc. | Gym | (*) How are you moving? | |
| Phone calling; in chat WhatsApp | Shop supermarket | By subway | |
| Reading a book; listening to music | Pizzeria pub bar restaurant | By car | |
| Movie Theatre Concert Exhibit ... | Movie Theatre Museum | By foot | |
| Housework | Workplace | By bike | |
| Shopping | Other place | By bus | |
| Sport | Outdoors | By train | |
| Rest/nap | | By motorbike | |
| Hobbies | | Other | |
| Work | | | |

Figure 3.2: The questionnaire in SmartUnitn Two.

## Smart Unitn One

*2016*

- Two weeks
  - week #1: survey every 30 mins
  - week #2: only sensor data collected
- Participants:
  - 312 students were contacted
  - 104 students fulfilled criteria
  - 75 students took part in
  - 72 valid participants

## Smart Unitn Two

*2018*

- Four weeks
  - week 1-2: survey every 30 mins
  - week 3-4: survey every 2 hours
- Participants:
  - 12000 students were sent emails
  - 273 students registered
  - 237 students took part in
  - 184 valid participants

*Listening is not enough, we must ask. Knowing is not enough; we must apply.*

# Any Questions?