Crowdsourcing Smart Home Data

Tutor: Diogo gomes (dgomes@ua.pt)

Group Size: 4

Description: The main objective of this project is to create a platform to can gather through a crowdsourcing mechanism anonymous information from homes. The base of this project will be the Open Source Software (OSS) project Home Assistant (https://home-assistant.io) and the Portuguese community of users (http://cpha.pt).

The project will involve the development of 3 main components: a Data Lake, a Dashboard and a custom component for Home Assistant. Since we will be dealing with information that can be considered sensible, unless it is not anonymized, a great deal of importance will go towards providing informed decision making (choosing what to share), right to be forgotten and anonymization techniques.

Requirements:
- Development of a Data Lake capable to store heterogenous information (mostly JSON docs)
    - Ingest API publicly exposed
    - Query API for consumption by the Dashboard in near realtime
    - Export API for data dumping in CKAN compliant formats
- Development of a Dashboard
    - Showing information collected and health of the platform (uptime, size of the data lake, number of volunteers)
    - Should contain lots of graphics and be web-based
- Development of an Home Assistant Custom Component
    - This will follow a recorder pattern in Home Assistant and will aggregate authorized information before sending periodically to the Data Lake
    - Will required development of Custom Lovelace Card to collect an informed consent from the user
- Adherence to international standards such as CKAN

Possible challenges:
- First contact with Data Science technologies
- Involvement with an Open Source Software project (Home Assistant)
- Getting users to participate

Expected Results:
- User-friendly platform
- Gather na interesting dataset that can be used in other subjects and eventually MSc and Phd thesis.