

# Crowdsourcing Smart Home Data

Elaboration Phase

Group 5





# State of the Art

## *Home Assistant Data Science Default support*

Support is given for users to analyze their own data. This analysis is made with the help of a jupyter notebook add-on that can be added to the home assistant used by the user. To share information with third parties, information needs to be exported to a file and the file shared

Issues:

- ❑ Doesn't have native support for sharing and aggregation;
- ❑ Data isn't anonymized;

Source: <https://data.home-assistant.io/>



# State of the Art

## *Home Assistant Analytics*

Anonymous analytics of home assistant users who opted in on data sharing about their Intalls, Stats, Integrations and ADD-ONS.

Issues:

- ❑ Limited data scope,
- ❑ The scope doesn't take into account the users sensors or deeper configurations;

Source: <https://analytics.home-assistant.io/>



# State of the Art

## *UK-DALE Dataset*

Open-access dataset from the UK recording Domestic Appliance-Level Electricity.

Issues:

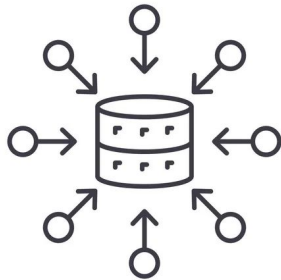
- ❑ Limited region scope, only evaluates UK;
- ❑ Limited entities scope, only five houses were evaluated;
- ❑ Lack of data visualization;
- ❑ Lack of real-time data updates;

Source: [https://ukerc.rl.ac.uk/DC/cgi-bin/edc\\_search.pl?GoButton=Detail&WantComp=41](https://ukerc.rl.ac.uk/DC/cgi-bin/edc_search.pl?GoButton=Detail&WantComp=41)



# Goals

- ❑ Gather and transmit home assistant data efficiently
- ❑ Store and make it available for easy access, deletion and analysis
- ❑ Anonymize all consenting users
- ❑ Make opting in and out easy and simple





# Actors

## Volunteer

Person providing data from their smart home to the system



## Researcher

Person who wants to extract value from gathered data





# Use Cases

## *Home Assistant Component*

### Volunteer



- ❑ Send data from their smart home
  - ❑ Choose what data to send
  - ❑ See what data is going to be sent
- ❑ Opt-out and be forgotten
  - stop sending data and delete all previous data sent
- ❑ Receive notifications on sensitive data
  - be notified when sensitive data is redacted, to build trust



# Use Cases

*Dashboard*

## Researcher



- ❏ See gathered data as graphs and other formats readily and easily available through the dashboard
- ❏ Export data in CKAN compliant format for further research and treatment





# Functional Requirements

- ❑ Allow users to formally consent to data collection
- ❑ The anonymization procedure should be on the volunteer side
- ❑ Only work with aggregated data
- ❑ Ignore/blacklist certain kind of information to send (like location data)
- ❑ Guarantee that the system complies with the GDPR, transparently
- ❑ Follow the CKAN standard for data dumping in the Export API



# Non-Functional Requirements

[HIGH]

- ❑ If there are irregularities in data upload reliability between volunteers, then fill missing info using an AI model
- ❑ The data uploading to the Ingest API won't be real-time, it will be done in a batch-processing fashion (to reduce load)
- ❑ The Query API has to provide real-time processing/transformation of data that is sent to the dashboard
- ❑ Pagination in the Query API
- ❑ All development code should be open-source, as well as the tools used
- ❑ Both the Home Assistant Aggregator and Lovelace Card should be available on the Home Assistant Community Store (HACS)

[MED]

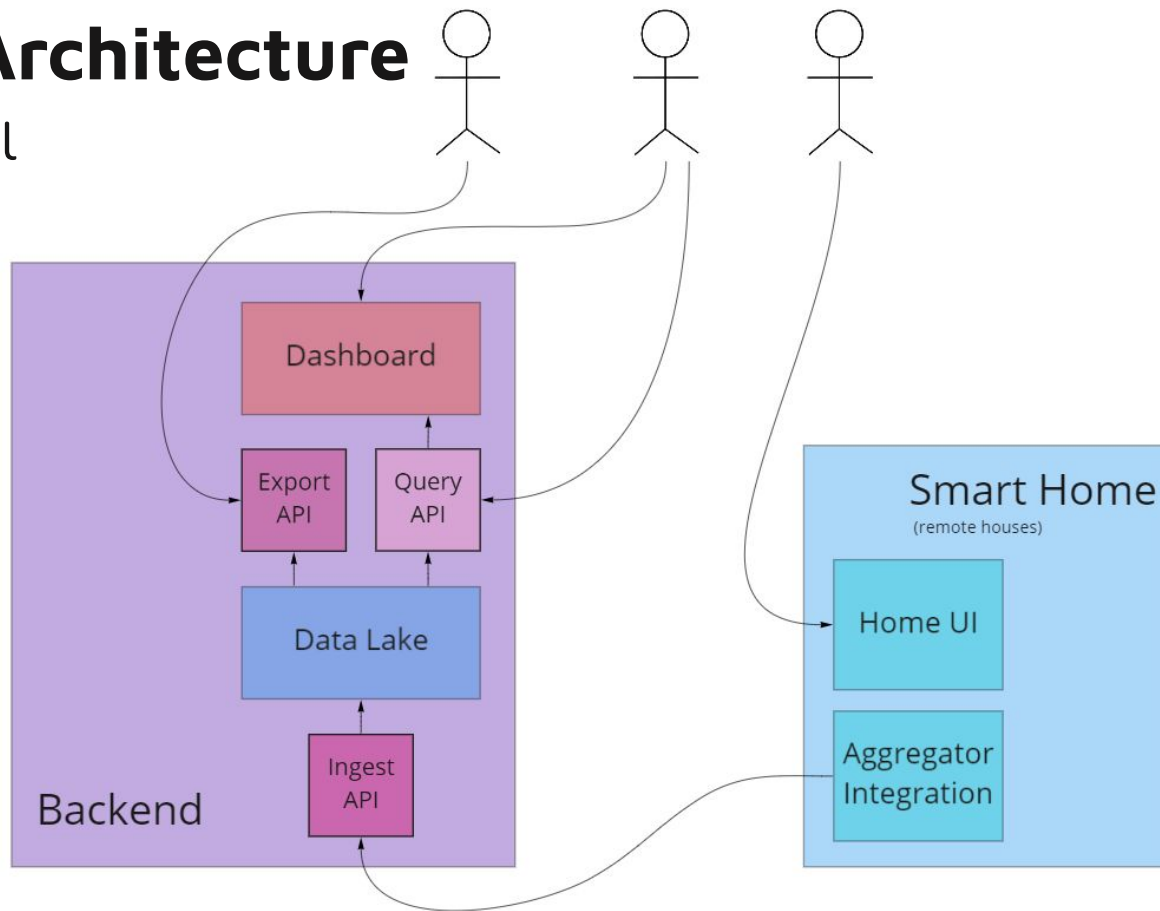
- ❑ Usage of the Home Assistant integration should be properly documented
- ❑ Prevent third parties from abusing the individual record deletion using random/specific IDs
- ❑ Protect against upload malicious/manipulated data (worms, spammers, etc.)

[LOW]



# System Architecture

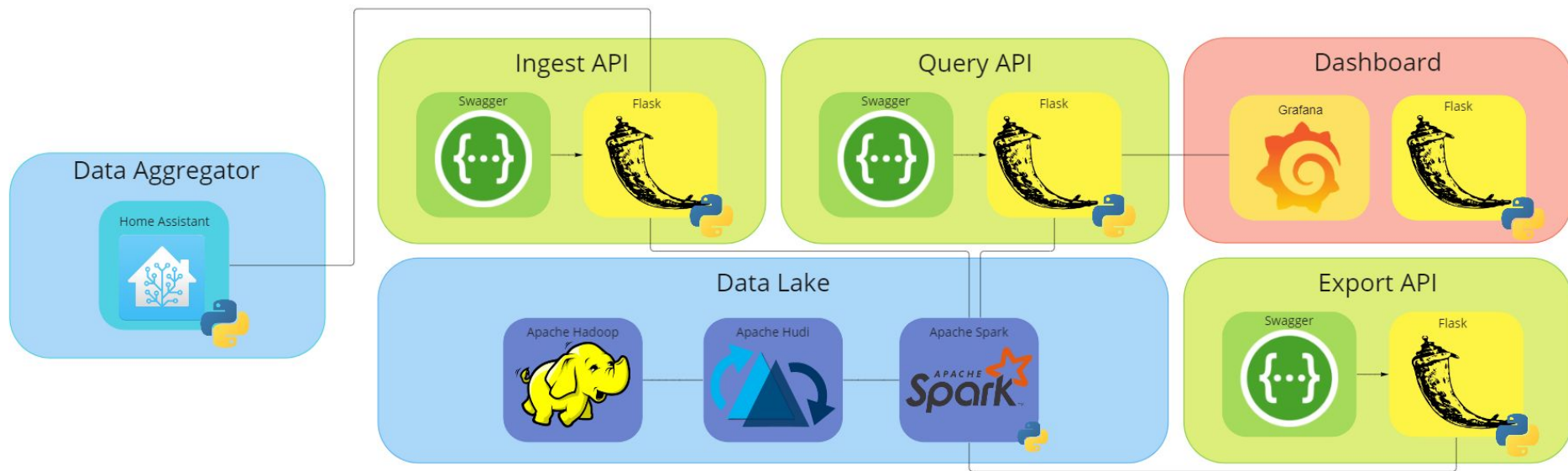
## Logical Model





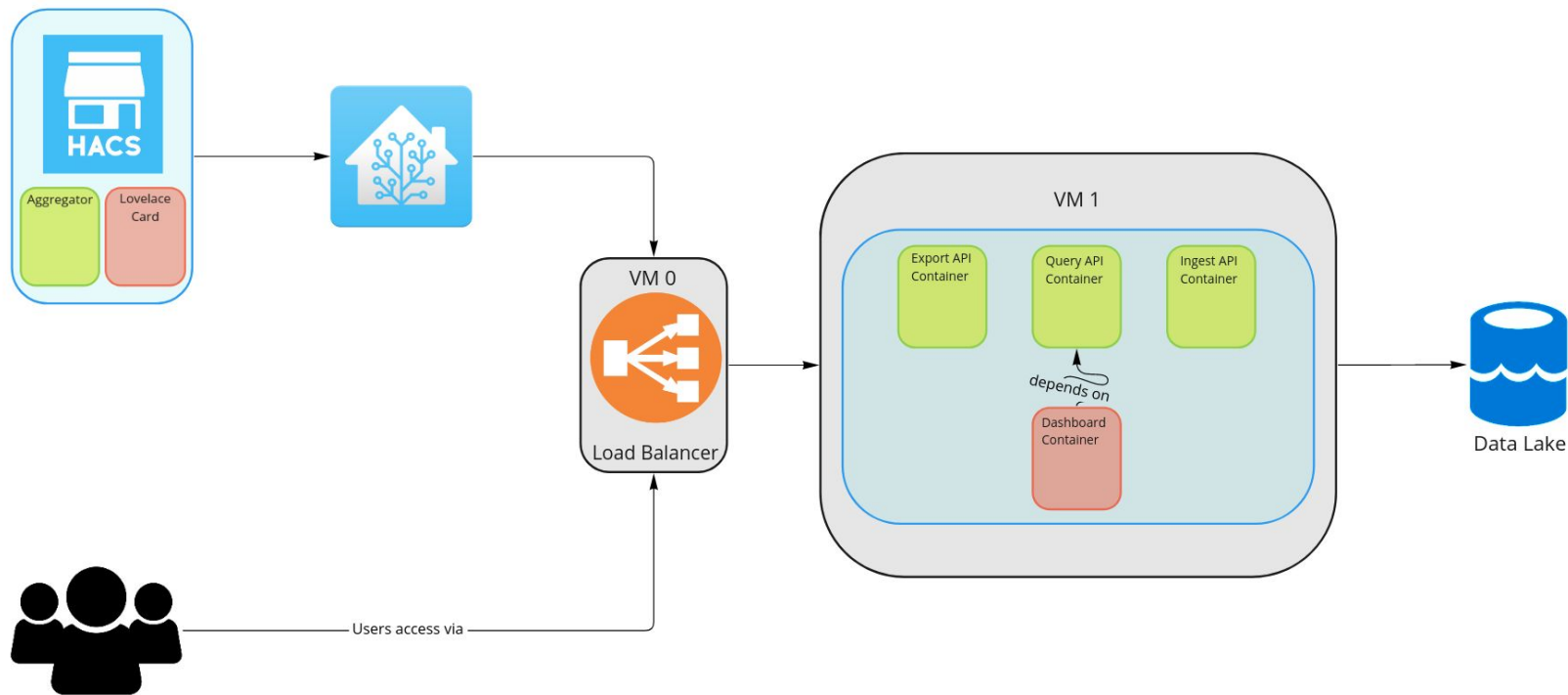
# System Architecture

## Implementation Model



# System Architecture

## Deployment Model





# Home Assistant Integration Mockup

## Main view

Data Collector

Menu

Data collection is enabled

Size of last sent package:18MB

Total size of sent packages:132MB

Date of last sent package:24/07/2022

Sending data since:05/06/2022

Manage collection settings

Delete my data

[Terms and conditions](#)



# Home Assistant Integration Mockup

Manage collection settings

Menu

## Data Collector

Menu / Collection Settings

← Back

I accept the [terms and conditions](#) and consent to anonymous data collection



Select what data you're willing to share

These are the sensors we detected in your home

> Temperature [0/3 sensors]



> Energy Consumption [0/12 sensors]



> Lighting [0/15 sensors]



> Motion [0/6 sensors]





# Home Assistant Integration Mockup

Manage collection settings

Menu

## Data Collector

Menu / Collection Settings

← Back

I accept the [terms and conditions](#), and consent to anonymous data collection



### Select what data you're willing to share

These are the sensors we detected in your home

#### ▼ Temperature [3/3 sensors]



Living Room



Kitchen



Bedroom



#### ➤ Energy Consumption [12/12 sensors]



## Data Collector

Menu / Collection Settings

← Back

I accept the [terms and conditions](#), and consent to anonymous data collection



### Select what data you're willing to share

These are the sensors we detected in your home

#### ▼ Temperature [2/3 sensors]



Living Room



Kitchen



Bedroom



#### ➤ Energy Consumption [12/12 sensors]







# Home Assistant Integration Mockup

Delete my data

Menu

## Data Collector

Menu / Delete Data

← Back

If you proceed, all your sent data will be deleted from our storage.

Continue?

Yes, delete my data

Cancel

**Thank you**