

## Data Collection Overview

This document offers a comprehensive overview of the sensory data gathered through the Air Quality monitors and sensing devices deployed within the meeting rooms. Additionally, it gives a detailed description of the data collection process and data retention periods.

## Project information

**Principle investigator:** BSc Danny de Vries (14495643)

**Contact Email:** danny.de.vries@student.uva.nl

**Project:** Master Thesis (IS)

**University:** University of Amsterdam (UvA)

**Master:** Information Studies Information Systems (track)

**Institute:** Informatics Institute

**Faculty:** Faculty of Science (FNWI)

**Research Group:** Digital Interactions Lab (DIL)

**Supervisor(s):** Dr. H. Seied Alavi PhD & S. Rao Ph.D. Candidate

## Description of the project

### 0.0.1 Title of the project

A breath of fresh air: towards optimal Indoor Air Quality (IAQ)

### 0.0.2 Aim of the research

This master thesis investigates the integration of sensory measurements of indoor air quality (IAQ) into specific indoor spaces to raise occupants' awareness and help occupants take preventive measures against poor air quality. Conducted at Lab42 the study employs human-building interaction principles and persuasive technology to visualize IAQ data. Key research questions address methods for collecting accurate IAQ measurements, integrating environmental data into visual representations, and evaluating the impact of physical visualizations on occupants' understanding and behavior. The overall aim is to inform the development of effective interventions, contribute to healthier indoor environments, and inform building design decisions.

### 0.0.3 Location of project execution

Meeting rooms within the Lab 42 building at the UvA Science Park Campus

### 0.0.4 Expected research duration

1th of February 2024 until 30th of June 2024

### 0.0.5 Names of organisations involved

Digital Interactions Lab (DIL)

### 0.0.6 Principle investigator and Supervisor

D. de Vries (danny.de.vries@student.uva.nl) and H. Seied Alavi (h.alavi@uva.nl)

# Data collection

## 0.0.7 What data is being collected and generated?

There are 'two' components during the research that will gather data in some form (from sensors or occupants):

1) Measuring Indoor Air Quality (dataloggers): Analyzing current space with regards to air quality by utilizing the data measured from the automated systems already installed within the building exposed by the building API and augmenting this dataset with the aforementioned commercially available indoor climate data loggers. .

2) Occupant behavior (sensing devices): Collecting information about occupants within the specific spaces within the building. Surveys are anticipated to be the primary data collection method distributed across the course of the thesis projects. Additionally, conducting one-on-one interviews using open-field questionnaires to explore occupants' comfort levels more comprehensively will also be used.

## 0.0.8 Technical set-up

Two commercially available data loggers (air quality monitors) are installed within the meetings rooms, calibrated as specified by the manufacturers and installed in the position as specified by the manufacturers. One Atal Atu-CT indoor climate datalogger (grey housing, green front, lcd display) and a AirCheq Touch Aero (black housing, coloured oled display). Both are installed within the meetings rooms using a mounting frame with a lock attached. On the windows Aqara smart window sensors are installed to gather information about when the windows are in an open or closed state.

## 0.0.9 Scope of data collected

No data that can be considered personal is collected. Only environmental properties about the building and information about the number of occupants within a meeting room are gathered.

- Carbon Dioxide (CO2 ppm) measurements
- Particulate Matter (PM2.5 and PM1.0) measurements
- Humidity measurements
- Temperature measurements
- VOCs measurements
- Open/close state of window sensor
- Number of occupants within the meeting room

Sensor devices and prototypes will store data locally (no 'over-the-air' data transmission) and will be collected weekly. After collection, the memory and storage of these devices will be wiped (reformatted). During analysis, the data will be locally stored on encrypted laptops (e.g. password-protected Macbooks with FileVault enabled).

Data such as full name, contact information, identification numbers, date of birth, health information, biometric data, financial information, demographic, and employment information will not be collected during the research.

## 0.0.10 Does the remixing/mashing of data enable the identification of individuals?

Not to the knowledge of the researchers. As stated above, no personal data is being collected and also no location-specific data such as GPS coordinates or IP addresses are collected.

## 0.0.11 How much data do you plan to collect?

Text files (.pdfs of surveys, interviews, and field observation), Spreadsheets (.csv of surveys and interviews), Images of the building, prototypes, and sensing devices (.jpeg, rasterized images) Sensory data (.json objects, timestamped.txt files). In terms of actual file size somewhere between 0 - 2 GB.

#### **0.0.12 Where will the data be stored during the research project?**

Data will be permanently stored on cloud service providers provided by the University of Amsterdam (such as Research Drive and Surfdrive) as outlined in the Research Data Management guidelines <sup>1</sup>.

#### **0.0.13 Who will have access to the data?**

Only researchers in possession of the encryption methods and passwords will be able to retrieve the data from the sensors. Data from the user studies from the cloud storage will only be accessible to the researcher and supervisor(s).

#### **0.0.14 How will the data be destroyed?**

The data will be removed, after collection of the data, by formatting the storage of the sensory devices. The user data will remain in the cloud storage while the thesis research is ongoing. The data will be accessible for a certain period after the study has been concluded and then deleted after graduation.

#### **0.0.15 What data will be published as outputs from the project and when?**

Upon completion of the master thesis aggregation of the analysis will be published in the results section of the thesis paper. Notebooks of the analysis performed will be published in a public GitHub repository as well as the source code of the sensory devices and prototype. The data used for these analyses will not be published on GitHub and will be kept private in the cloud storage and on local devices.

#### **0.0.16 How long will the data exist in the repository?**

The GitHub repository of the output analysis will remain published and public during the research and will remain on GitHub after completion of the master thesis project. Contact information will be supplied in the documentation (e.g. readme.md) on how to reach out to the researchers for possible requests for code removal.

#### **0.0.17 What consent is needed for subsequent data use?**

The outputs of the analysis and the code of the prototype are hosted on GitHub and subsequent use or references of these analyses must adhere to the licenses within the repository. If the input data for future research needs to be acquired a request can be made to the researchers and will be reviewed for approval.

---

<sup>1</sup><https://rdm.uva.nl/en/looking-after/storage/storage.html>