Data preparation and contribution instructions for the indoor air quality database

First of all, thank you for your time and effort in submitting your data. Collectively, we are creating the first global pilot IAQ database that will allow for answering crucial research questions towards a better and more sustainable built environment. This document provides instructions on how to prepare metadata and IAQ data.

1. Instructions on submitting metadata

The aim of collecting metadata is to label the actual IAQ data with relevant information related to buildings, spaces, occupancy, and instruments, so the end user can filter and analyze the data according to their interest. Also, information related to your study will be optionally collected to facilitate attribution. Metadata are classified into three levels of importance.

Mandatory: metadata that are crucial for creating a structured database, without which data submission will not be accepted.
Recommended: metadata that are highly useful for end users exploring the database.
Optional: metadata that are useful but difficult to collect, or metadata purely for acknowledging data contributors.

1.1 Two approaches to provide metadata

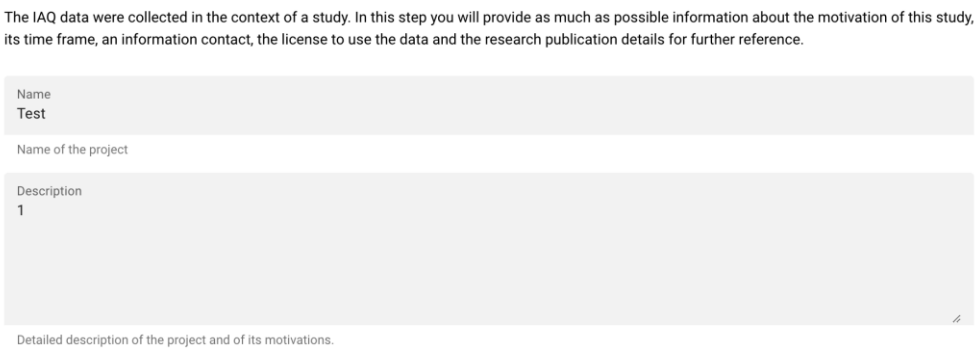
1.1.1 Using the online data entry form

The recommended approach is to fill in metadata using the web tool:
https://iaqdb-dev.epfl.ch/#/contribute

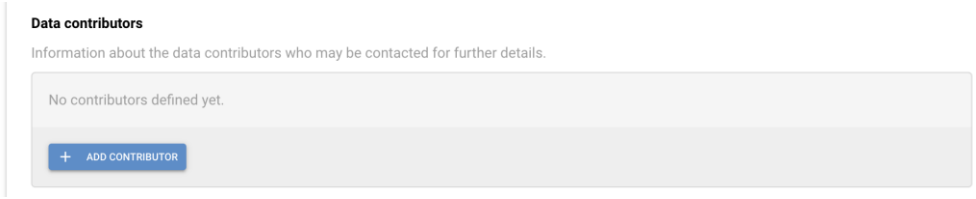
The tool contains three main pages. It will always save your input locally on your browser until you clear your caches.

1) Study information. Here, mandatory information includes name of your study and at least one contributor with a valid email address, for the database management team to contact you in case further information is needed. You can choose whether to disclose your email to the end users.

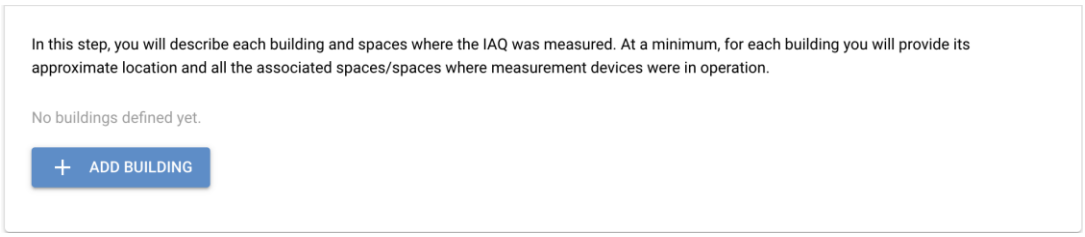
   Fill in description of your study for visibility and impacts.



   Click “Add Contributor” to add name and email address. Up to two contributors can be added per study.

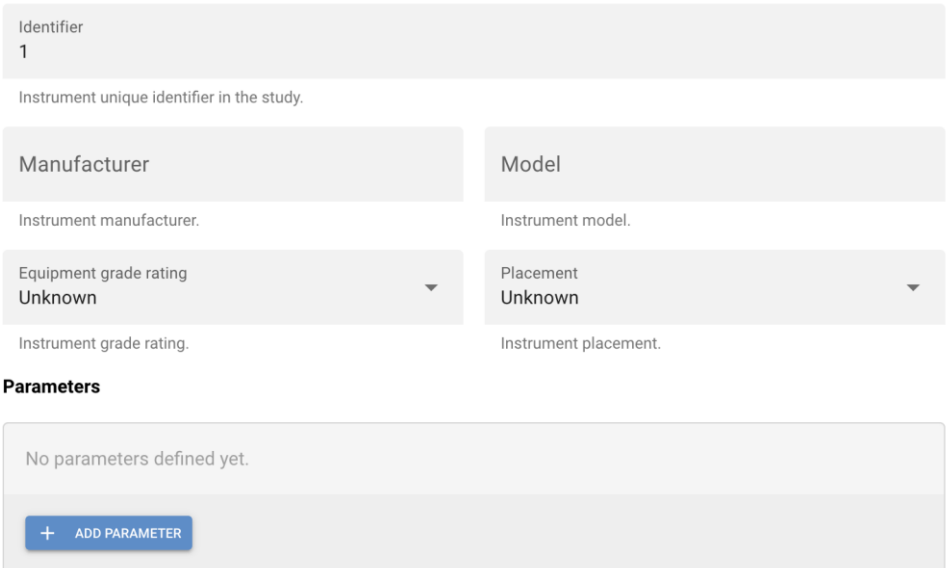


2) Buildings and spaces information.
Please start from “Add Building”. Mandatory information includes identifier (must be unique within your study, for associating actual IAQ data), Country, City, and Building type. It is also possible to duplicate information for similar buildings by clicking the icon.



Within a building, you may add multiple spaces. Adject outdoor space or personalized monitoring is also considered a “space” here. However, only add them if there is an adjacent indoor space where IAQ data are available. Mandatory information includes identifier (must be unique within the building, for linking with actual IAQ data) and space type. It is also possible to duplicate information for similar spaces by clicking the icon.

3) Instrument information. Please add instruments used in the study and provide information about detection/analysis method and uncertainty. All fields are optional but highly recommended as provide information about reliability of measurement for the end users.

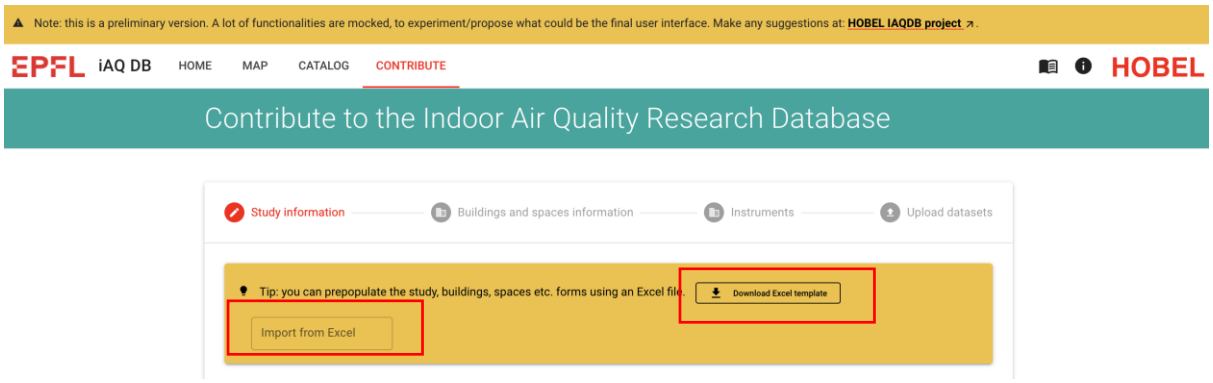


4) Afterward, please proceed to upload your actual IAQ data.

1.1.2 Using the spreadsheet template

For larger studies consisting of data from many buildings, it is also possible to fill in information in the provided excel spreadsheet and the upload it to the web tool. Information will be automatically recognized and filled in corresponding fields.

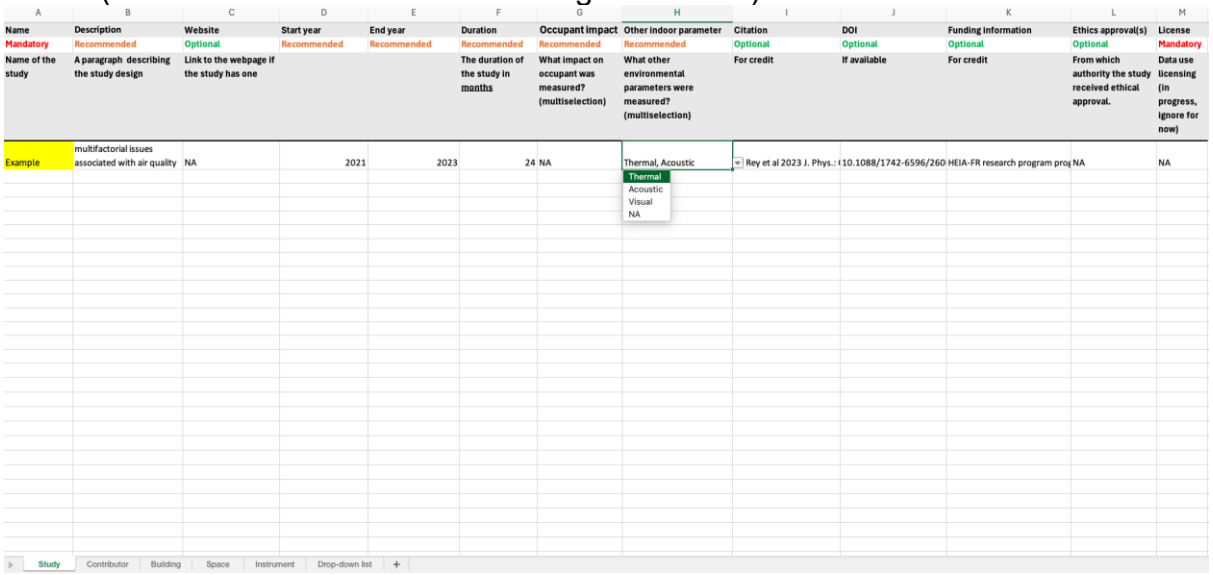
<u>The template uses macro functions so please choose to enable macros when opening the file.</u>



An example is provided to illustrate the expected input format. Please **<u>delete</u>** the existing rows when filling in the sheets. **<u>Please do not add or delete columns or change their names.</u>**

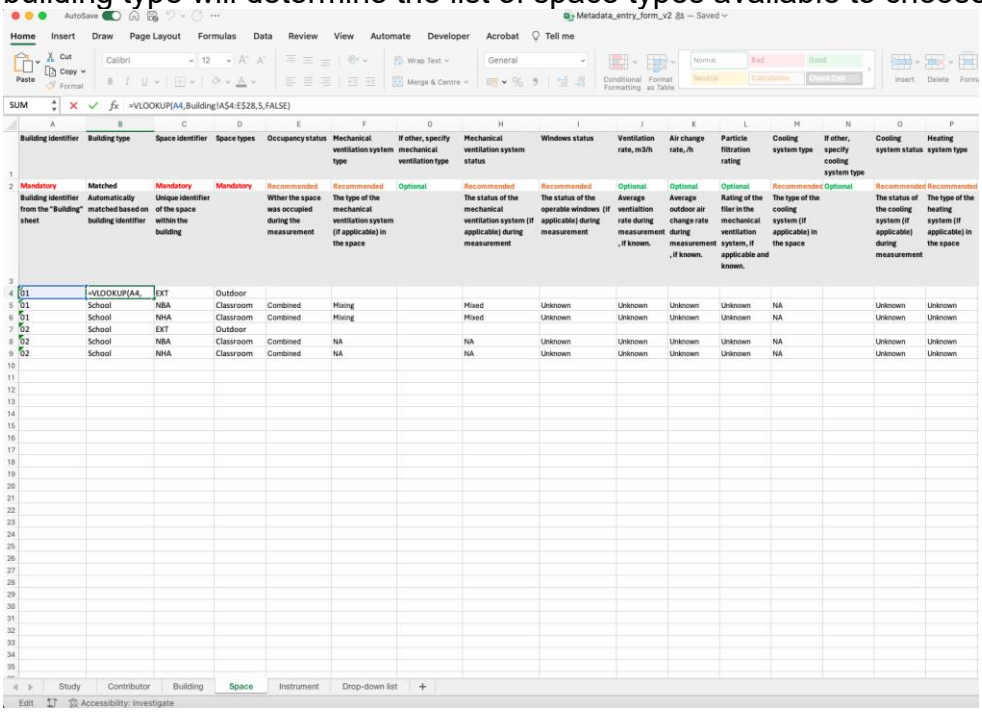
Please fill in the first five sheets **<u>in order</u>**, ranging from Study to Instrument.
Special notes:

1) Under sheet Study, the form asks for whether the study investigated any indoor environmental impact on occupants (e.g., comfort and health) and any other environmental parameters. These two questions allow for multiple selection. However, for removing an answer, you will need to clear the cell and start again (the cell doesn't allow for removing an answer).



2) Under sheet Contributor, up to two names can be provided. Please provide at least one email address for us to contact you in case of incomplete data. You can indicate whether you hope to have your email address available to end users or not here.

3) Under sheet Building, please assign a unique identifier to each of the buildings in the study, which can be string or numeric.
4) Under sheet Space, please carefully check to which building each of the spaces belong. Then, please assign a unique identifier to each of the spaces within the building, which can be string or numeric. Then, please match the building type according to building identifier using the provided “VLOOKUP” function. The building type will determine the list of space types available to choose from.



5) Under sheet Instrument, please try your best to provide complete information about the sensors used for measurements.

After you upload the spreadsheet template, the web tool will capture the relevant metadata. You will be able to upload actual data after confirming the metadata are complete and correct, before proceeding to upload actual IAQ data.

2. Instructions on standardizing IAQ data

To facilitate data collection and database construction, we use the true long format to store data, including four columns of data (Timestamp, Parameter, Value, and Unit) and three columns of identifiers (Building_ID, Space_ID, and Instrument_ID, with the first two being mandatory). An example is given below.

Timestamp	Parameter	Value	Unit	Building_ID	Space_ID	Instrument_ID
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.5	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.3	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2
18.03.2018 13:22	Ozone	35.4	ppb	1	OA	Dasibi 2

Note: for time-integrated measurement such as VOC concentration from samplers, please use the end of measurement as timestamp.

We strongly suggest that you standardize your IAQ data according to the provided long format before uploading the data file, which significantly facilitates the scalability of the database. However, we will accept other data formats as long as such information is complete and easily identifiable.

Contact information

Dr. Bowen Du
Human-Oriented Built Environment Lab (HOBEL)
School of Architecture, Civil and Environmental Engineering
École polytechnique fédérale de Lausanne (EPFL)
bowen.du@epfl.ch