ClairCity data portal
http://claircitydata.cbs.nl

User Guide

Version 20161206

Introduction

The goal of the ClairCity data portal is to organize data produced and consumed by the ClairCity organizations within the ClairCity project. This document is a brief guide for getting started. Additional guides explaining more advanced topics may be added later.

The data portal in a nutshell

The ClairCity data portal organizes data in so called *Datasets*. Each dataset has some standard metadata fields such as a *title*, *description*, *tags*, *visibility*, *quality aspects* and *license*. Each dataset is owned by exactly one of the 16 ClairCity organizations or by the ClairCity organization as a whole.

Datasets in the ClairCity data portal have some ClairCity specific metadata fields. Each dataset may be added to one or more *Groups* from the ClairCity *conceptual model* (see next page), such as Agriculture, Transport or Health and may be assigned to one of the ClairCity *Cities* (Liguria, Aveiro, Ljubljana, Amsterdam, Bristol).

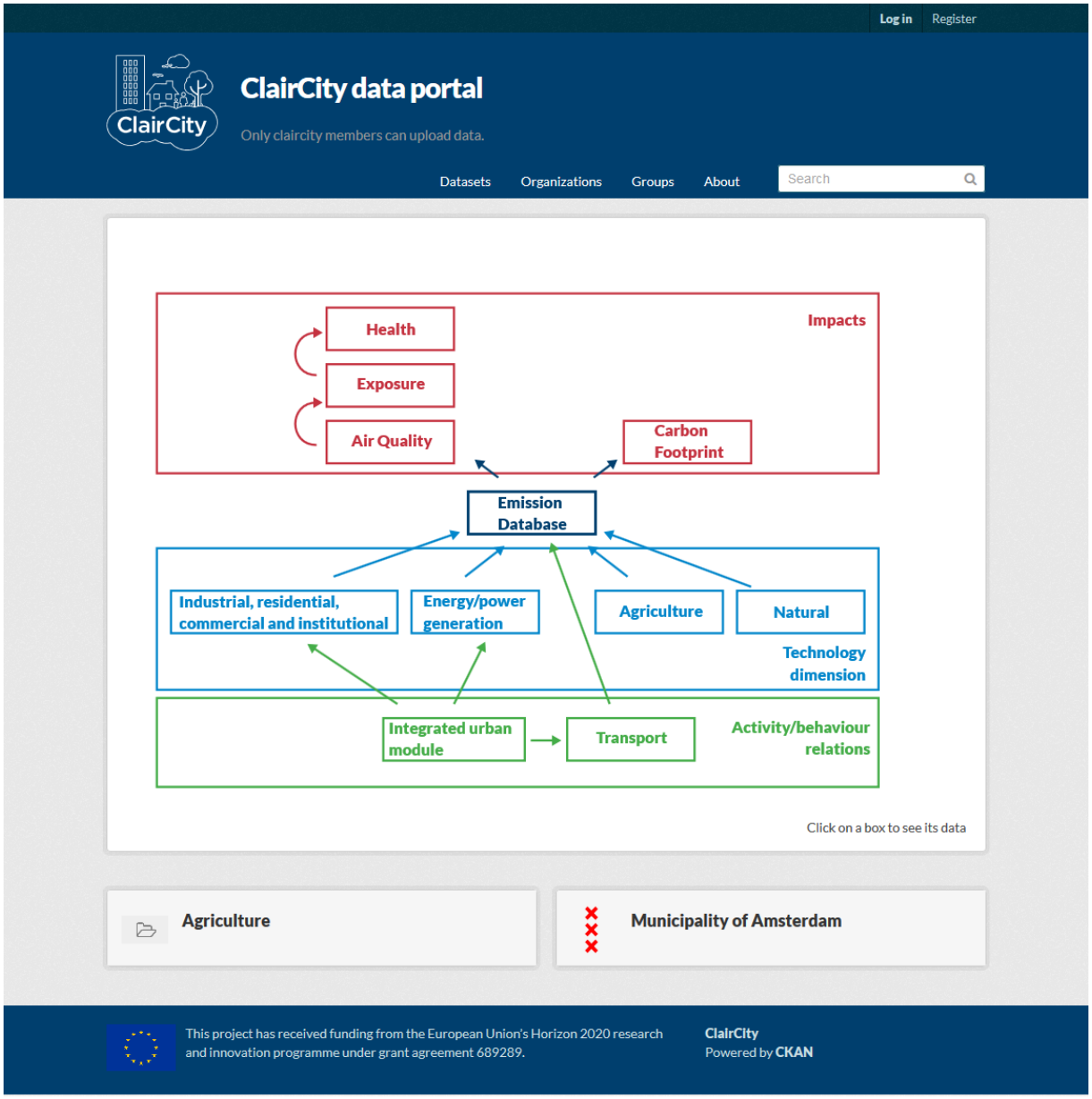
Only members of the ClairCity project are allowed to add datasets. They can effectively choose from three levels of access control: *public*, *private* to the Claircity organization and *private* to the owner organization. More details about this mechanism can be found in the section “Uploading a dataset”.

The data portal offers a data preview in a datagrid, chart or map. This works if the data was recognized (csv, json or xml) and contains the right type of variables for the preview type (like lat/lon coordinates or a geojson field for maps). Other preview facilities may be added later. If the portal cannot recognize or parse the data, it is stored as a blob for download. We advise to use one of recognizable formats as much as possible. If you have ideas or suggestions for new ClairCity specific previews for your data, please contact us.

It is not allowed to add any privacy sensitive data to the data portal. It is the responsibility of the user to verify that a dataset does not contain any privacy sensitive data and that it does not infringe intellectual property rights.

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement 689289.

Below is a screenshot of the front page of the data portal as it is displayed for non registered users. The main menu (Datasets, Organizations, Groups, About) and the search field is always present. Right under the main menu there is the ClairCity conceptual model. Next to the conceptual model a selection of data groups and / or organisations is displayed depending on the public data available in the portal.

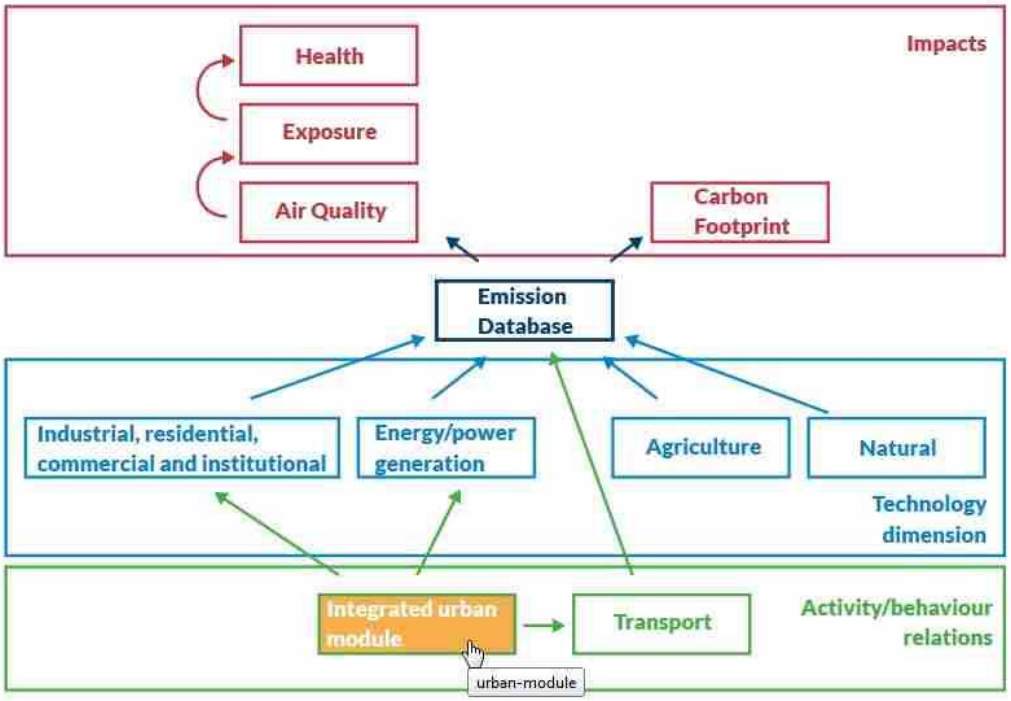


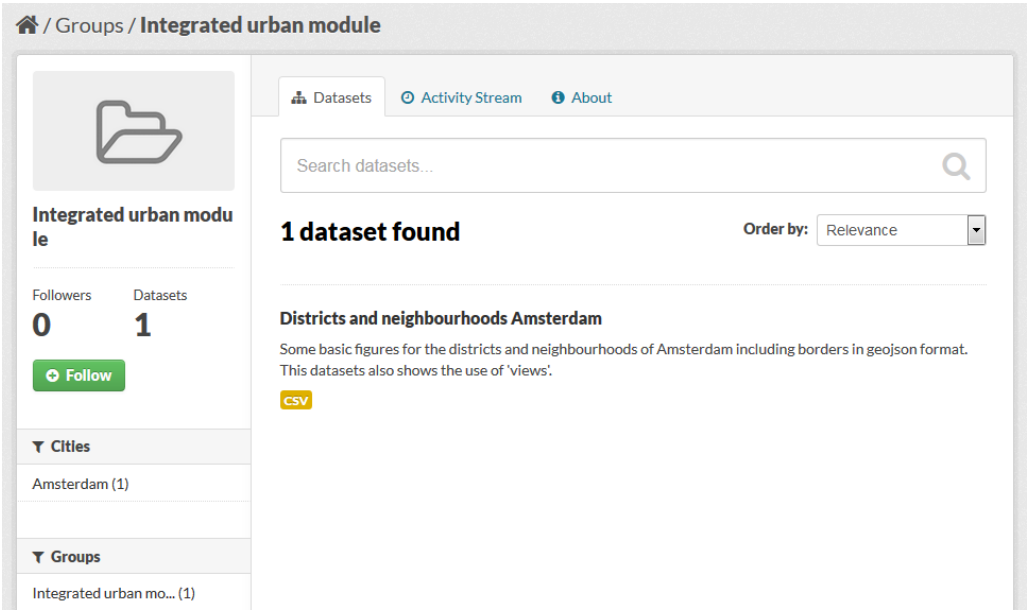
Searching for data

The data portal offers various mechanisms to search for data:

Using the ClairCity model

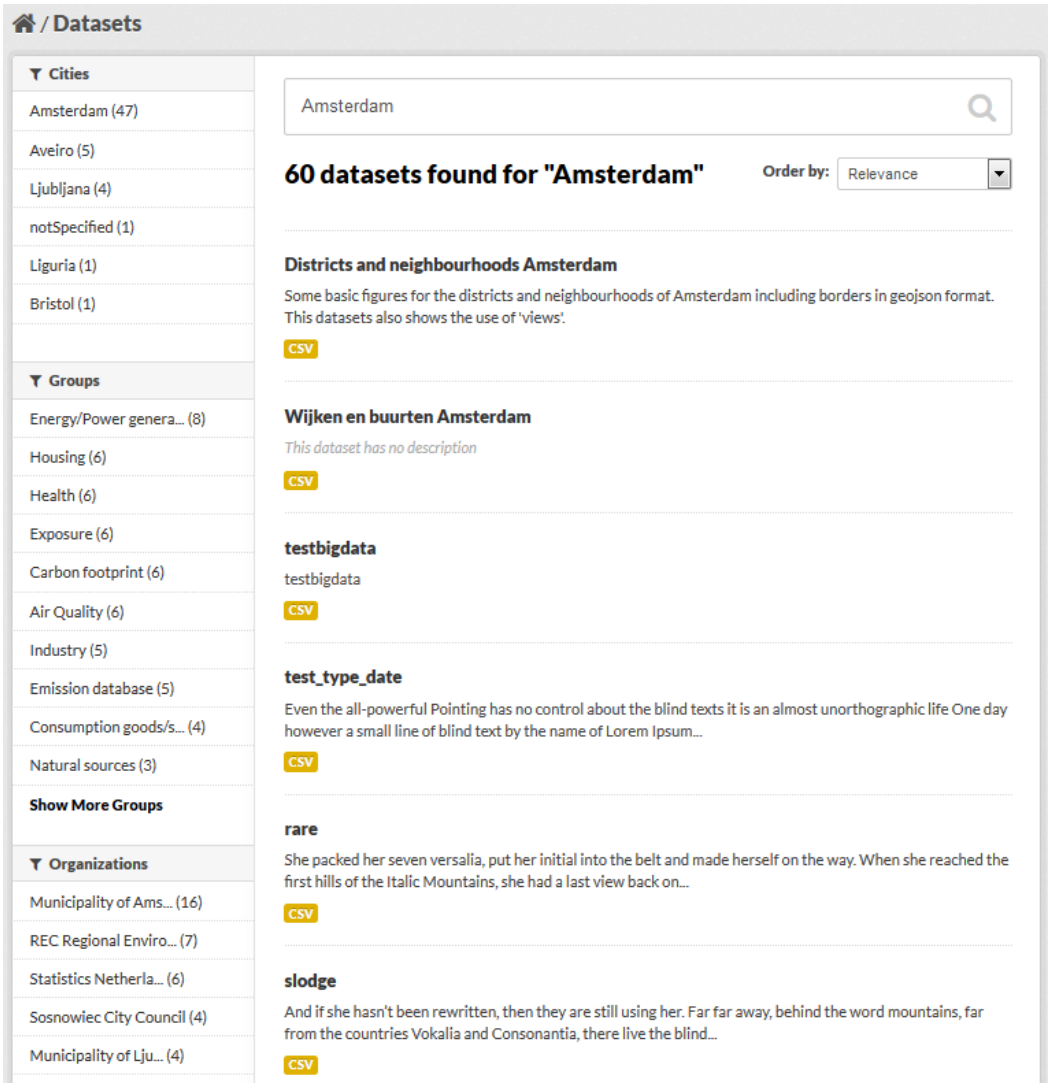
In the conceptuel model from the ClairCity project, displayed on the home page of the ClairCity data portal each box represents a group of datasets. A click on a box brings the user directly on a page displaying all datasets of that group visible for that user. The screenshots below the selection of the datatypes for the “Integrated Urban module”:





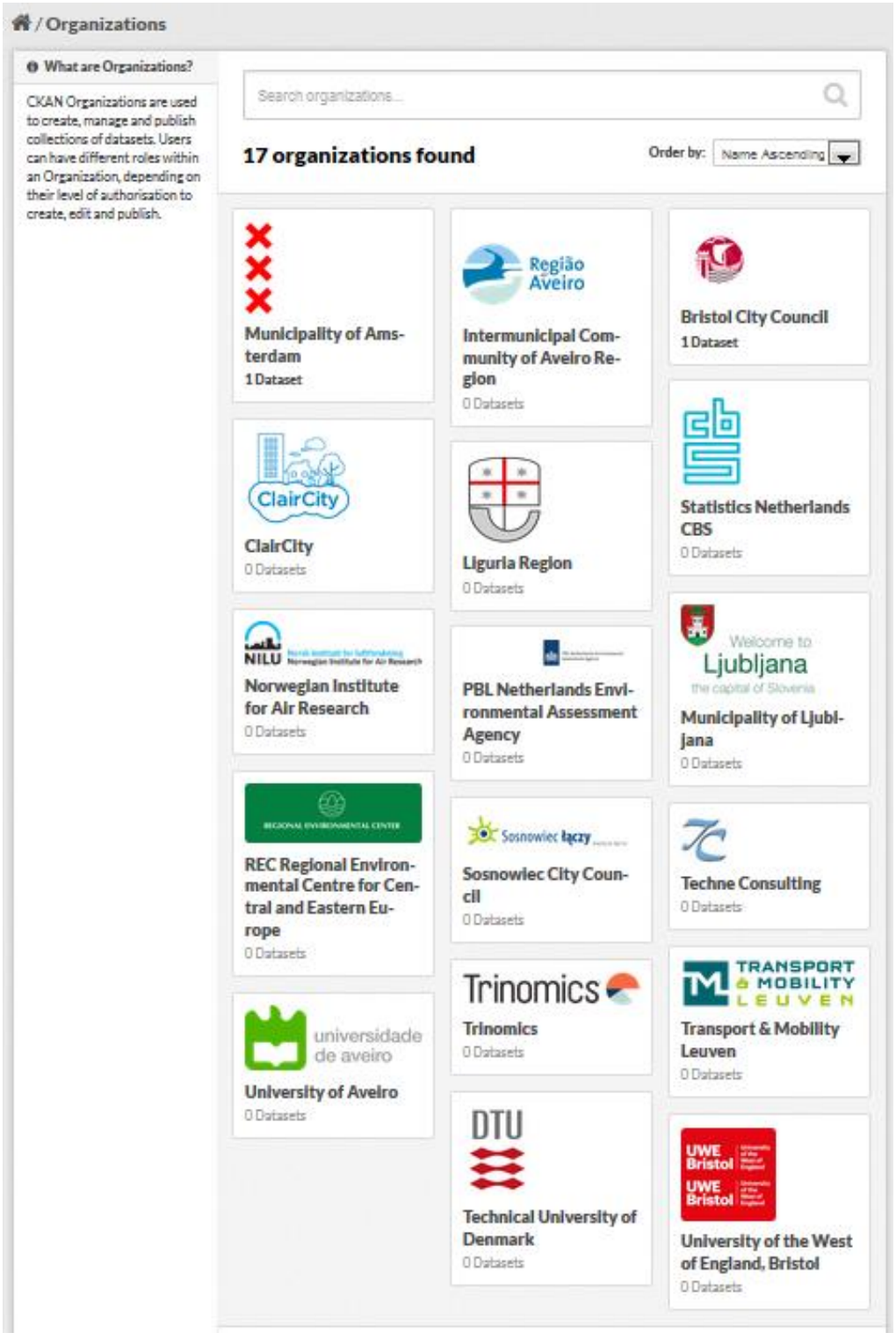
Using the Search engine

The ClairCity data portal contains a seach engine. Using the search box or clicking the “Datasets” tab gives a search results page with facets for cities, groups, organizations, tags and formats. The screenshot below shows such search results page for the search phrase “Amsterdam”:



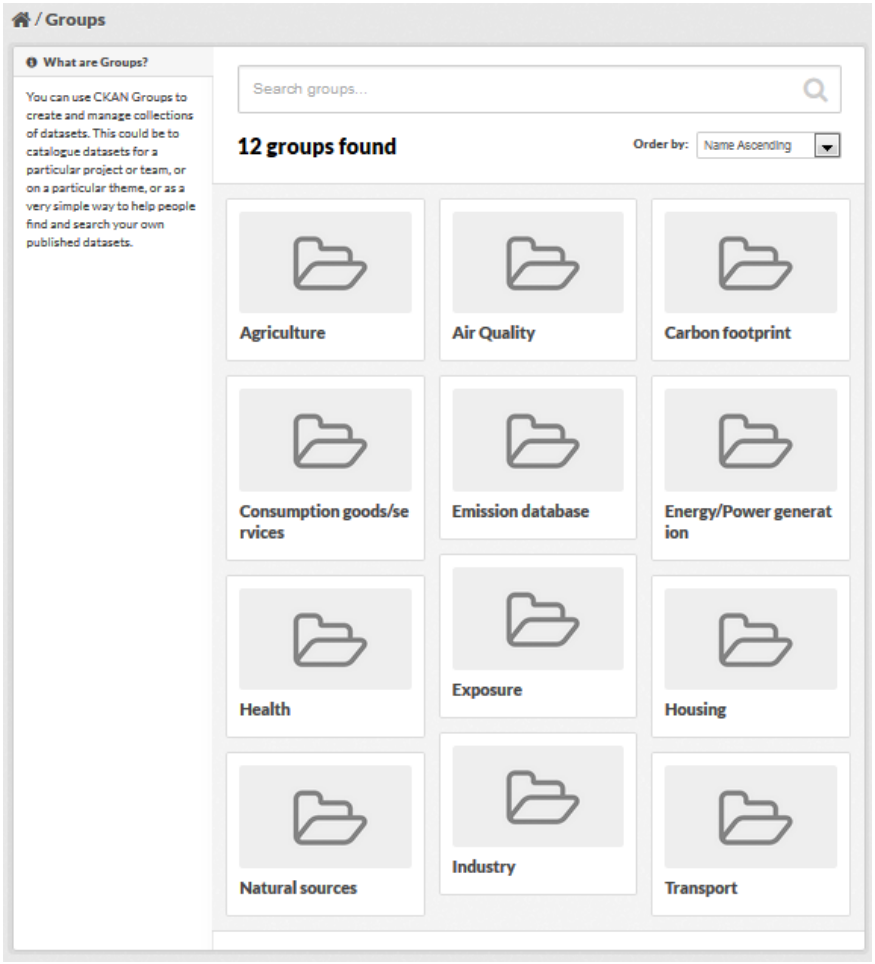
Using the list of organizations

The “Organizations” tab brings you to the page which lists all ClairCity organizations and their respective number of datasets. A click on an organization brings you to all datasets owned by that organization. Here is a screenshot:



Using the list of groups

The “Groups” tab brings you to the page which lists all groups of datasets from the conceptual model. A click on an group brings you to all datasets for that group. This is the same functionality as clicking on a group box on the conceptual model on the home page, which might be a faster and more intuitive way of selecting a group. For completeness, here is a screendump:



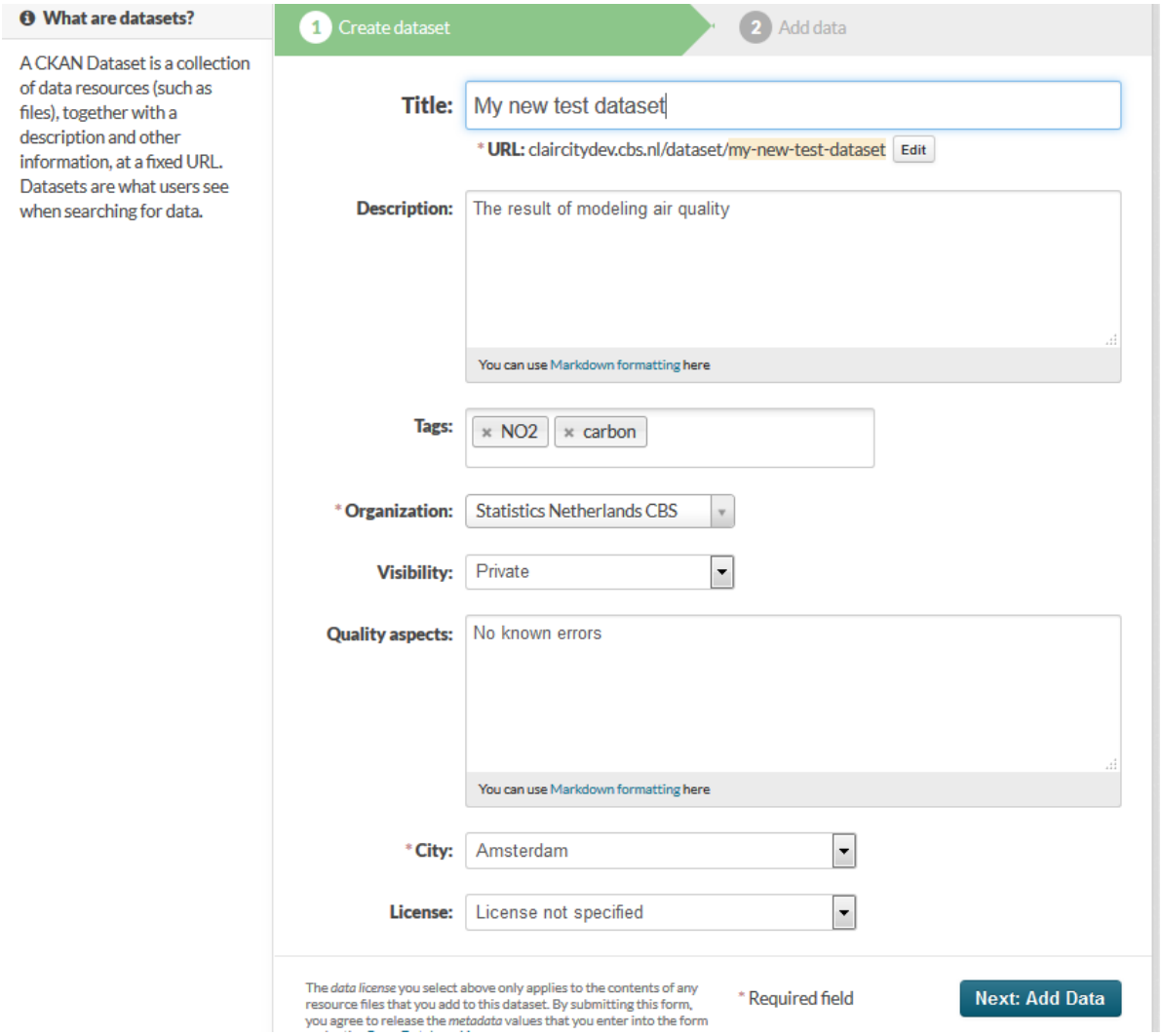
Uploading a dataset

Registered users assigned to one of the ClairCity organizations have an “Add Dataset” button on the datasets page. This brings you to the “Create dataset” form where you specify the metadata of your dataset. The fields are:

- **Title**
- **Description**: short descriptive text
- **Tags**: you can create your own
- **Organization**: either your own organization or the ClairCity organization.
- **Visibility**: Public or Private. Public makes your dataset visible to the world, Private will make your dataset visible to the organization you specified above.
- **Quality aspects**
- **City**: the city the data applies to (use not specified if it is not specific to one city)
- **License**

Note: Choosing the right combination of the Organization and Visibility fields you can choose between three levels of access control as specified in this table:

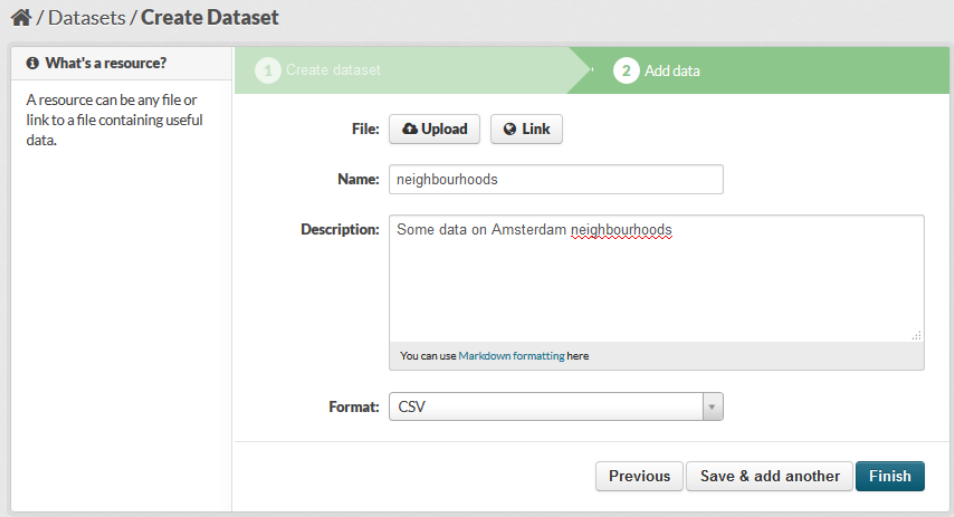
Organization	**Visibility**	
<your organization>	Public	Open to the world
<your organization>	Private	Visible to the members or your organization
ClairCity	Public	Open to the world
ClairCity	Private	Visible to all members of ClairCity



After creating your dataset, the portal invites you to add data resources to this dataset in the “Add data” form. The fields are:

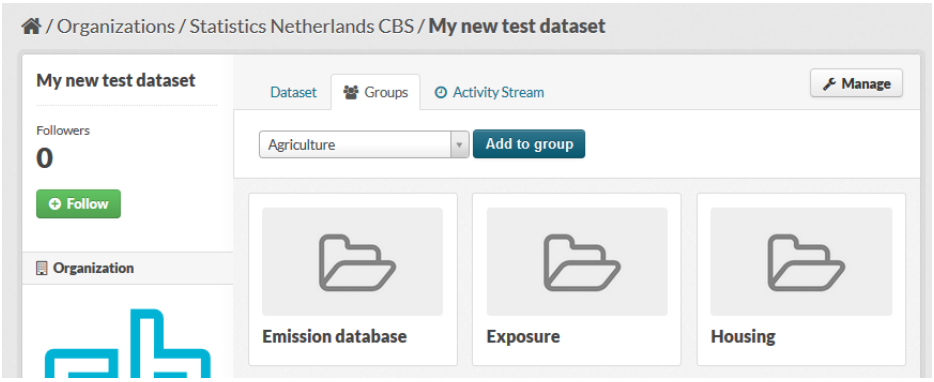
- **File**: either an upload or a link, we advise not to upload any data if you can link to it.
- **Name**
- **Description**
- **Format**: CSV, JSON or XML. If empty the system tries to recognize it. Leave empty in case of another format

A dataset can have multiple data resources, which might be useful to group several files on the same subject together. One example is to provide the data in a CSV as well as a textual description into one dataset Another example is to provide the data itself as well as a link to where the data was derived from into one dataset.



Assigning your dataset to the right Group(s)

The “Create dataset” form does not offer you a way to assign it to one or more groups immediately. So after creation it is not assigned to any of the groups. You can however assign your dataset to the right Group(s) afterwards using the “Groups” tab of the dataset page as displayed here:



The Application Programming Interface (API)

The data portal offers programmatic access to a resource of a recognized format via an application programming interface (API). This so-called DataStore API can be used for reading, searching and filtering data without the need to download the entire file first. One can use this mechanism to execute queries against data stored in the data portal. To do so one needs an api key, which by default has any registered user or can can be regenerated from your personal settings page. More information on the possibilities of the DataStore API can be found here:

http://docs.ckan.org/en/ckan-2.6.0/maintaining/datastore.html#the-datastore-api

Contact us if you want to interact with the ClairCity data using this API. We have example software to do this and can together decide on the optimal strategy for your use case.

How to register

Click on the register button and follow the wizard. Once you created an account please send an email to claircity@cbs.nl mentioning your username and the ClairCity organization you work for. We will assign your account to this organization in the data portal and to the ClairCity organization. From then you have the right to upload datasets on the ClairCity data portal on behalf of your organization.

Contact

Questions, suggestions, remarks can be sent to claircity@cbs.nl