# Circular economy data hub manual

# Table of Contents

DIGICIRC

# Structure

**Dataset:** The data publishing unit in the CE data hub is called "dataset". A dataset is a parcel of data - for example, it could be the locations of car sharing stations in a specific city. A dataset consists of "metadata" and a number of "resources", which hold the data itself. Data formats can include CSV or Excel spreadsheets, XML file, PDF documents, image files, linked data in RDF format etc. A dataset can contain any number of resources.

**Metadata:** "Metadata" are loosely defined as "data about data". Metadata document and describe all aspects of a specific dataset (i.e. the who, why, what, when and where) that allow understanding of the physical format, content and context of the data. The metadata accompanying each dataset are described in Chapter 3.1.

**Domain:** The datasets are assorted in one of the three DigiCirc domains, i.e., Circular cities, Blue economy, Bioeconomy. For the first version of the data hub, the datasets are in the Circular cities domain.

**Sectors:** Each domain consists of sectors for further categorizing the datasets in the data hub. Each dataset may belong to one or more sectors. Currently seven sectors are available and they represent a simple way to help users search and access data thematically.

**Administrator:** The data hub has an administrator who is responsible for managing the content, managing the users and allocating editing authorization rights to the users.

**User:** The CE data hub can be publicly accessed but if you want to contribute by uploading data you have to register. The administrator will authorize you to upload data related to a specific domain, by assigning the role Editor to you.

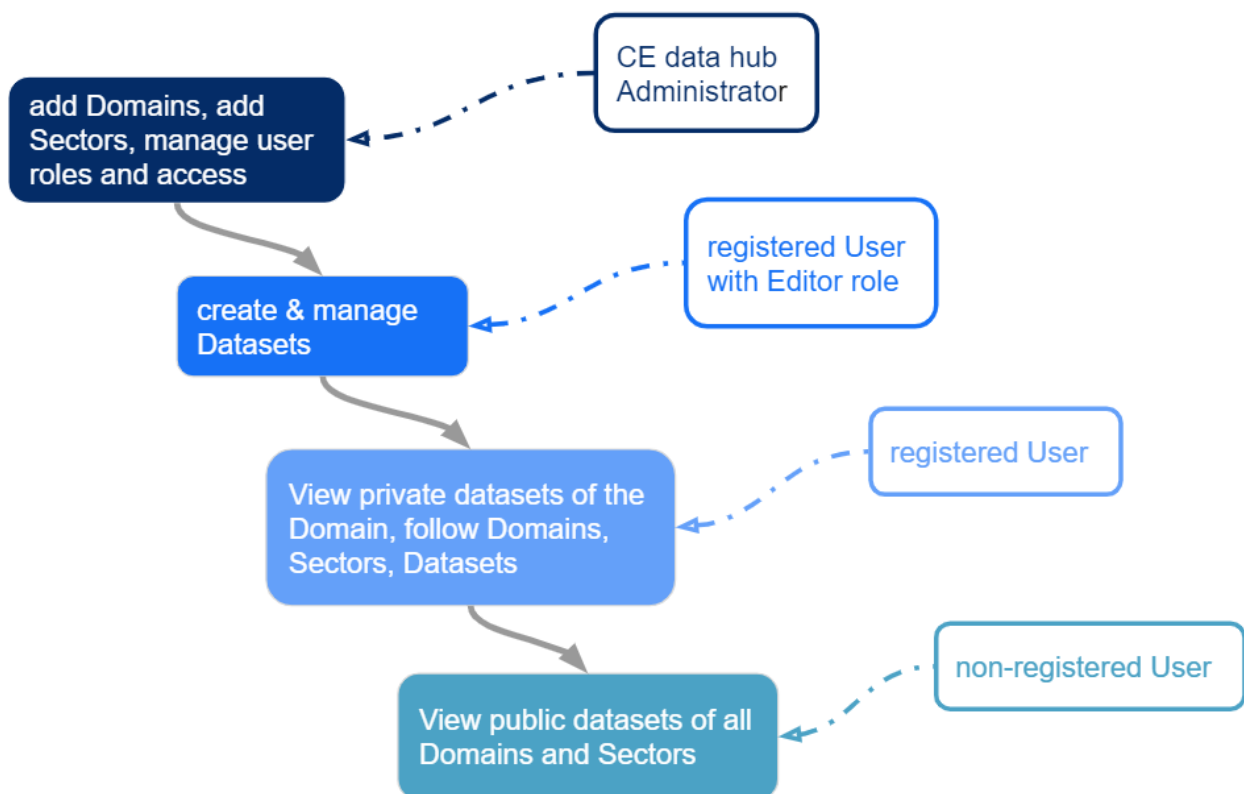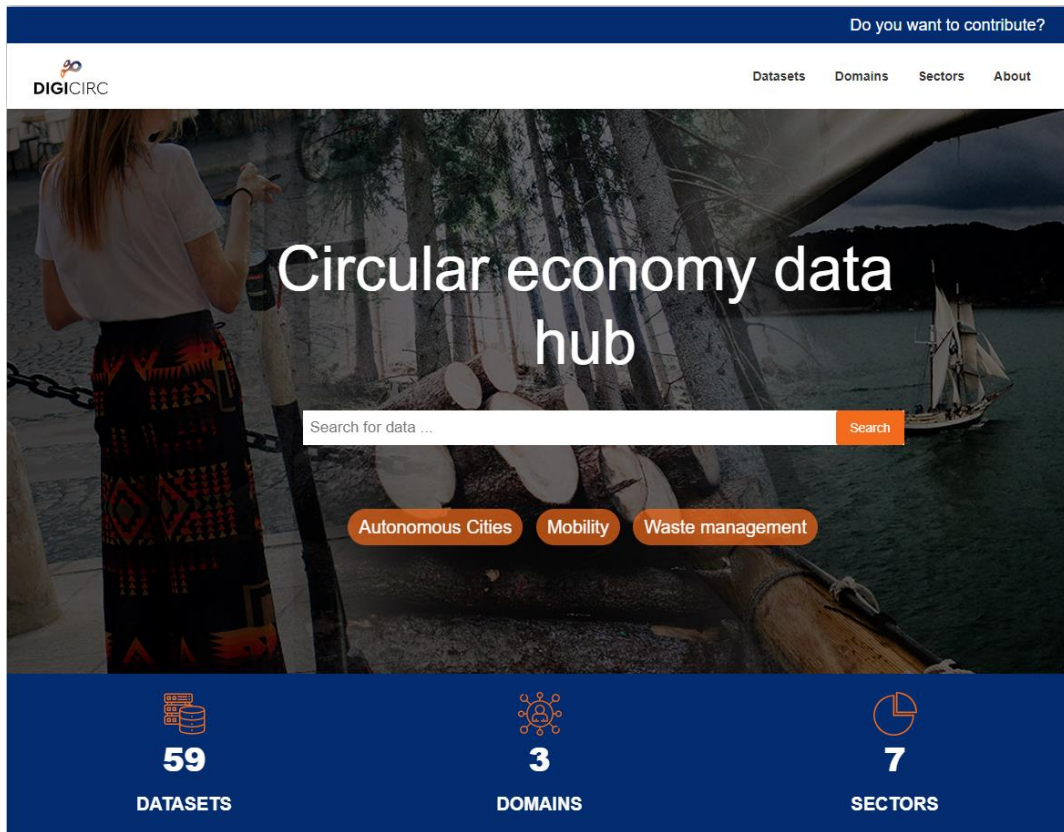The following diagram depicts the structure of the Circular economy data hub schematically.



Figure 1: Structure of the Circular economy data hub

# Welcome page



Figure 2: Welcome page of the Circular economy data hub

The welcome page of the CE data hub includes a menu for allowing you to easily navigate to the Datasets page, the Domains page, the Sectors page and the About page. The main feature of the welcome page is the Search functionality, that directly produces a list of datasets based on your search input. The most common tags are available under the search bar, directing you to the datasets related to those keywords. Statistics about the data hub are included in the next part of the page. Furthermore, descriptions about the three DigiCirc domains and links to respective domain pages are included.

# Domain

You can navigate to the Domains page of the data hub by selecting "Domains" from the menu (Figure 3). Each domain has a dedicated page, where you can find information about the domain (tab "About"), search within its datasets (tab "Datasets") and look at the latest activities relating to the datasets of that domain (tab "Activity Stream") (Figure 4).
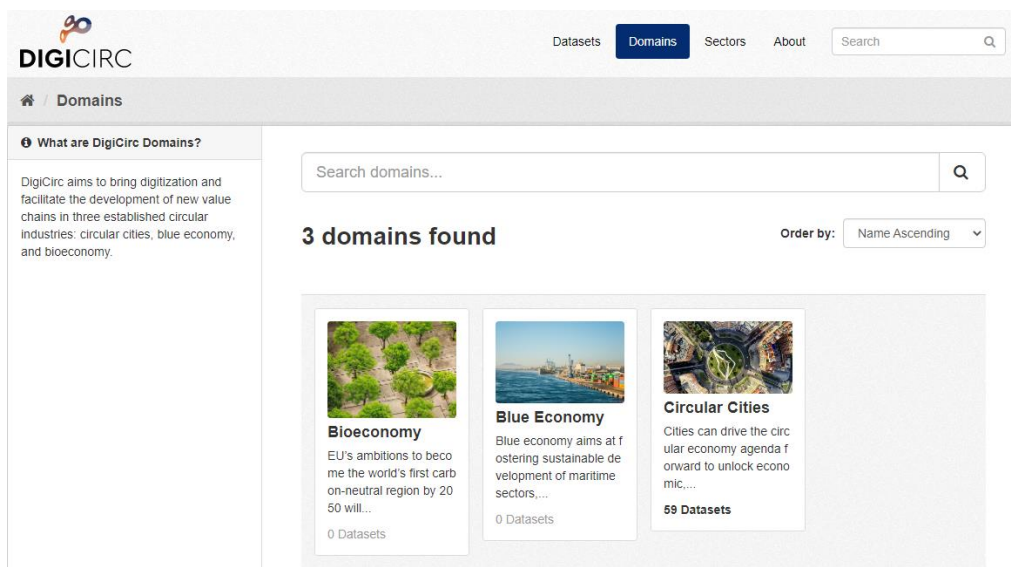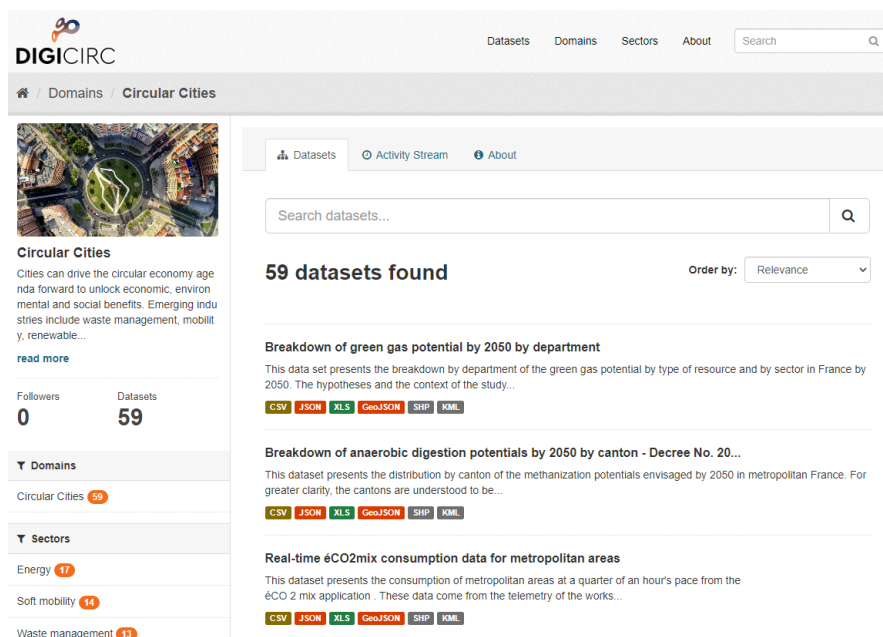


Figure 3: Domains page



Figure 4: Page of Circular cities domain

# Sector

In the CE data hub, you will find different sectors related to the certain thematic aspects of the three DigiCirc domains, by selecting "Sectors" in the menu (Figure 5). Each sector has a dedicated page, where you can search within its datasets (tab "Datasets") and look at the latest activities relating to the datasets of the sector (tab "Activity Stream") (Figure 6).
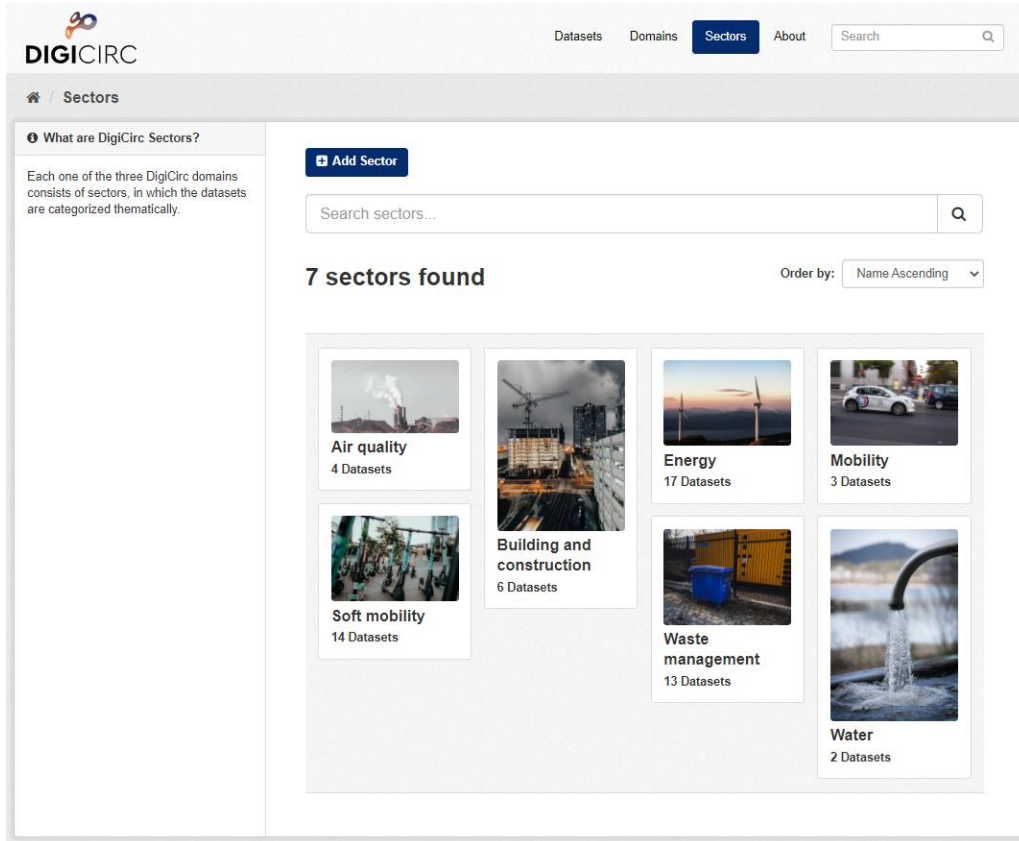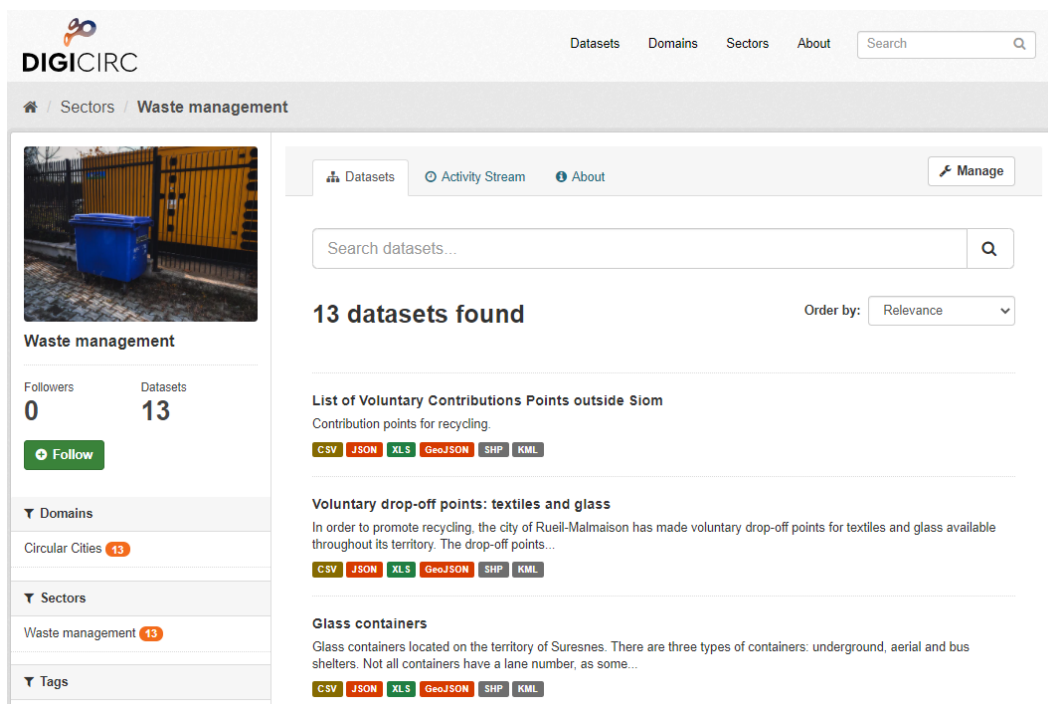


Figure 5: Sectors page



Figure 6: Page of Waste management sector

# Search for datasets

To find datasets in the CE data hub, you can type any combination of search words (e.g., "water", "recycle", etc.) in the search box on any page. The CE data hub will then return all corresponding search results as a list (Figure 7).



Figure 7: Datasets page with the search field and the search results

On the search result page, you can sort the results according to relevance, name, modification date or popularity by selecting "Order by". You can also limit the results using the filters on the left column (Domains, Sectors, Tags, Formats and Licenses). You can combine filters, selectively adding and removing them, and modify and repeat the search with existing filters still in place.

Additionally, it is also possible to search and filter for datasets by using the "search by location" option. You can click on the edit button (pencil) on the map in the top right corner. You can zoom in and zoom out using the "+/-" buttons and move the map with the little hand symbol. Then draw a square over the area you are interested in and press the "Apply" button (Figure 8). You will get a list of fitting results.
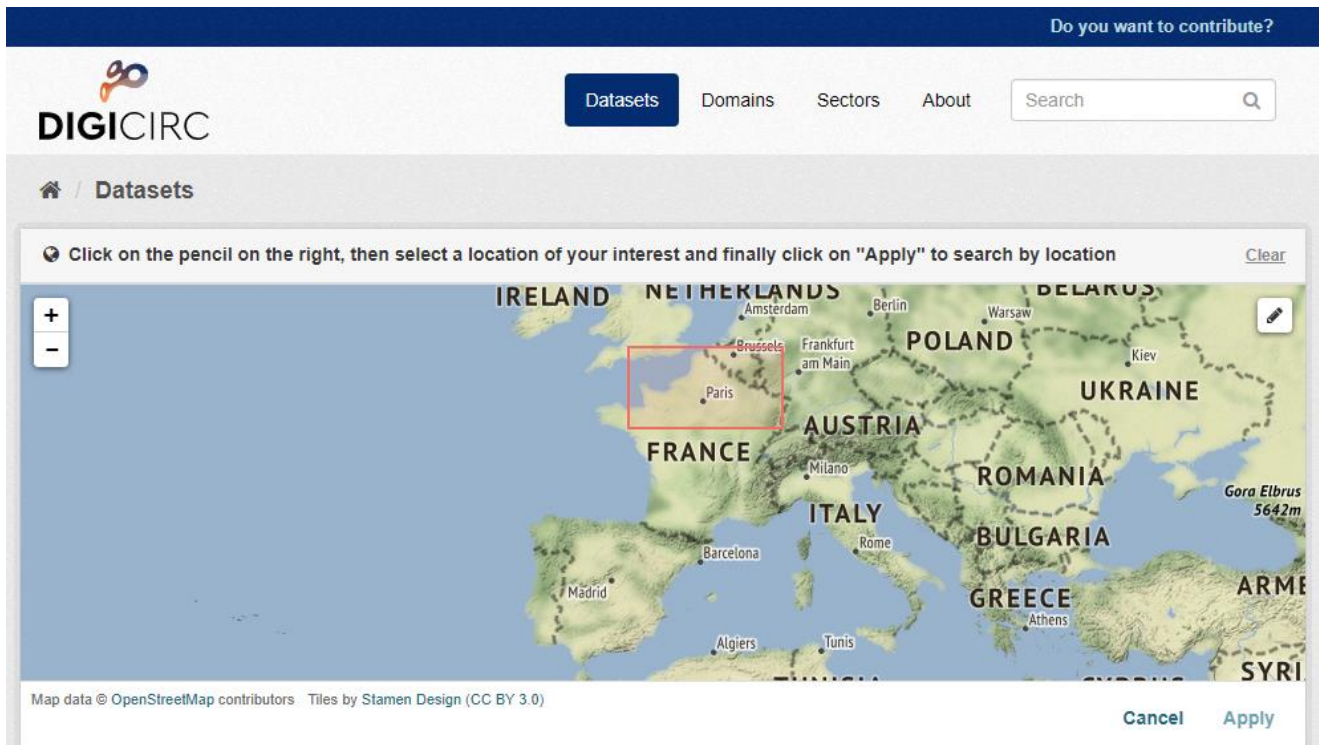
Figure 8: Search by location

Additionally, you can select "Domains" from the menu to view the three DigiCirc domains and then select the domain you are interested in and you will be directed to the domain's page. By typing a search query in the main search box on the page the data hub will return search results as described above but restricted to datasets from the specific domain. Apart from typing in the search box, you can explore the datasets in that domain. Respectively, you can select "Sectors" from the menu and follow the same process to explore the datasets thematically.
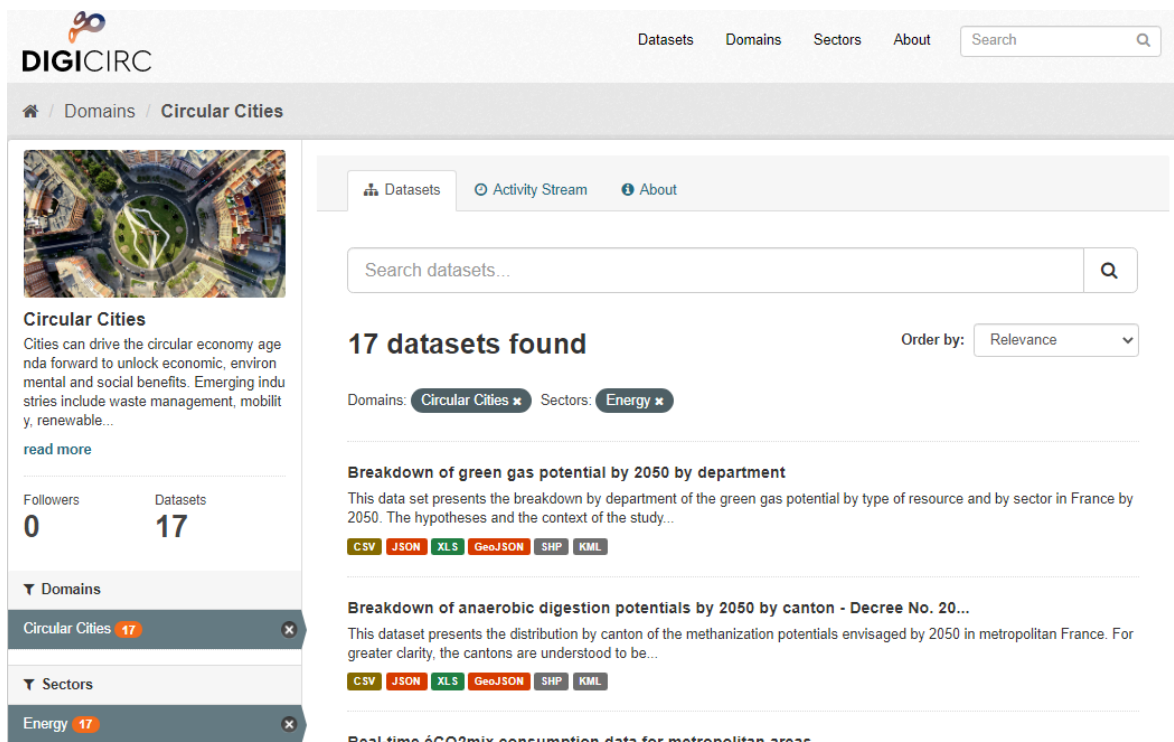


Figure 9: Datasets in the Domain's page with applied filters

# Dataset

**Once you find a dataset you are interested in and select it, the CE data hub will display the dataset page (**



Figure 10). On the overview page of a dataset, you will find three tabs: "Dataset", which shows the data and resources belonging to this dataset as well as additional info (metadata), "Sectors", which shows the sectors this dataset belongs to and "Activity stream", which shows the history of recent changes to the dataset. On the left part is a static column that displays the title of the dataset, the domain that it relates to and the license of the dataset.

On the "Dataset" tab you can see all the information of the dataset including the title, the description, the list of data and resources, the keywords associated to the dataset and the additional info. The map shows the wider area of interest for this dataset. The "Explore" button on the right of each resource offers the following options: "Preview", which shows the page of the resource including additional information, a preview and the API and "Download", which downloads the file directly. The list of keywords is collected both from the template prepared for the dataset and also from its original data source. The additional information presents the metadata of the dataset, collected both from the template prepared for the dataset and also from its original data source.

Figure 10: Dataset overview page

# Data preview and visualization

In the resource page, information for the specific file is presented and you have several preview options. The type of the file is presented on top, along with a link to its original source and the description of the dataset. Files in the format of CSV and XLS spreadsheets can be previewed in a grid view, with map (after definition of the coordinate data fields in the file) and graph views also available if the data is suitable. The resource page will also preview resources if they are common image types, PDF, or HTML.

By selecting the "Grid" option, you can see in a table view the contents of the file. Sorting functionalities are available. Below the table, you have access to the data dictionary, that includes information for the different fields the file consists of (Figure 11).

DIGICIRC

Figure 11: Dataset explore page, preview of data in grid

The "Graph" option allows you to create diagrams of five different types and select which fields of the file they want to display on the two axes of the graph. As a result, you are able to create dynamically any combination of data and get valuable insights for the dataset (Figure 12).



Figure 12: Dataset explore page, preview of data in a line graph

If the resource contains geospatial information, then a third option is available, to display the contents on a map by clicking on "Map" (Figure 13). The map is created automatically, but you have the option to select the fields you want to display on the map (Figure 14). Clicking on points or polygons on the map displays the information of each asset.

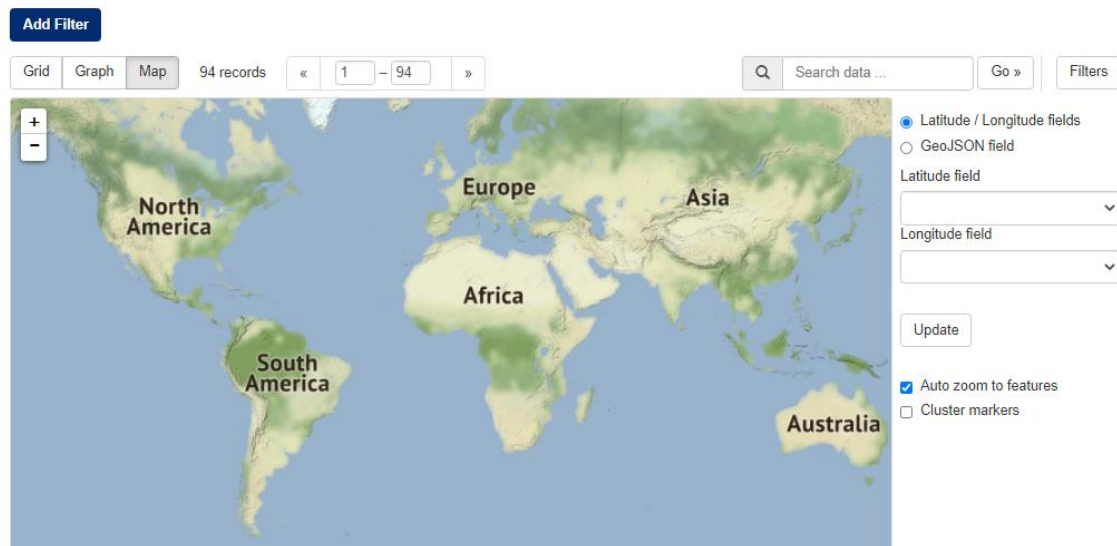Figure 13: Dataset explore page, preview of data on a map



Figure 14: Dataset explore page, preview of data on a map with user selection

# Download and API

Apart from downloading a resource directly from the dataset page, you can navigate to the resource page where you are presented with more options. The download button is still available at the top of the page, where you can select the file format you want to download (Figure 15). Furthermore, the API option is available, which allows you to access the data directly for use in you solutions. By clicking on the "Data API" option, you are presented with a pop-up, where all the information for the API is available. More specifically, the available Endpoints are presented, along with several examples to guide you. Links to documentation are also included (Figure 16).

DIGICIRC

Figure 15: Download files in resource page



Figure 16: API available in resource page

# Register and login

Users of the CE data hub can search for and find datasets without any restrictions. However, in order to upload a dataset, a login with appropriate permissions is needed. Anyone – not only DigiCirc project partners – can apply for registration. Registration is available through the welcome page when clicking "Do you want to contribute?" at the top right of any page. The required information for the registration is username, full name, e-mail address and password (Figure 17).

After the registration, the Administrator will assign editing authorization rights to you, for them to be able to upload a dataset. The Administrator will add you to the three domains and will assign to you the "Editor" role, so you can edit and publish datasets to each domain (Figure 18).
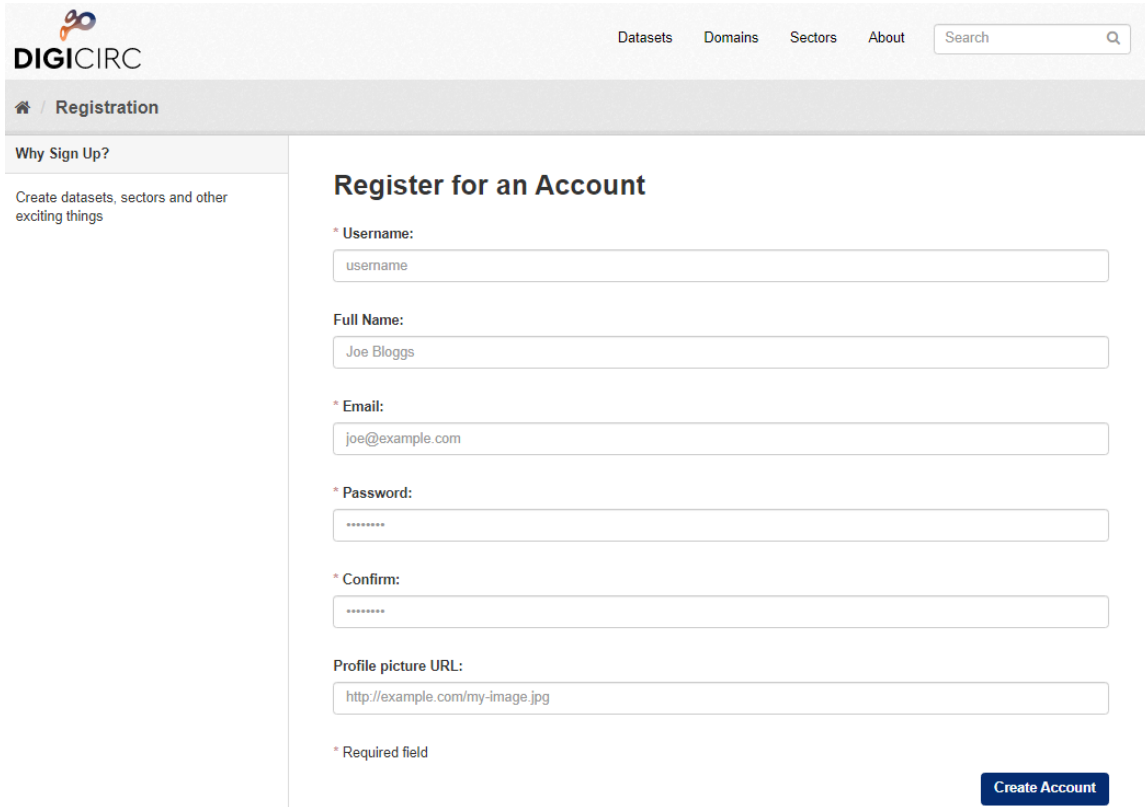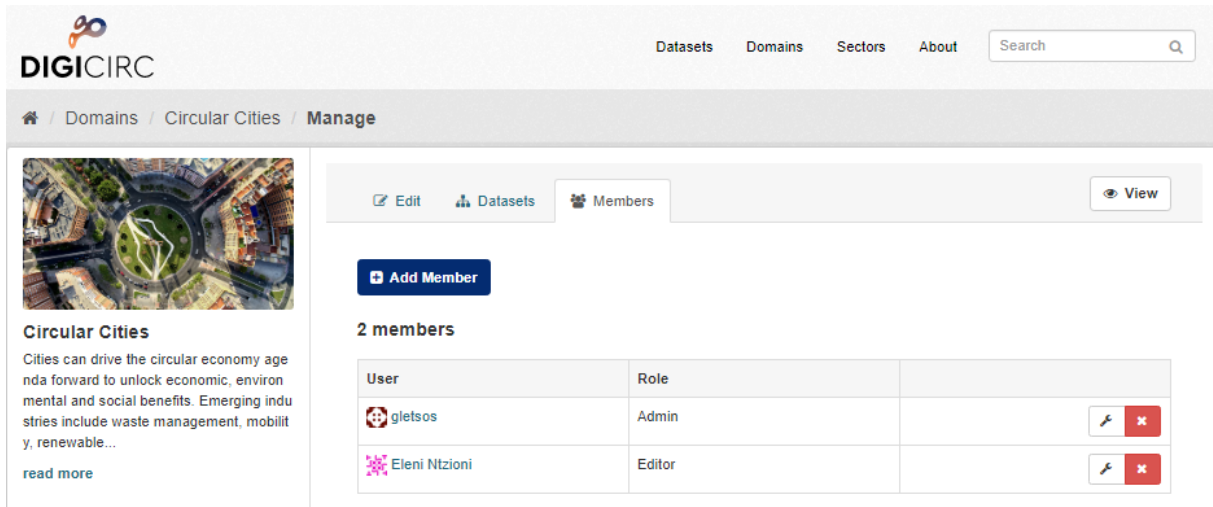
Figure 17: Registration page



Figure 18: Domain's page in the editing mode, tab "Members"

# Upload

In order to upload a file, you log-in and then you can either select the "Add Dataset" button in the Datasets page (Figure 19) or navigate to the domain that the dataset is in relation to and then select the "Add Dataset" button above the search box (Figure 20). The "Add Dataset" button will be available after you have been assigned editing authorization rights by the Administrator.
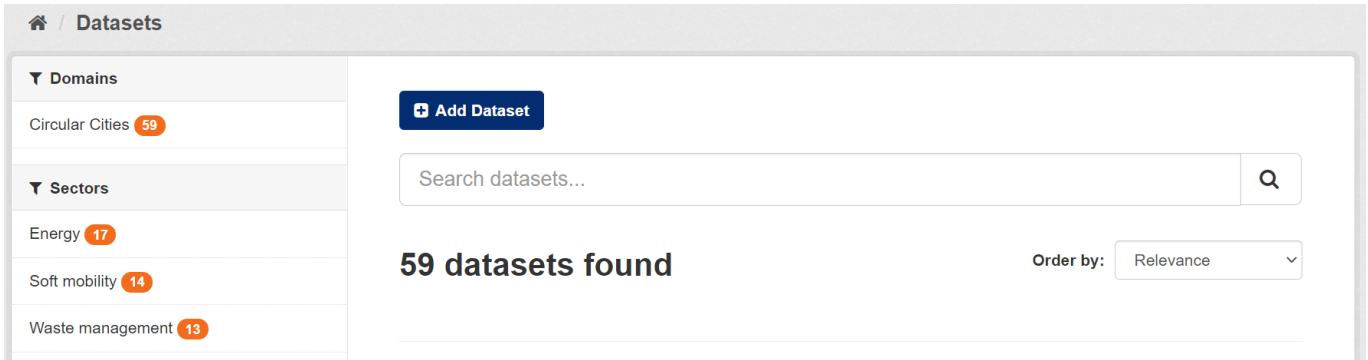
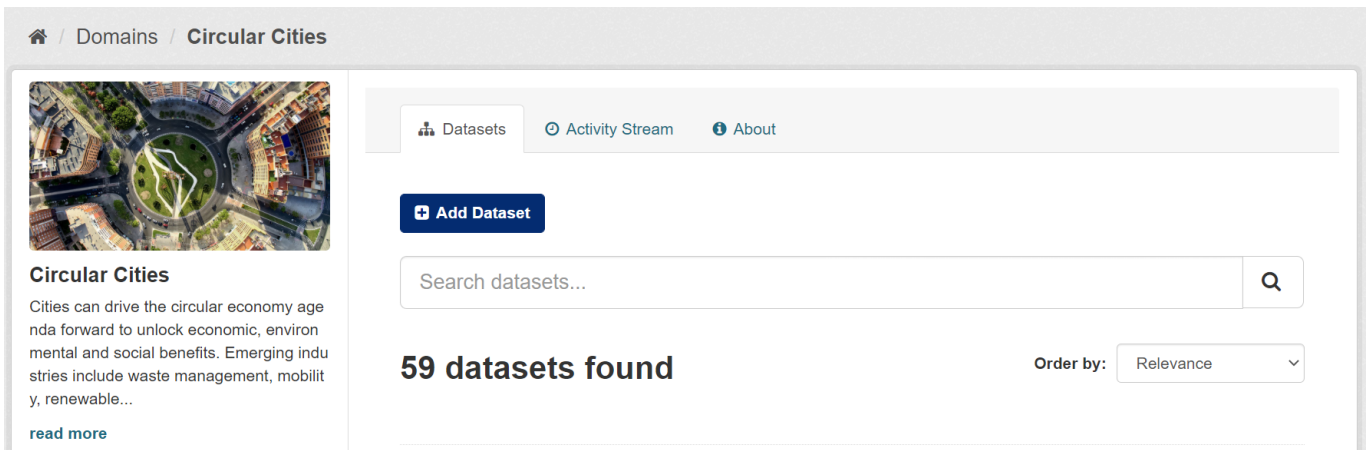Figure 19: "Add Dataset" in Datasets page



Figure 20: "Add Dataset" in a Domain's page

For each new dataset, you have to enter specific information. During the metadata entry you can select if a dataset should be private or public. Before uploading any resources containing the actual data, you need to describe the dataset by providing metadata on the "Create dataset" page (Figure 21,). On the "Add data" you can add one or more "resources" which contain the data for this dataset. To add a resource, you can either choose a file from your local directory by pressing the "Upload" button or link to your data resource by clicking the "Link" button. Additionally, you should add a name for the resource - different resources in the dataset should have different names – and define the format of the resource, e.g., CSV, XLS, JSON, PDF, etc. This field can also be left blank as it will be guessed automatically (Figure 22).

Figure 21: Create dataset page - metadata fields



Figure 22: Create dataset page, upload of resources section

# DIGICIRC

End of Document