DELIVERABLE 6.2
PARCOS DATA EXPLORER

"ParCos – Participatory Communication of Science"
A HORIZON 2020 RESEARCH AND INNOVATION ACTION

Consortium: Lappeenrannan-Lahden teknillinen yliopisto (FI, coordinator), Katholieke Universiteit Leuven (BE), Vlaamse Radio- en Televisieomroeporganisatie (BE), and Knowle West Media Centre LBG (UK).

Webpage: https://parcos-project.eu
Duration: 1/2020 – 12/2022
Grant: H2020-872500 (Call H2020-SwafS-2019-1)

Contact (co-ordinator):
Asst. Professor Antti Knutas & Asst. Professor Annika Wolff
LUT University
e-mail: parcos.project@lut.fi

Disclaimer: This document's contents are not intended to replace consultation of any applicable legal sources or the necessary advice of a legal expert, where appropriate. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user, therefore, uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors' view.

DESCRIPTION OF THE DELIVERABLE

Overview	Details
Authors	Natasha Tylosky (LUT), Finland
Reviewers	Antti Knutas (LUT), Finland
Number of Deliverable	6.2
Title of Deliverable	ParCos Data Explorer
Type of Deliverable	Demonstrator
License	CC BY 4.0, see https://creativecommons.org/licenses/by/4.0/
Attribution	CC BY 4.0 ParCos, http://parcos-project.eu H2020-872500
Dissemination Level	Public
Contractual delivery date	2021-10-31
To be cited as	Tylosky Natasha. (2021). Parcos Data Explorer, Deliverable 6.2. of the Horizon 2020 project ParCos, EC grant agreement no 872500, Lappeenranta, Finland.

SUMMARY

This report describes the technical set up behind the Data Explorer, as well as the rationale behind it, and a written description of what everything in the Data Explorer does. It also provides links to a video walkthrough on how to use the Data Explorer as a user, the Data Explorer GitHub repository, and a link to the live interactive current version of the Data Explorer (Data Explorer v1).

TABLE OF CONTENTS

1 INTRODUCTION

1.1. Purpose of the Data Explorer

The Data Explorer is a data visualization system that was created under the Participatory Communication of Science (Parcos) project. The Parcos data explorer is a novel way of visualizing a diverse range of data sets. The data explorer is meant to be a tool for both scientists and non-scientists to be able to interact with and curate data. The idea behind the data explorer is to provide local people with a means of interacting with data that relates to their environment. (Loukissas, 2019)

Users of the Data Explorer can interact with and curate data by creating their own "data stories" from data sets chosen by and curated by experts. The idea behind the data explorer is to collect a diverse range of curated data sets related to a specific topic, and then to display all of those data sets in the same platform so that non-experts may explore those data sets and form their own hypothesis about the data.

The goal of the data explorer is to promote citizen science and to provide a way for researchers to include non-experts in the research process. Data inclusivity is especially important when researchers are outsiders working with local communities, and the Parcos data explorer could be a tool for fostering trust and understanding between researchers and the general public. The data explorer is based on the Bristol Approach (Hudson et al., 2020), which aims to foster and support a people and issue-led process for citizen science and bridge the gap between research and how the general public understands science.

1.2. Test Case

The test case for the prototype of the Data Explorer is biodiversity data related to Helsinki. Participants will explore various data sets and pre-made data stories related to the biodiversity of flora and fauna in Helsinki. Users will be asked to come up with their own data stories related to biodiversity in Helsinki.

1.3. Data Comics

The visualization method chosen for this study was data comics. Data comics are a new visualization method that is ideal for communication large amounts of complex information in a simple way (Bach et al, 2017). However, any type of 2D data visualization could be used in conjunction with the data explorer. You can see one of the data comics found in the data explorer below.
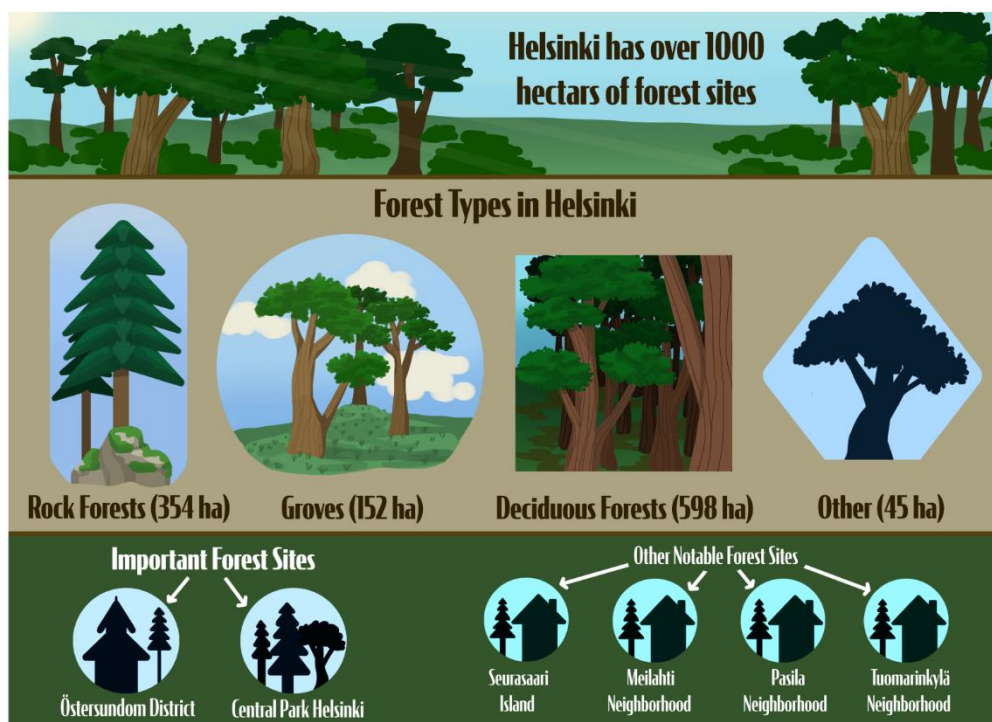
Figure 1: One of the data comics made for the Data Explorer prototype v1

2   TECHNOLOGY USED

Prototype Data Explorer v1 was built using the game engine Unity. The prototype was built using Unity version 2020.3.11f1. If you would like to edit the Data Explorer directly in your own version of Unity it should work in any version from 2020.3.11f1 onward.

3   LIVE VERSION

To see the Data Explorer v1 live, go to: https://natashatylosky.itch.io/data-explorer-v1

4   GITHUB VERSION

You can find the full current back-end version of the Data Explorer (Data Explorer v1) on GitHub. If you want to edit the Data Explorer directly and make your own changes to it or version of it, you should have some programming knowledge already. Unity experience or C# knowledge will be particularly useful. If you do not have any programming knowledge or Unity experience, there are plenty of wonderful tutorials out there to learn from. Please refer to https://unity.com/learn for more information.

To see the project's GitHub go to: https://github.com/tashisha/The-Data-Explorer

5 DATA EXPLORER WALKTHROUGH

5.1 Using the Data Explorer

The Data Explorer is laid out in such a way that each white point or "star" represents a data set. The rings around the stars represent a "cluster" of data sets used in a data story. And the lines connecting the stars serve to show when a data set is shared across multiple data stories. There can be two or more stars that contain the same data set, the connecting lines show which stars share data sets with each other. In this initial prototype there are only six data sets, with some of these data sets being used across different data stories.
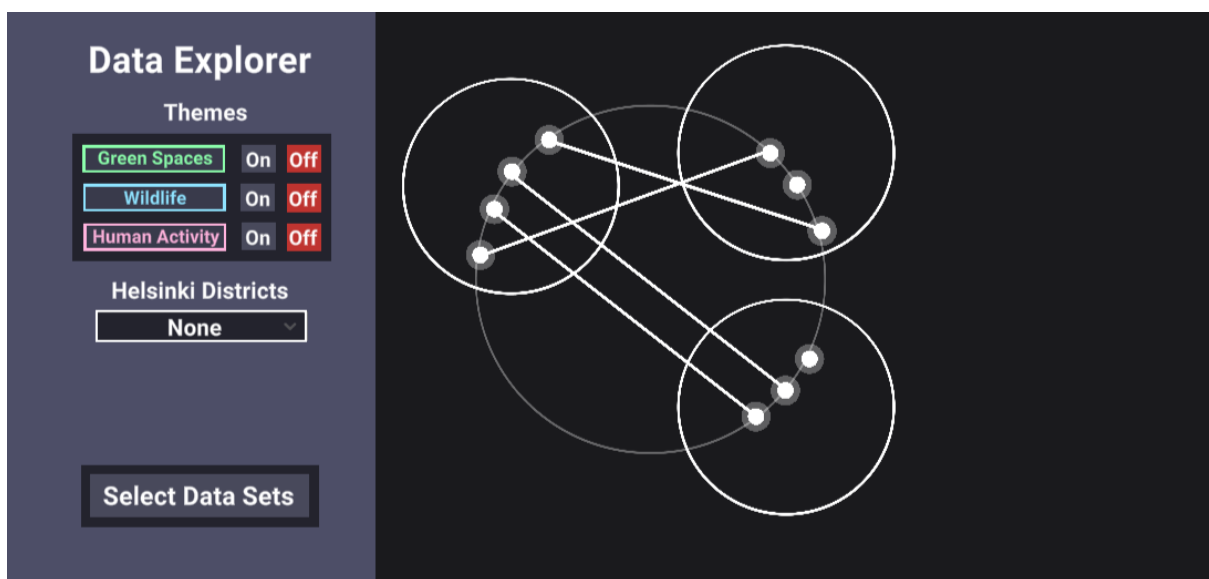
Figure 2: Data Explorer, opening screen.

When hovering over each star the star will grow brighter indicating it is clickable. Likewise, the data stories clusters also grow brighter when hovered over, indicating that they are also clickable.
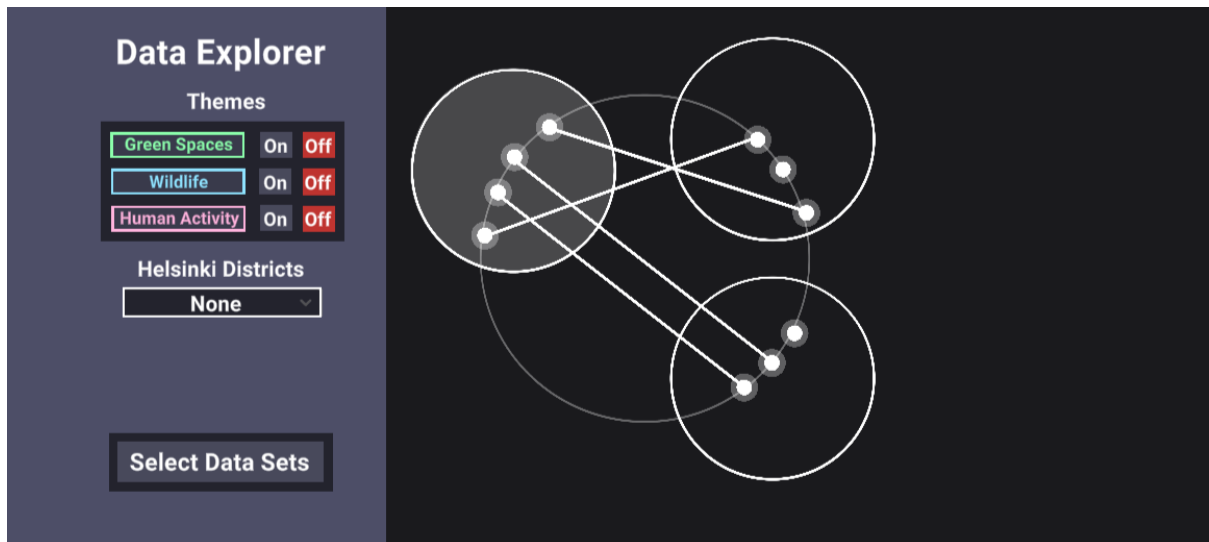
Figure 3: Data Explorer, hovering over a data story.

When a star is clicked on the related data set window will be opened. The data set window will contain the title of the data set, and add and remove button, a visualization of the data, and links to the data set(s) from which the visualization was made from. In this case the visualization is as described below, in Figure 4.
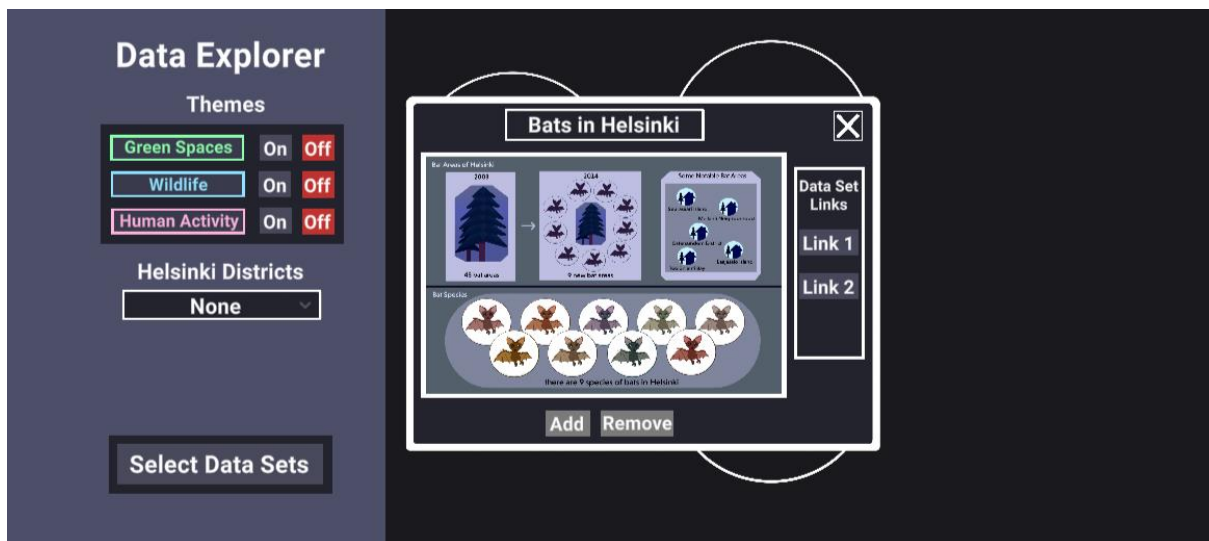

Figure 4: Data Explorer, opening a data set window.

A user can open a Data story by clicking on the highlighted circle. This will open a window where they can see the data story that the user created as well as the data sets that the user selected for their data story. The data stars are clickable and will open the corresponding data set window.
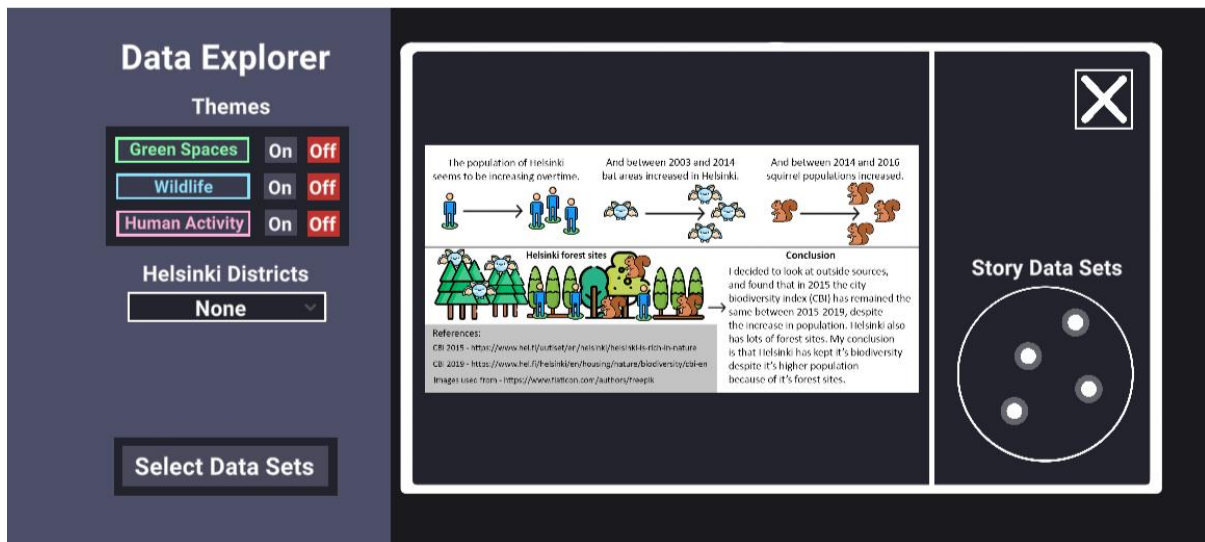
Figure 5: Data Explorer, opening a data comic window.

The data sets in the Data Explorer can be sorted by theme. Each theme is represented by a color. In this prototype the three themes a user can turn on are "Green Spaces", "Wildlife", and Human Activity". Each data set can have one theme.
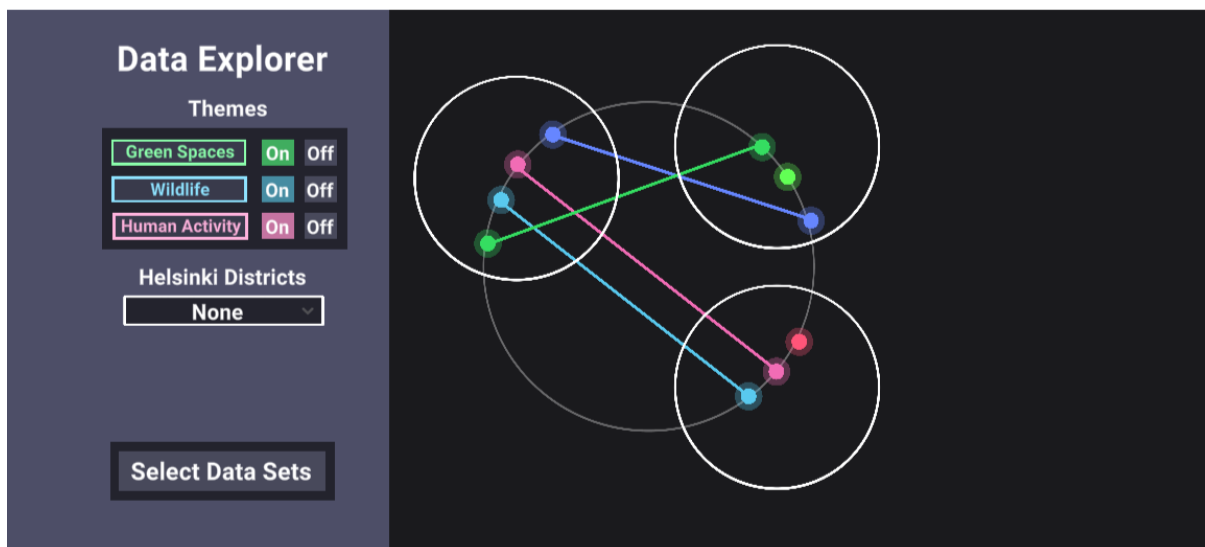

Figure 6: Data Explorer, activating the themes buttons.
The data sets can also be sorted by location, which the user can do by selecting a location in the dropdown menu. Here in this prototype the locations correlate to the districts of Helsinki. But any locations could be used. When a data set pertains to a location the "star" will change from round shaped to diamond shaped.
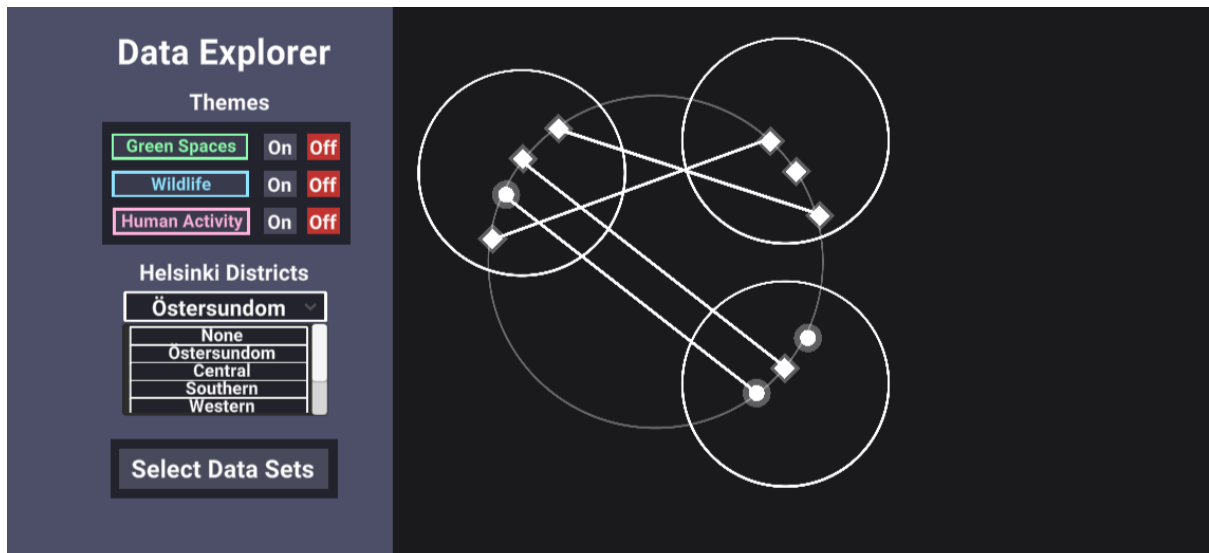
Figure 7: Data Explorer, activating a location filter from the dropdown menu.

Themes and locations can be activated at the same time as well.
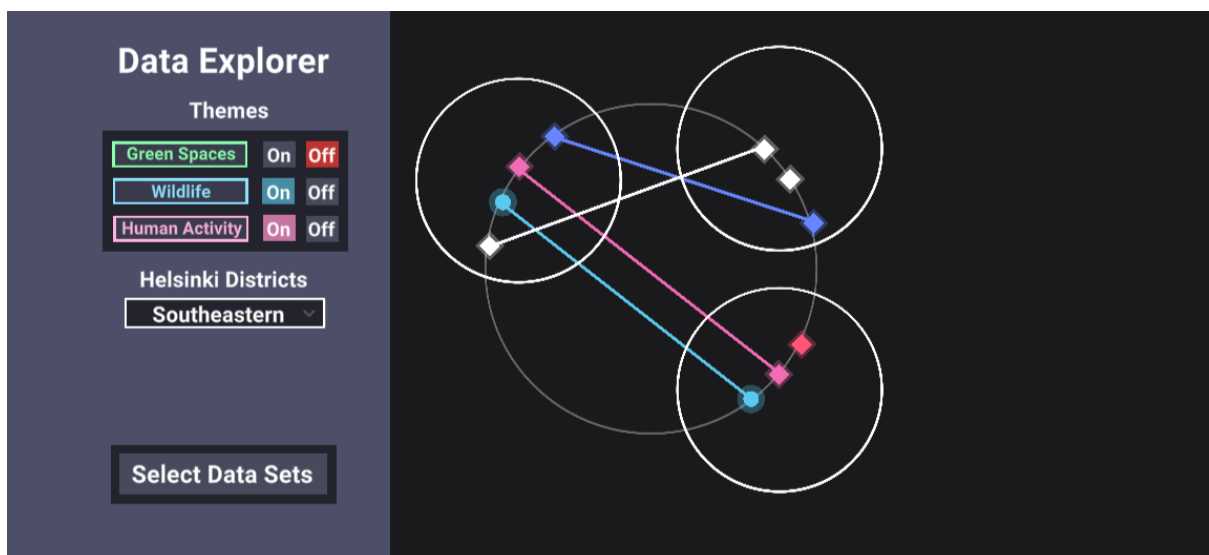

Figure 8: Data Explorer, activating the themes and locations buttons.

After a user has explored the data sets and the example data stories, they can click the "Select Data Sets" button to begin choosing data sets for their own data story. This will bring up the "My Data Sets" Window.
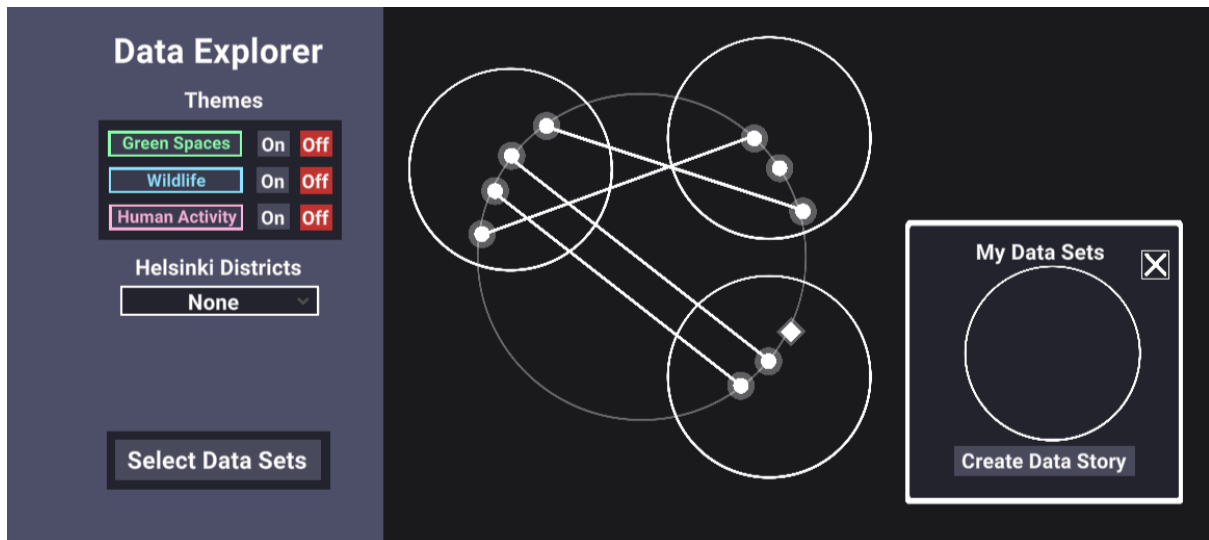
Figure 9: Data Explorer, opening the my data sets window.

The user can then click on one the stars to open its data set window. Then they can click on the "Add" button to add that data set to their collection. They can also remove data sets with the "Remove" button. The "Add" and "Remove" buttons are greyed out and usable until the "Select Data Sets" button is pressed. The stars in the "My Data Sets" window can also be clicked on.

Figure 10: Data Explorer, clicking the add button.

Once the user has selected their data sets, they can then click the "Create Data Story" button. This will open a window that contains their data sets and a window where they can upload the image that they want to use for their data story. This allows the user to create their own materials, for their data story, even using pen and paper, then to take a photo of their data story and upload that photo to the system. To upload their data story the user clicks the

"Select Image" button. The user can then click the "Upload" button to upload their data story to the system.
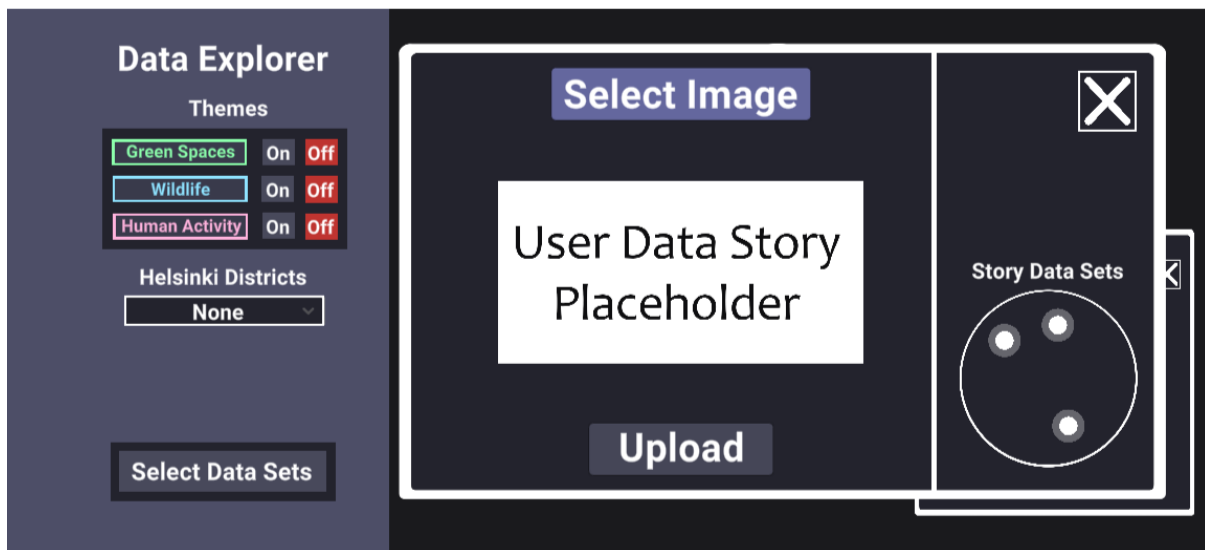

Figure 11: Data Explorer, user uploading their data story.

This will make their data story cluster appear alongside the other data stories in the system. The user can now click on and view their own data story, as well as see the where data sets that they chose were used in other data stories.
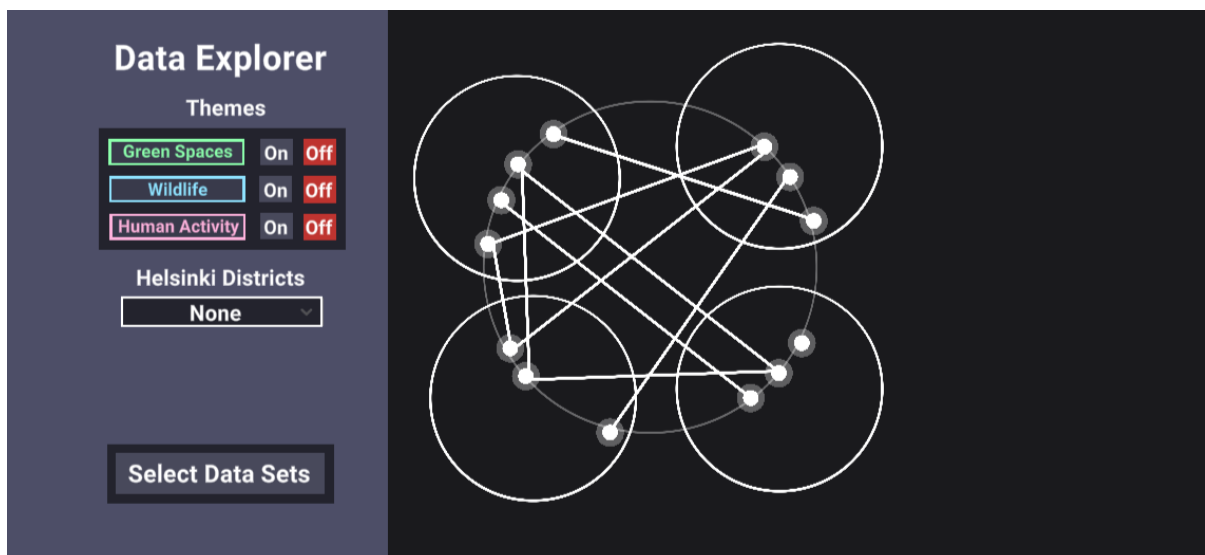

Figure 12: Data Explorer, the user's data story and its connections being visually represented as a data story cluster.

The user can also view the themes and locations associated with the data sets in their data story.
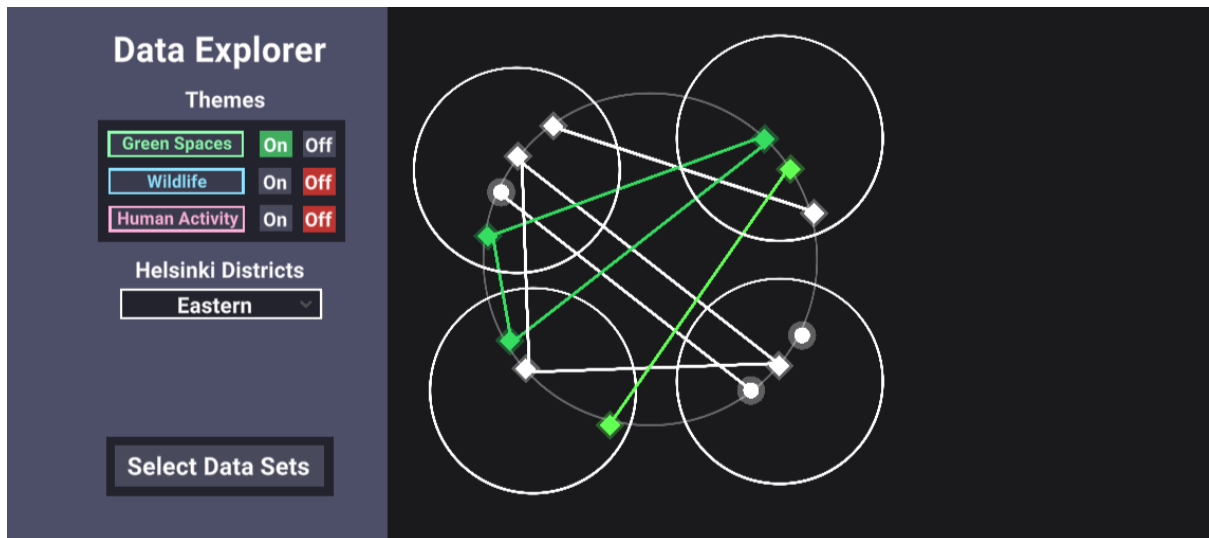
Figure 13: Data Explorer, the user's data story and its connections being visually represented as a data story cluster, with a theme button and location filter applied.

5.2. Video Walkthrough

To see a video walkthrough on how to use the data explorer go to: https://www.youtube.com/watch?v=6ajD7V1vT-M

6 CONCLUSION

The key advantage inherent in the design of Data Explorer is that this set up could be extrapolated to encompass a very wide range of data sets. For example, while this prototype is built around the theme of "biodiversity in Helsinki" The "constellation" of data stories used in this data set could potentially connect to another "constellation" about mental health in Tampere for example, and so on. Essentially this system can be scaled up at either a macro or micro level looking at data in an individual neighborhood or data in a whole country. And it can be used to bridge the gap between diverse data types that don't have an immediate apparent connect.

REFERENCES

Yanni Alexander Loukissas. 2019. All Data Are Local: Thinking Critically in a Data-Driven Society. The MIT Press.

Hudson L., Evans, P. and Banks Gross, Z. (2020, draft). The Bristol Approach to Citizen Science, deliverable 2.1. of the Horizon 2020 project ParCos, EC grant agreement no 872500.

Bach, Benjamin & Henry Riche, Nathalie & Carpendale, Sheelagh & Pfister, Hanspeter. (2017). The Emerging Genre of Data Comics. IEEE Computer Graphics and Applications. 38. 6-13. 10.1109/MCG.2017.33.