INCEPTION REPORT

Created By

Jose Guillen Santos and Wynne Wu

COMP 3220
Object Oriented Analysis and Design
School of Computer Science

University of Windsor

June 25th 2020

1 Vision and Business Case

The goal is to develop the next generation open data access portal, complete with the ability to collect data, manage it, distribute it in a various format, and deliver high-quality API libraries for developers. We are to present Windsor's stakeholders a state-of-the-art tool for open data.

The successful project will benefit the user through:

- Increased transparency and openness through better data availability and data accessibility.
- Better data understanding and data management for the City of Windsor.
- The cataloguing, collection, analysis and publication of maintained data sets to promote better understanding of data assets for related projects.
- Produce output and results. An open data platform will promote the re-use of knowledge assets thereby increasing the value of Windsor data.

2 Glossary

Information: Means information resources or Records protected by copyright or other information or Records that are offered for use under the terms of this licence [2].

Dataset: An organized collection of data. Datasets can be presented in tabular format or non-tabular formats such as an extensible mark-up language file, a geospatial file or an image file, etc [1].

Open Data: The idea of making data openly available for everyone to use and republish as they wish, without the restrictions from copyrights and patents [1].

Open Government Data: Open data produced by the government [1].

3 Risk List and Management Plan

1. **Lack of Stakeholder's Involvement and Lack of Communication**

   Most of the time stakeholders are interested in a project, but they do not participate in the development process. Stakeholders' involvement is crucial to the project's requirements analysis before the construction of the project. Failing to identify or adequately engage with stakeholders can cause failure to build a successful project. Therefore, developing team must make sure stakeholder is

actively engaged and have a clear communication plan with them by setting up a meeting every so often

2. **Request Big Change Late in Project**

   If stakeholders decide to make major changes to the site, it will delay the project's expected delivery date and may even require a major overhaul of the website. This issue can be managed by having a proper requirements analysis before the project starts and the paper statement must be signed and accepted by stakeholders and the project team. If huge changes are required, then it is important to have mutual agreement with stakeholders that any major changes proposed 2-3 weeks before the testing phase may be accepted but there might be a delay on delivery date and increase in budget.

3. **Urgent leave/Quit/Sick of the Key Project Team Member**

   Without the key project team member, the project will suffer a huge delay. Way to overcome this issue is to have an adequate buffer time added into the project's schedule. This way the team member can know whether everything is on time and other senior programmers will be able to give assistance.

4 Supplementary Specification

1. Hosting system can be hosted either on a platform provided by city of Windsor or by the service provider

2. The system must be stable both operationally and functionally 24 hours a day. The system must be able to access using low bandwidth connections and DSL communications without experiencing performance degradation. The bandwidth and latency of any connectivity within the cloud and the direct connection to the Internet services should be stable and as specified.

3. There should be scheduled and agreed occasional maintenance for the website. Any emergency maintenance, or recovery of the website, should be given consent and approved ahead of time. The measure of a stable and safe back-up process should also be indicated.

4. There should be a team that provides maintenance and support for the website.

5. Devices that can access the open data portal will include desktop, laptops, tablets and mobile devices. Therefore, the site has to be adaptive, have a responsive design and a user-friendly interface.

6. City of Windsor's ambiguous budget presents considerable risk to the development of the project.

5 Development Case

There are four phases in Unified Process which includes, Inception, Elaboration, Construction and Transition. In the Inception phase, the development team communicates with stakeholders regarding the risks, cost, expectation, estimated delivery time and functional requirement of the project. Then in the Elaboration phase we would start creating and analyzing the full functional requirements (use-cases) of the project. The use-cases and class diagram will be shown to the stakeholders.

Then the next phase is Construction phase. In this phase, the development team will start coding the project and implement the features for each iteration. In this case, we will be using Python and CKAN. CKAN is an open data portal site build that provides many features such as searching, adding, and visualizing datasets. CKAN has many libraries built in Python—which we will use to build upon those features.

Lastly, at the transition phase the development team will fix any bugs that occur in the project and the finished project will be shown to the stakeholders. The stakeholder will give feedback regarding the project and assess whether any change needs to be made. The Development Case Table can be seen in Figure 9.

6 Use Case Model

Windsor's Open Data Portal would provide the functionality to add and delete datasets in large batches. It would do this by using the CKAN platform as the foundation of the site and using it's FileStore API to handle the data. Using these tools, we can also provide a way for users to search the site for datasets, view data visualizations, and download the data. Another use case would be to build on top of the CKAN platform and create APIs that would enable the developer to query, search, and download datasets. A Use-Case Diagram can be seen in Figure 6 which outlines a rough sketch of the relationship between use cases in this project.

7 Prototypes and Proofs of Concepts

A major requirement for the project was the ability to access the data faster by using usage statistics. This feature was prototyped by building a mock open data portal with CKAN that uses Apache-Solr as a search platform. With these technologies, the user can search for datasets with a search bar that suggests the most popular dataset categories. This mock CKAN site we built also proved the feasibility of other major requirements. The site currently has the functionality to add datasets, create organizations whose datasets will be hosted on the site, and create administrator accounts who have the permission to add datasets manually. Currently, only sample datasets have been added to the site, but it is possible to migrate Windsor's data

to the site manually at this point. Screenshots of the mock site and its features can be seen in Figure 1 through Figure 5 and are included in a folder along with this report.

8 Phase and Software Development Plan

Please refer to Figure 7 and Figure 8 for the phase/software development plan. These Gantt charts outline the steps and timelines for the next 3 iterations.

9 Iteration Plan

Start Date: June 25th, 2020
End Date: July 16th, 2020

High Level Objectives

- Address usability issues for users logging into the site.
- Deliver features adding the ability to transfer and add large datasets to the site.
- Enable data visualization for key datasets of interests to the stakeholders.
- Begin development of API features for developers that extends our current available features and gives access to Windsor datasets.

Evaluation Criteria

- Admins have favorable response to the site, can successfully add to datasets, tinker with site aesthetics, and can create organizations.
- Over 60% of the datasets available on the current Windsor data portal have been transferred over, and there is a feature or concept in place that allows admin to add datasets in batches.
- 15% of key datasets have a data visualization available to view when users query their data.
- Developers have a favorable response to the current build of API, and can access some datasets.

References

[1] The Corporation of the City of Guelph. “Guelph Open Data”. In: (2000).

[2] The Corporation of the City of Windsor. “City of Windsor’s Open Data Terms of Use”. In: (2000).

Appendices

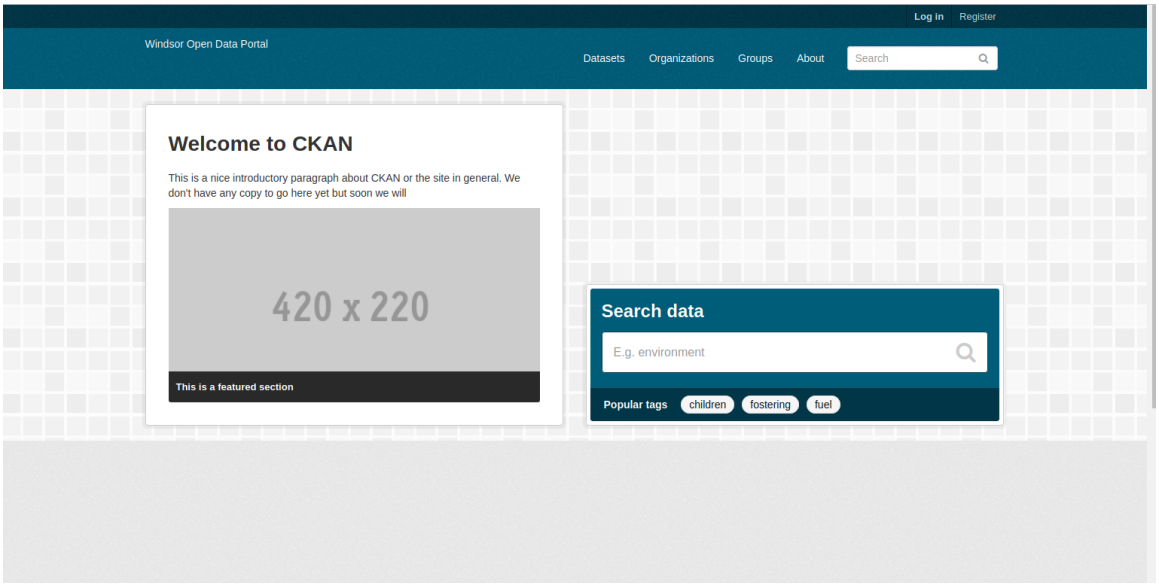

Figure 1: Mock Site Image 1

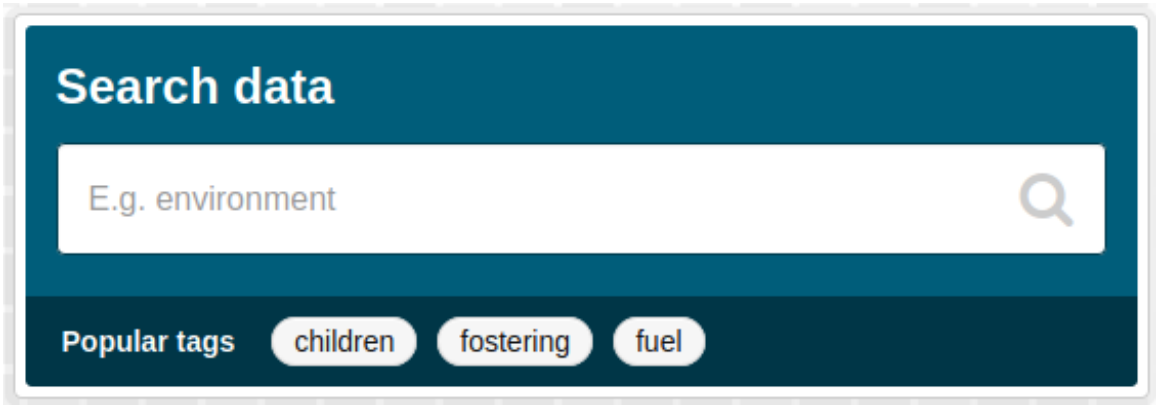

Figure 2: Mock Site Image 2

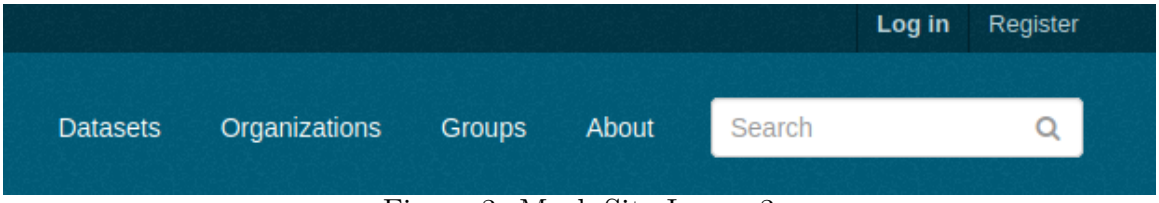

Figure 3: Mock Site Image 3

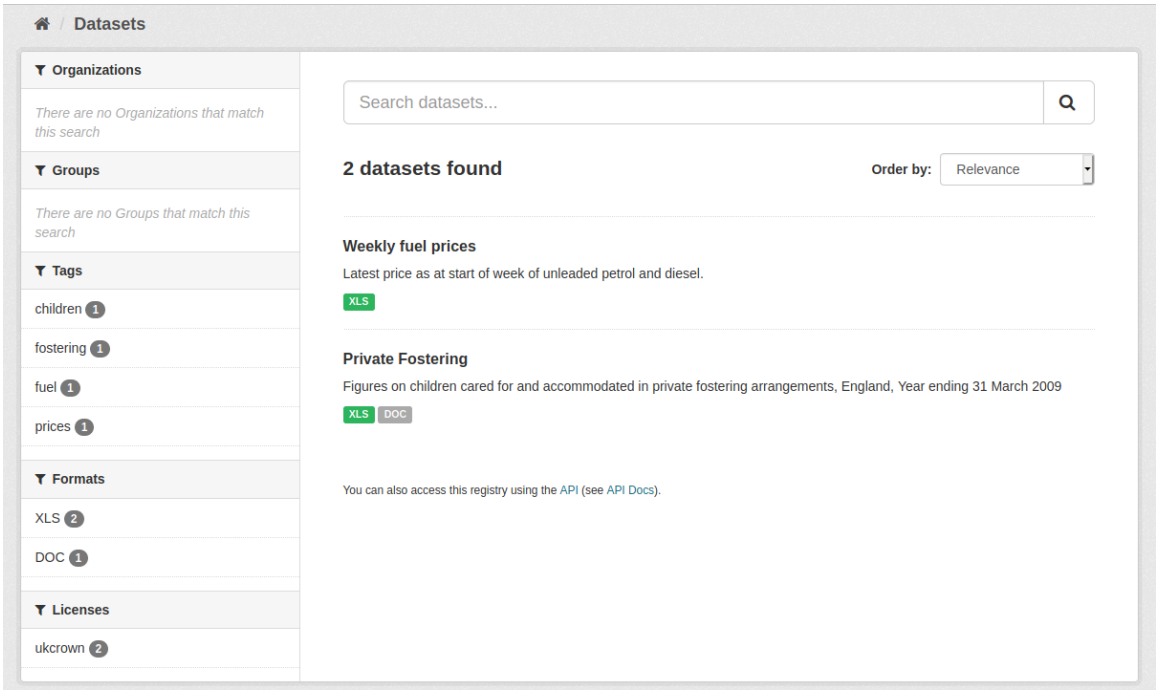

Figure 4: Mock Site Image 4

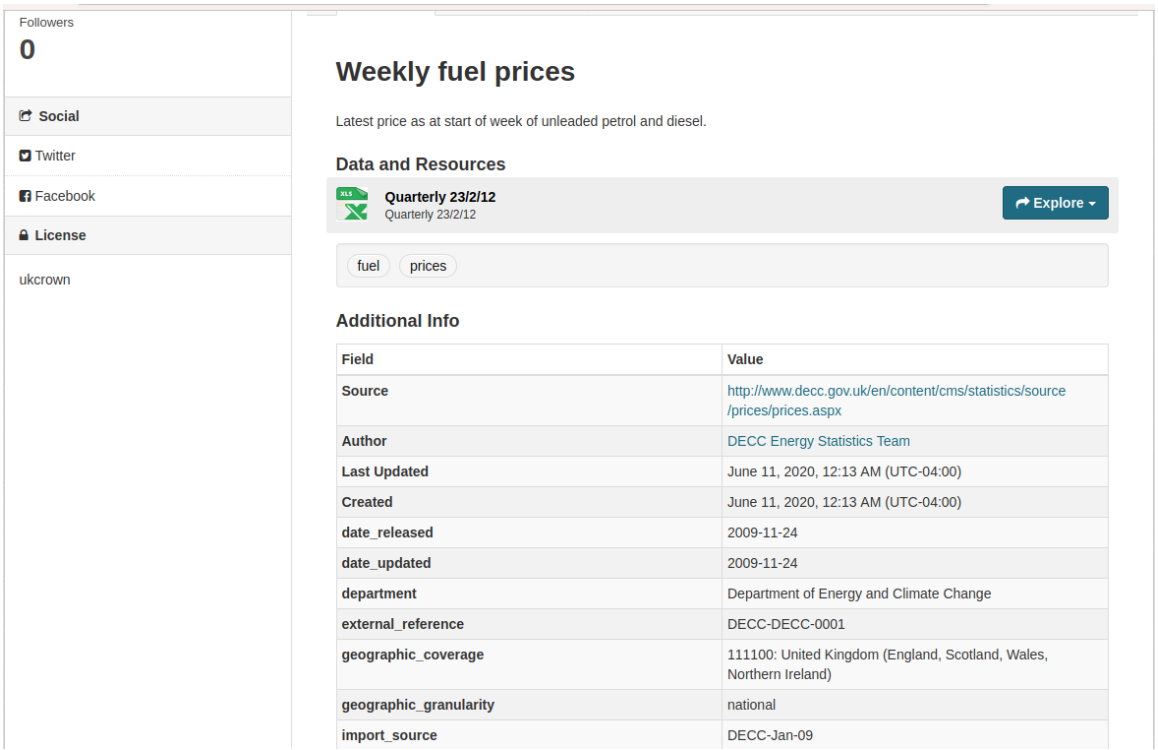

Figure 5: Mock Site Image 5

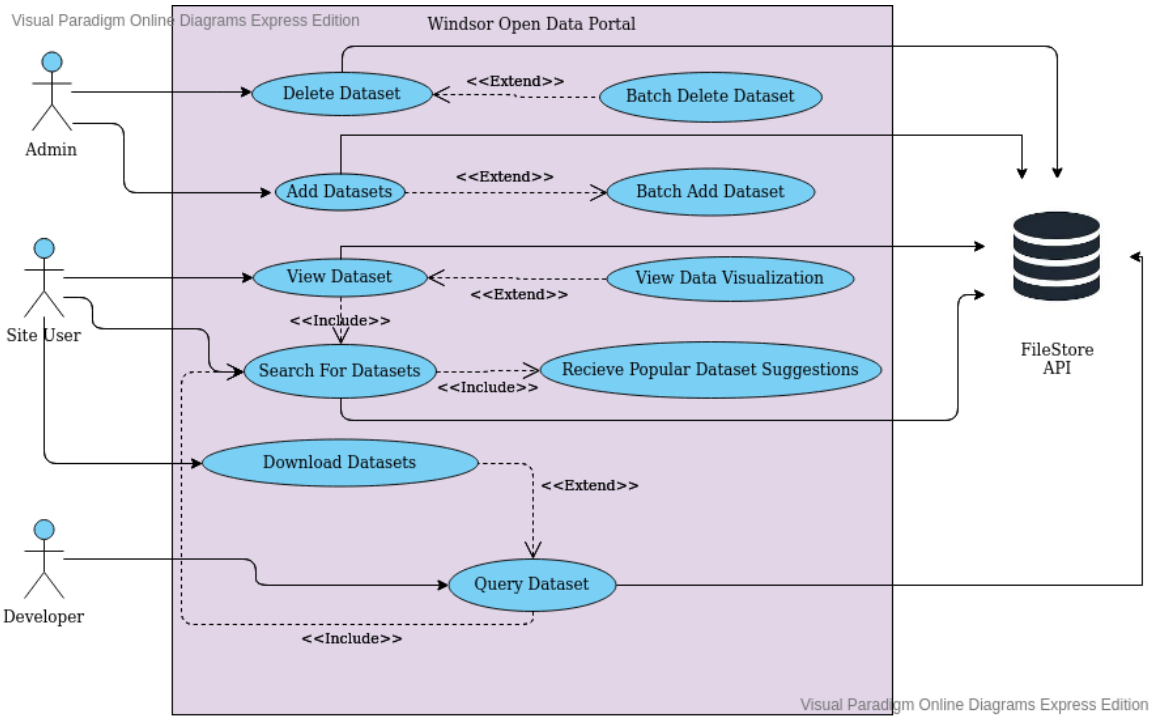

Figure 6: Use Case Diagram

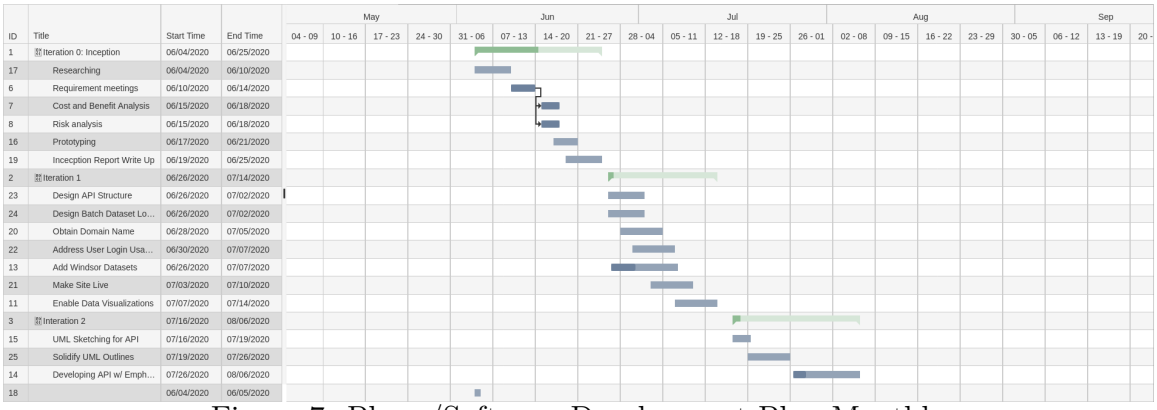

Figure 7: Phase/Software Development Plan Monthly

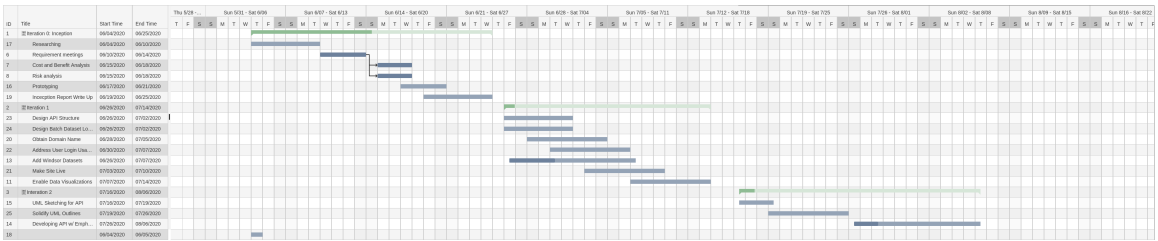

Figure 8: Phase/Software Development Plan Weekly

Artifact	Tool Used	Formal Deliverable
Vision and Business Case	Latex	Yes
Use-Case Model	Visual Paradigm 16.1	No
Supplementary Specification	Latex	Yes
Glossary	Latex	Yes
Risk List and Management Plan	Latex	Yes
Prototypes and proof of concepts	Latex	Yes
Iteration Plan	Latex	Yes
Phase Plan and Software Development Plan	Latex	Yes

Figure 9: Development Case Table