

Group 5, Government of Canada Data Portal

Group Members: Nadine Abdul-Hamid, Brett Adams, Calder Johnson, Arya Kumar, Omer Latif, Hisan Salimi, Yize Tang.

Abstract

In the modern era, data is foundational to all scientific inquiry, business decision making, and knowledge seeking. The Government of Canada offers a rich catalog of data sets, reliably describing everything from population history, economic trends, and everything in between. However, the existing interface is difficult to access, has poor user experience design, and lacks modern features such as data visualization and easy API access. Using HTML/CSS/Javascript and Firebase, our team developed a superior access point to the valuable datasets offered by the government. By focusing on excellent user experience, next level tooling, and the creation of an easy to navigate interface, we sought to improve ease of access to the data and empower business owners to make informed data driven insights to guide their endeavors.

Keywords

5 keywords that describe our project are: Data Portal Enhancement, User Experience Design, Business Empowerment, Accessibility, Performance.

Acknowledgment section

Omer Latif - Developed backend, wrote experimental setup.

Yize Tang - Managed team tracking software, wrote related work and future work.

Arya Kumar - Created documentation and UML diagrams, wrote introduction and conclusion.

Nadine Abdul-Hamid - Created documentation and UML diagrams wrote approach.

Calder Johnson - Developed front end, managed team, wrote abstract and discussion.

Brett Adams - Had minimal team contributions (identified design patterns).

Hisan Salimi - Had minimal team contributions (made 1 change to front end).

Introduction:

Imagine trying to find important information about Canada, like data on agriculture, education, economics, health, and labor. It's like searching for puzzle pieces scattered all over the place! That's the problem many Canadians face when they need government data. But what if there was a solution?

The Government of Canada wants to make accessing data easier for everyone. They're working on a project to create a new website called the Open Data Access Portal. This portal will be like a treasure trove of information, where people can easily find, use, and even upload data. The idea is to empower Canadians, researchers, businesses, and developers with valuable resources to make informed decisions and drive innovation.

Our mission is to assist the government in developing this next-generation portal. We'll create a live dashboard and user interface that allows users to collect, manage, and distribute data in various formats. Additionally, we'll provide high-quality API libraries for developers to create new applications and services.

This report will outline our journey through three project phases: Inception, Elaboration Iteration 1, and Elaboration Iteration 2. Each phase brings us closer to our goal of building a state-of-the-art tool for open data access.

Related Work

In order to better prepare for the development of the Government of Canada's next generation open data access portal, we have made a careful and extensive review of existing products, research papers, resources, and articles. These materials are related to open data platforms and tools that are used to gather insights and best practices to inspire the design and development of our product.

Generally speaking, open data sets refer to datasets that are freely available to the public, without restrictions on access or usage. These datasets encompass a wide range of information collected and maintained by governments, organizations, and individuals across various domains such as agriculture, education, economics, health, and more. The fundamental principle behind open data sets is transparency and accessibility, aiming to empower individuals, businesses, and governments to leverage data for innovation, decision-making, and accountability. By making data openly available, governments and organizations foster collaboration, encourage innovation, and enable informed decision-making by citizens, researchers, and policymakers. Open data sets are instrumental in driving social and economic development, facilitating research and analysis, and enhancing public services and governance.

Data.gov stands as a compelling example for our project. Introduced in late May 2009 by Vivek Kundra, the Federal Chief Information Officer (CIO) of the United States, Data.gov represents a pivotal initiative managed by the U.S. Government. Its primary objective is to

democratize access to a wealth of valuable, machine-readable datasets originating from the Executive Branch of the Federal Government. Offering a user-friendly interface, Data.gov hosts searchable catalogs encompassing both "raw" datasets and an array of powerful tools. Within the "raw" data catalog, users can explore datasets in various formats, including XML, Text/CSV, KML/KMZ, Feeds, XLS, or ESRI Shapefile. Additionally, the tools catalog provides links to sites offering data mining, extraction tools, and widgets, enhancing users' ability to derive insights from the available data. Search functionalities enable users to navigate datasets and tools by category, agency, keyword, or data format. Upon selecting a dataset or tool of interest, users are directed to detailed metadata and information pages, facilitating informed decision-making and usage. These distinctive features of Data.gov offer invaluable insights into structuring our portal's data repository and refining its search capabilities to maximize usability and accessibility.

By synthesizing information from above, we gained a comprehensive understanding of the current landscape of open data platforms and tools. This knowledge serves as a foundation for designing and developing an open data access portal that meets the needs of stakeholders and empowers developers to create innovative applications and services for Canadians.

Approach

The landscape of our procedure centralizes on enhancing the Government of Canada's data access portal by utilizing HTML/CSS/JavaScript and firebase technologies. The main goals are to ameliorate user experience, and enable data to be accessed in a flawless fashion. The Approach involves a simple but concise description of the methodology in terms of assumptions, project setup, and the leading to the algorithms for the prototype in which has been proposed and developed.

Our team has gathered a list of assumptions that help with the decision-making, brainstorming and the execution process. Below are the collected assumptions:

Technical Assumptions: the members of our team are skilled in the HTML/JavaScript and Firebase technologies that ameliorates data access. Otherwise there will be a result of new implementations such as training courses. This could cause delay in the series of events (timeline) and final results.

User essentials: Users search for a design that concentrates on the facility of navigation, log in , and accessing authorized data.

Stakeholder availability: stakeholder will hold responsibilities in the decision-making and give weekly assessments. If stakeholders lack involvement this can cause slow downs in the project, affecting the timeline.

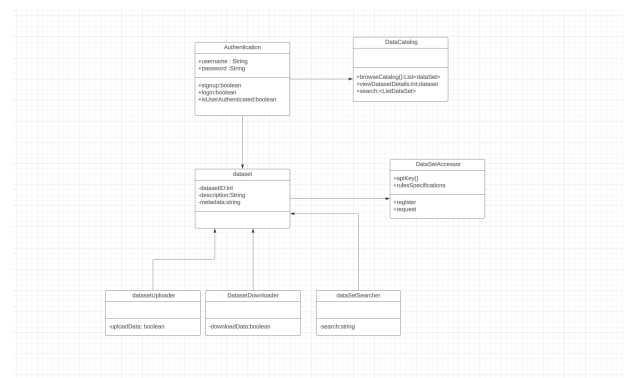
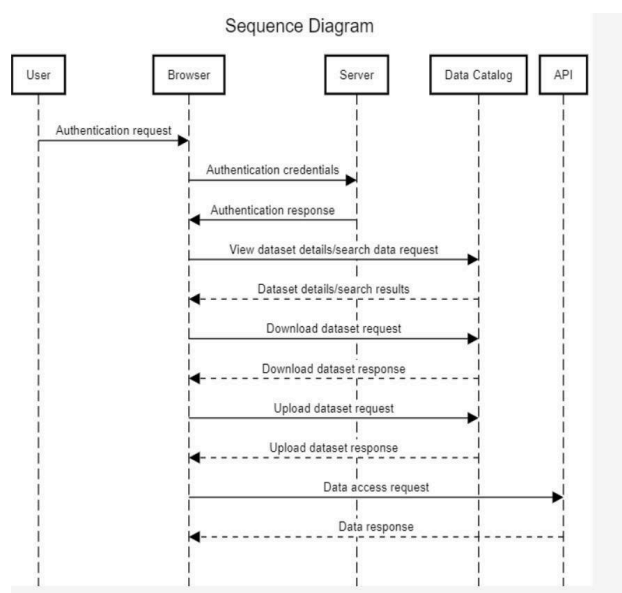
Resource Availability: A fair budget and financial support will help with all required resources and cover any unforeseen expenses. This is important for embarking people with experience, acquire tools, or implement necessary activities leading to the developments and delivery of the project.

External Dependencies: Third-party APIS or required information will be provided and accessible through the project term. Otherwise if external APIS experience data unavailable, or encounter modifications this may cause disruption in the activities that depend on those external dependencies, affecting delivery and practicality.

Regulatory Compliance: The application adheres to all rules and regulations such as privacy laws, and data protection expectations ruled by the government of Canada. If there is failure in that, this will cause legal issues such as fines or ruin the reputation of the Canadian government.

In the first part of our project setup we established timelines and discussed resources as well as the funding for those resources. In addition we created an accumulation of different features and functionality using user requirements as reference and project goals.

The next part of our set up focused on the physical architecture design of our system with creating UML diagrams which include class diagrams and sequence diagrams.



For the third part, team members initiated the frontend and backend using HTML/CSS, JavaScript and firebase. With that, our team members focused on implementing an adaptable interface with design principles that are in accordance with today's technology.

Also, user authentication and authorization features that secure crucial information are created.

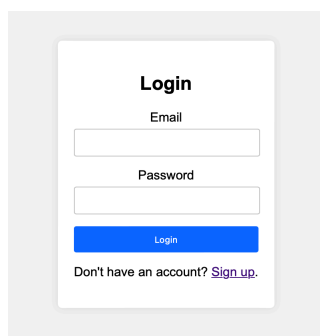
For the next part we conducted unit testing for frontend and backend to ensure soundness and functionality in the application. Also, we ensured user authentication and authorization features to verify data privacy and security.

The last part in our project setup is the deployment. Our team deployed the prototype to the pre-production environment for user acceptance. During this phase any issues or bugs identified during testing were addressed by team members to ensure the prototype matched standards.

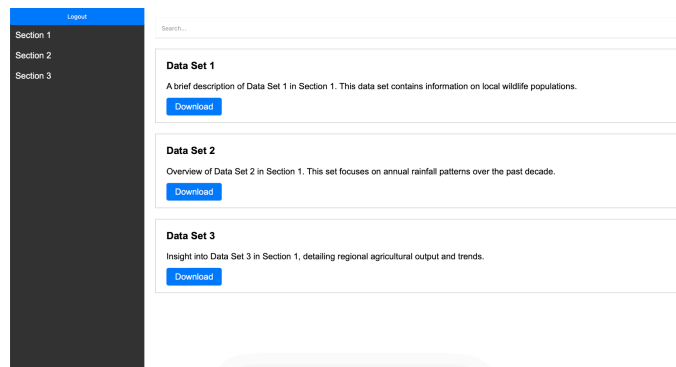
Experimental setup or Demonstration

The application features a multi-faceted interface that includes a login page, a signup page, and a user dashboard for navigating and downloading datasets. This setup was constructed using a combination of HTML, CSS, and Javascript for the frontend, with Firebase serving as the backend infrastructure. The choice of these technologies was guided by our aim to create an intuitive, secure, and easily navigable interface for users, enhancing their experience while accessing government datasets.

The login and signup pages serve as the entry point to the application. Both pages are designed with a focus on simplicity and user-friendliness. The signup page requests essential information from the user, ensuring a straightforward registration process. Once registered, users can log in through the login page, which features a clean layout with fields for username and password, accompanied by a submission button. These pages incorporate basic validation checks to improve data integrity and user experience.



The centerpiece of the application is the user dashboard, a carefully designed interface that allows users to navigate through various datasets. The dashboard layout is intuitive, ensuring that users can easily find and access the information they need. It features a categorized list of datasets available for download, with each dataset accompanied by a brief description, its size, and the option to download. This design prioritizes user empowerment, enabling them to make informed decisions based on the dataset details provided.



The dashboard also includes search functionality, allowing users to quickly locate specific datasets by keywords. This feature was developed to enhance accessibility and ensure that users can efficiently find data relevant to their needs without navigating through the entire catalog.

For the purpose of demonstrating the application's functionality, we conducted a series of user interactions simulating typical use cases. These interactions showcased the seamless flow from signup to dataset navigation and download, highlighting the application's efficiency and ease of use. Feedback from this demonstration emphasized the application's user-friendly design and the value of having a centralized portal for accessing government datasets.

The experimental setup and demonstration underscored the project's success in creating a more accessible and intuitive platform for data retrieval from the Government of Canada.

Discussion

The project underwent many changes during the development process. The final output of elaboration was a robust and reliable data portal with a good degree of usability, but with a smaller amount of features than we had initially planned. In this section, the development process will be explored, including our management of expectations, key risks, and technology choices.

When risks, technology choices, and requirements were initially analyzed, an ambitious plan for the elaboration iteration was settled on. The backend was projected to be built from scratch using Python, Flask, and PostgreSQL, and the frontend with ReactJS. A large backlog of desired features was drawn up, such as user authentication, API integration, data visualization tools, advanced searching features, and a range of non functional considerations such as a multilingual interface. Key risks were considered, such as the need for scalability, policy considerations, and security of user data.

Afterwards, we proceeded to the first elaboration phase. Upon further consideration of our vision, we realized that in order to implement scalability, data integrity, and user data security, building the server side architecture from scratch as was planned would not be suitable due to the limited number of developers we had available (only 2 developers). Firebase was instead opted for, and the wide range of prebuilt and configurable components was leveraged in order to create a robust and powerful backend. This ensured scalability would be seamless moving forward, the database would be highly secure, and the application would be very reliable due to the high quality of Firebase components used.

The user interface was also constructed during this phase. Another change was implemented in this stage, ReactJS being swapped for a straightforward approach with HTML/CSS/Javascript. This was also done due to the limited number of development resources.

At the end of the 2 week timeboxed first elaboration iteration, the result was a complete prototype of the application with core features built. Software verification and validation was then conducted. Extensive user testing was performed, ensuring all functional requirements implemented were correct and complete. Bugs were discovered with searching functionality and user authentication, and fixes were implemented. Non functional requirements were also tested, ensuring the application adequately met expectations for performance usability and reliability.

Certain features that were planned to be integrated into the application were not constructed in time, such as data visualization, data uploading, and API access to datasets. These features were placed on the backlog, to be completed in future iterations. The creation of a robust set of core features with high quality and low failure rate was prioritized over the implementation of many features that would not be able to be fully tested and validated in time.

Important takeaways from the development process include the importance of fully considering development resources before planning an iteration. The limited number of active developers lead to more features being planned than could be implemented in time, and lead to a number of last minute technology choice changes that could have been detected earlier leading to a more streamlined development process.

Conclusion

Looking back, we set out to make it easier for Canadians to access government data and use it to make informed decisions. We've achieved a lot with our Open Data Access Portal:

1. Accessibility: We've centralized access to a wide range of datasets, making it easier for everyone to find the information they need.

2. Empowerment: By providing complete and timely data, we've empowered users to make informed decisions and drive positive changes in various fields.
3. Collaboration: Our platform has facilitated collaboration between data providers, users, and developers, fostering creativity, knowledge exchange, and partnerships to address complex challenges.
4. Economic Growth: Opening up government data has led to economic growth, job creation, and innovation in different sectors, contributing to the nation's development.

While we've made significant progress, there are still some limitations and challenges:

1. Resource Constraints: Limited time and budget have impacted the scope and development pace of our project.
2. Technical Challenges: Ensuring data integrity, scalability, and performance remains a challenge, especially with the growing user base and data volume.
3. Compliance: Adhering to regulatory standards and data governance policies requires ongoing effort to ensure legal and ethical compliance.

Despite these challenges, our prototype offers several benefits:

1. User-Centric Design: Our platform prioritizes user experience, making it easy for users to find and utilize data effectively.
2. Data Visualization: The built-in data visualization tool allows users to create charts, graphs, and maps from datasets, enhancing data exploration and understanding.
3. Security: Robust access control mechanisms ensure data security and privacy, protecting sensitive information from unauthorized access.

Overall, our Open Data Access Portal represents a significant step towards promoting transparency, accountability, and innovation in Canada's data ecosystem. While there's still work to be done, we're proud of what we've accomplished and excited about the positive impact it will have on the country's future.

Future Work

In addressing the multifaceted requirements of this project, our approach begins with meticulous API design. We will craft a set of RESTful APIs that not only facilitate seamless data access and management but also adhere rigorously to industry standards and best practices. This involves defining endpoints, request/response formats, authentication

mechanisms, and error handling protocols to ensure both efficiency and security in data transactions. Concurrently, we will establish an updated development workflow that encompasses version control, branching strategies, and comprehensive issue tracking. Embracing modern methodologies, we will implement continuous integration/continuous deployment (CI/CD) processes to automate testing and deployment pipelines, ensuring rapid iteration cycles and robust quality assurance measures.

To uphold code integrity and readability, we will define coding standards and conventions that prioritize maintainability and scalability. Concurrently, we will also institute robust code review practices to foster collaboration among team members and mitigate potential issues early in the development lifecycle. Our commitment to excellence extends to gathering feedback from stakeholders, users, and developers, which will be systematically analyzed to identify areas for improvement and guide the prioritization of new feature requests.

User experience lies at the forefront of our concerns, prompting us to refine and enhance interfaces based on user feedback and rigorous usability testing. Security is paramount in our architecture and development process, with encryption, access controls, and proactive threat monitoring integrated seamlessly to safeguard sensitive data and prevent unauthorized access.

Anticipating future growth, we will conduct thorough scalability testing and updates to ensure our solutions can gracefully accommodate a larger and more diverse user base. Ultimately, our deliverables will encompass a comprehensive API design document detailing endpoints, methods, parameters, and payloads, alongside documentation outlining our development strategy, CI/CD setup, coding standards, and code review procedures. Additionally, our commitment to delivering exceptional user experiences will manifest in tangible improvements and enhanced functionalities driven by user feedback and usability testing insights.