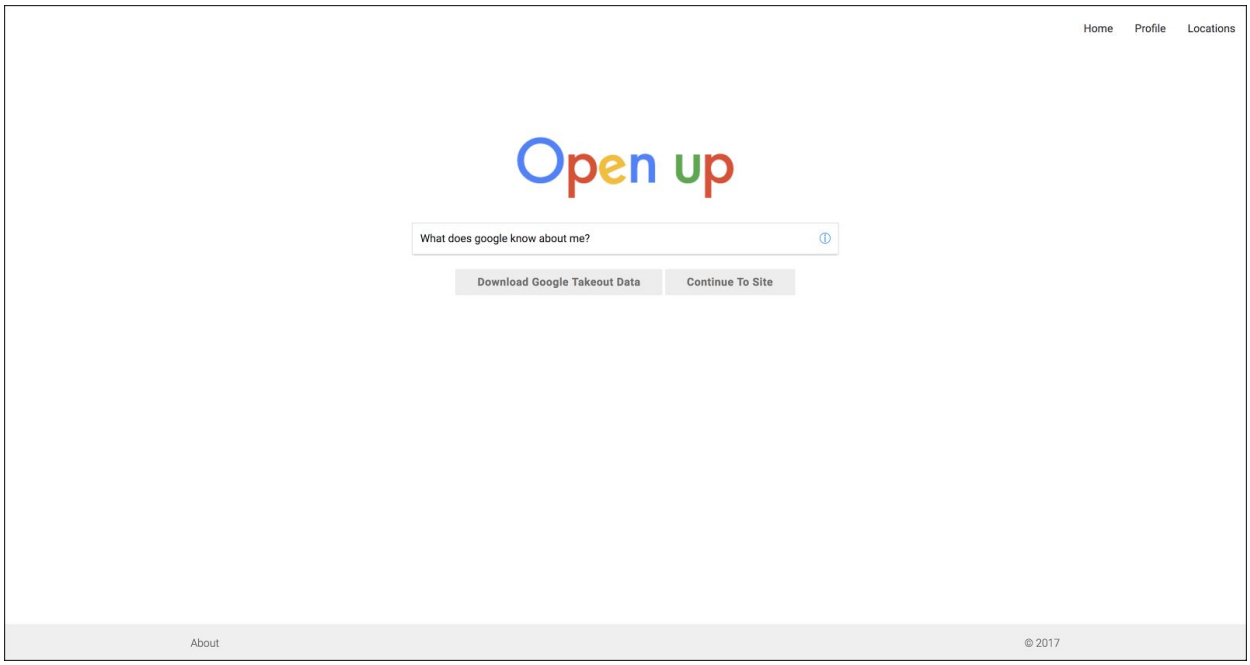
Open up

Google never forgets.



Ivy Calvert

MDDN412
Victoria University of Wellington
2017

INTRODUCTION

'Open Up' is a web based project that requires users to interact with their Google Takeout data. With the overall aim to highlight how much information users disclose online while logged into their Google account, 'Open Up' focuses on the connection of this information.

This project is live online at: https://ivycalvert.github.io/openup/openup-home.html

The code is viewable at: https://github.com/ivycalvert/openup

RESEARCH

Understanding how personal information is handled online was vital to the development of this project. Initially, this required research into multiple aspects, primarily this research focused on the privacy paradox phenomenon, "privacy zuckering", and the data broker industry (definitions below).

PRIVACY PARADOX: The privacy paradox refers to the phenomenon that has emerged when internet users privacy concerns are not reflected in their actions.

"PRIVACY ZUCKERING": This is when a website tricks users into sharing more personal information publicly than intended.

DATA BROKER INDUSTRY: Data brokers are companies that collect information from online activity, public records, and purchase history to sell it on to other companies for marketing purposes.

It was also important for the development of this project to relate it back to the defined Human-Computer Interaction (HCI) principles.

HUMAN-COMPUTER INTERACTIONS

The study of how people interact with computers is called human-computer interactions (HCI). Included with HCI is the extent computers have or have not been developed for the interaction with people. For this project, the three paradigms of HCI were investigated. The three paradigms are as follows: 1st Paradigm: Human Factors, 2nd Paradigm: Classical Cognitivism, and the 3rd Paradigm: Embodied Cognition.

Each paradigm aims to achieve a goal:

1. The 1st paradigm aims to optimise the fit between man and machine.
2. The 2nd paradigm aims to optimise the accuracy and efficiency of the transfer of information.
3. The 3rd paradigm has the goal of supporting a situation action in the world.

While this project exists within all three paradigms, it is primarily situated within the 2nd and 3rd paradigms. The 2nd paradigm applies to this project as the data visualisations created from the complex data files. This simplifies and presents the data in a clear manner that is easily understandable by the user(s). The 3rd paradigm applies to the social impact this site has. This relates to the impact and awareness the user(s) experience upon interacting with the site. These two paradigms work together to create a project that focuses on optimising how the brain reacts (2nd) and the overall interpretation of the project (3rd).

PROJECT OVERVIEW

The development process for this site followed a strict milestone process. Step one included the conceptual stage, step two was the creation of a proof of concept, step three (and the stage this project is at currently) was the development of a fully functional and aesthetic prototype (this is where the project is currently situated, with a functional demo live online). This strict methodology enabled the final demo of this project to be a fully functional prototype that provides the foundation for continued development on this project.

STAGE ONE: CONCEPTUAL

Early in the concept and ideation stage, there were two points of interest that drove the direction of this project:

1. The difference of how people interact online when things are private versus when things are public
2. Data visualisation

This ultimately drove the project in the direction seen today. Linking these ideas together, the early project concept evolved into a browser plugin that would create data visualisations of all the information the user would post online. Precedents that helped shaped this evolution included browser plugins ‘Ghostery’, ‘Lightbeam’ and ‘Collusion’ as

these extensions focus on internet anonymity and/or user awareness. An additional precedent that helped shape the project was the artistic project ‘Gate Peepin’.

During this stage however, it was established that in order to get access to the more personal data (which had potential to create more interesting visualisations) it would profit the project to move away from the plugin concept and instead move towards a website. This was decided as it was decided that the project would focus on the personal data users could retrieve from Google Takeout. This was due to the interest of the information retrieved from Google and also the potential to create dynamic visualisations from this data that would allow for the interaction of multiple data types.

STAGE TWO: PROOF OF CONCEPT

This stage was important for the development of the project as it established the conceptualised functionality was plausible and also allowed for the exploration into the technologies needed to execute the visualisations. The focus during this early development was placed on the site functionality rather than aesthetics. This early functionality included the core aspects that would later be required for the site, building the foundations for the project. The proof of concept included:

1. Turning raw data files into simplistic visualisations (with variations dependent on data)
2. Successful integration of the Google Maps Javascript API
3. User ‘Location History’ data interactions with the implemented API

While most of this code was rewritten during the next stage of development, this stage enabled the testing of the functionality, ensuring each element functioned as planned. The proof of concept was also set up with each “functional” element on separate pages of the website, allowing for controlled testing of these elements.

STAGE THREE: PROTOTYPE

Due to time restrictions on the development of this project, the prototype is the final stage which is reached (additional stages would include moving into the ‘future’ goals and developing a Final Version of the site). Early on in the development, a clear MVP was defined for the prototype:

1. Live online/accessible to users
2. Functionality of two data sets (Profile information and Location History)
3. Consistent aesthetics with Google
4. 100% legal (no storage of user information)

It was important to set clear goals for the prototype functionality (further discussed below) due to the time limitation of the development. This stage also enabled a revisit of the Proof of Concept stage to remove any hacks and bugs, tidying the code and enabling smoother integration of more functionality as this stage progressed. This stage was mainly used to assess the overall plausibility of the site (answering the question of “is the whole thing going to work?”) and test the response of the site in a demo setting where users could begin to interact with the functionality on offer.

The prototype included more than the functionality defined in the MVP, including:

1. Integration of up to three data sets -- one more than the MVP (Profile information, Location History and Search History)
2. Two data sets interacting to create a dynamic visualisation

During the development of the project, an edge case functionality was also implemented. This functionality arose when it became apparent that I, as the developer, did not have enough information to create one of the data sets (location history) required for functionality (mentioned below). The solution to this was to use Location History from one source, and Search History from another. While initially for the Proof of Concept stage this was effective, as the functionality was expanded during the Prototype stage difficulties arose. This was due to the functionality of the intended visualisation dependent on the timestamps of the data. Having come from two different sources, the timestamps were wildly different for each data set. This resulted in the inability to test functionality and bugfix while the data was (in this state) unuseable.

As a solution, a “new” functionality arose (initially intended for development purposes but soon developed into a feature of the site). This new functionality allows for users to still interact with the site even if their data has large discrepancies (which should only occur when the data is from two different sources, however to provide for the edge case situations this remained in the prototype). This functionality allows for users to ‘falsify’ the timestamps on their Search History data so correlations can be seen against the Location History data in the visualization (as shown below in more detail).

DIFFICULTIES

The difficulties of this project can be separated into two different categories: project difficulties and the difficulties experienced personally while developing the final prototype.

Project difficulties are defined as those that limited and had to be considered while

developing the site functionality. These included:

1. Gaining access to the personal information required
2. Allowing for multiple data sets on a single page without the need for server-side data storage
3. The lack of the correct data sets required to test and prove functionality (no Location History)

These difficulties were overcome through the careful implementation of these solutions (the numbers of the problem correlate to the numbers of the solutions; i.e. the solution to difficulty one is solution one):

1. The use of Google Takeout data
2. Local Storage implementation
3. Kind donation of this data set (Location History) from a friend allowed for functionality testing, bug fixes and the trial of the site during the proof of concept/prototype demo stages of development

Personal development difficulties relate to the limited knowledge of the code required from this project. Before development, much of the required technologies were unknown, resulting in the need to learn it during development in order to ensure the MVP was reached by the prototype stage. For the purpose of this project, the technology that was required to learn include:

1. Javascript
2. Google API implementation
3. JSON
4. Materialise css framework

While these limited the speed of development, overall they did not limit the planned functionality of the project.

AESTHETICS/DESIGN

The overall aesthetics and layout design of the project were decided when the decision to work with only Google Takeout data was finalised. It was important that the user would have a smooth experience while using the site, and as process to download Google Takeout data had multiple steps (including visiting specific Google sites), this meant that the ‘Open Up’ project would have to mirror the aesthetics seen on these pages.

Rather than building on these aesthetics and letting them shape the development of

unique aesthetics for this project, ‘Open Up’ embraced them. This project directly borrows aesthetics and layout design from the classic Google pages most internet users are familiar with.

RELEVANCE OF THE SITE

This site encourages curious minded internet users to interact with the data Google collects and stores while signed in to their Google account. This encourage awareness and increases knowledge about how much data is collected. As this information is seen through visualisations of interacting data, it enables users to see how their data can be used.

While this site will benefit from expanding on the data sets and interactions of these data sets with one another, the site is beginning to show how it can be used to create an interest for users to visit and interact with (much like the site ‘applymagicsauce’).

FUNCTIONALITY

TECHNOLOGY USED

This project implements a range of code languages and technologies available. These include:

1. Google Maps Javascript API
2. Javascript
3. HTML
4. CSS (standard and materialise css framework)
5. JSON
6. Google Takeout

All these technologies working together helped to create the range of interactions and visualizations seen throughout the website.

GOOGLE TAKEOUT

Google Takeout is a service provided by Google which allows users to download all the data Google has collected during the entire history of their account with Google. This services collects all this data into files that can be downloaded by the user on request.

For this project, it is required that users download this data from Google, ready to upload the specified files to the correct ‘Open Up’ pages in order to interact with the services

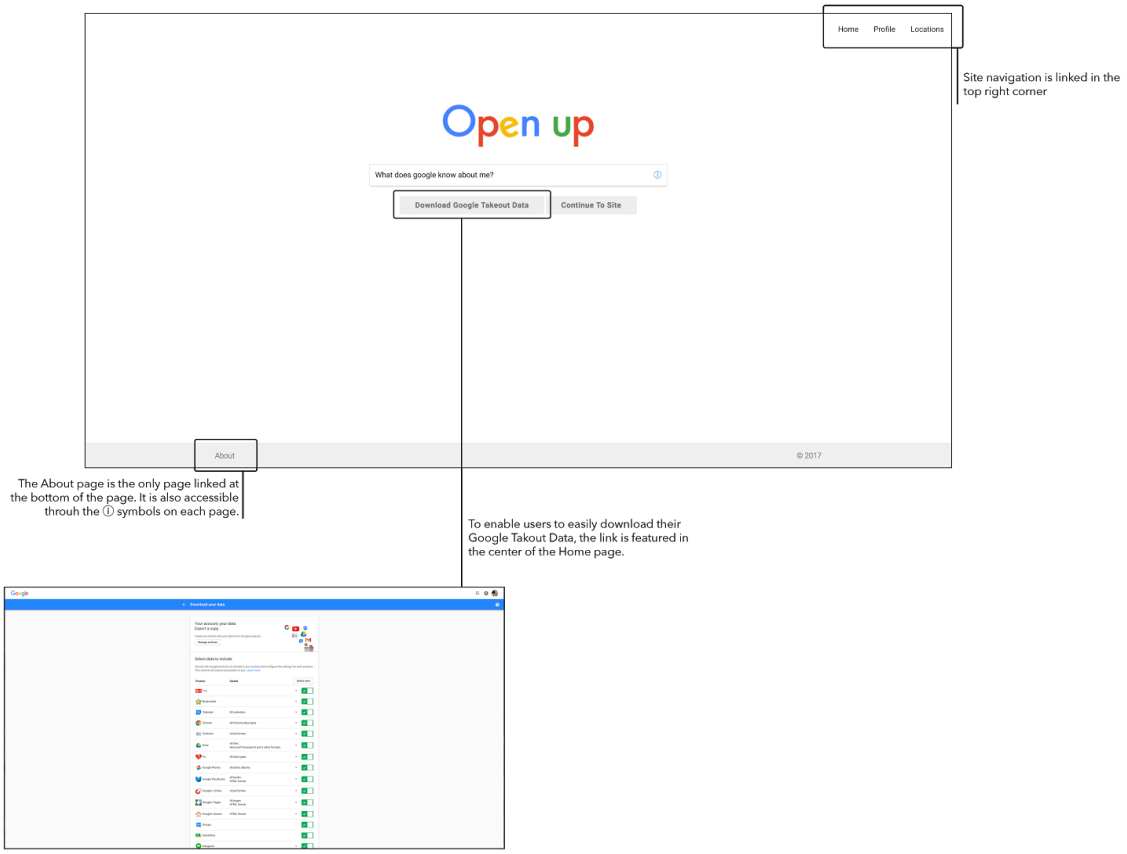
offered by this website.

THE SITE

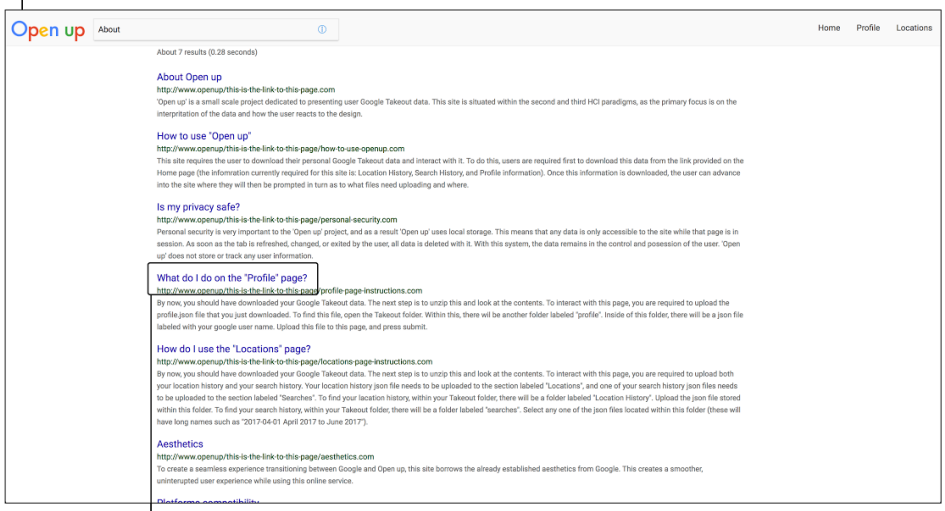
The site consists of four main pages:

1. Home
2. About
3. Profile
4. Locations

While the Home and About pages have limited interactions due to their more traditional nature, the Profile and Locations pages are were the true functionality of the site can be seen. See the below annotated screenshots of the website for explanations of the functionality of each page (pages are in order of the list above).

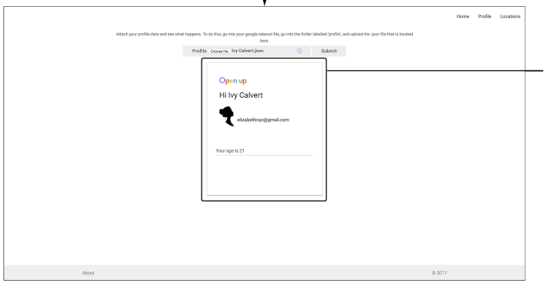


Structed like Google search results, the about page functions as both as an about page and also as a help page

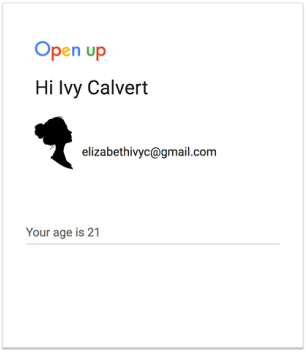


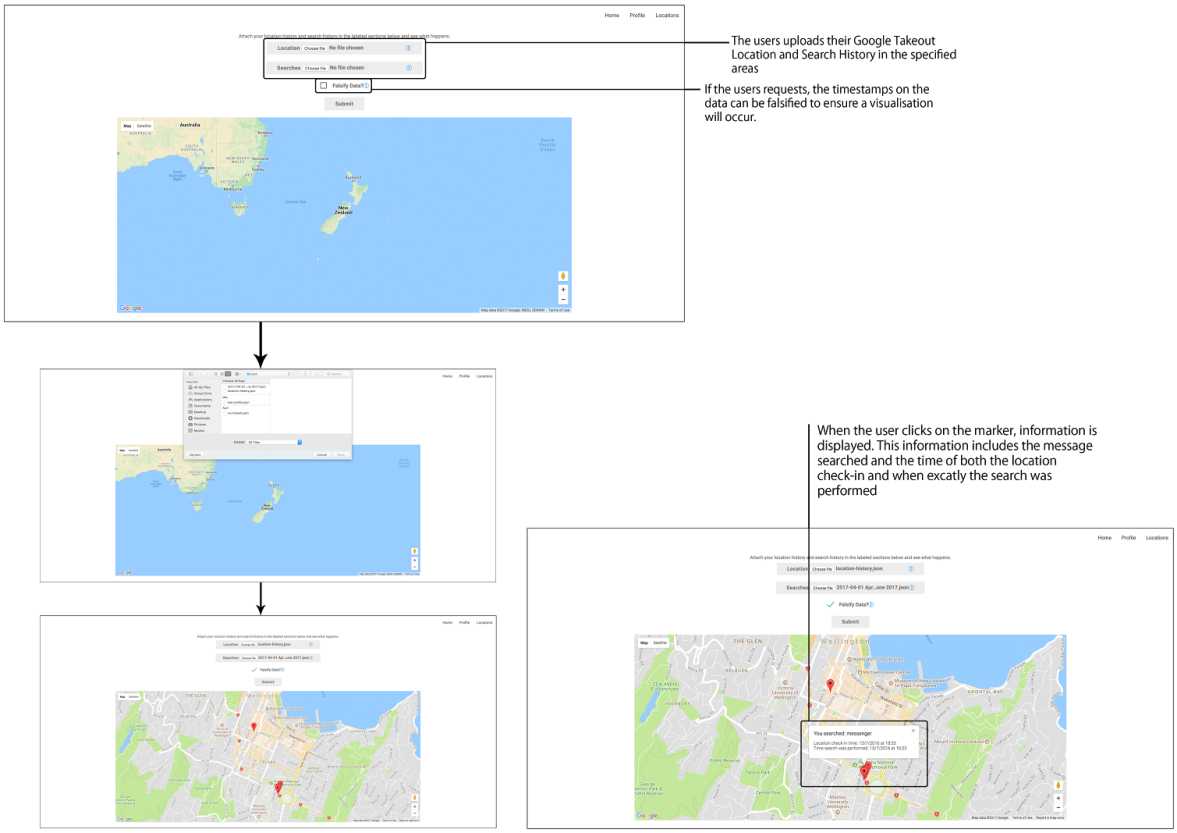
Each page has its own section on the About page. When the ⓘ icon is clicked on the Profile page, the About page links to this section specifically so the user sees exactly what they need immediately. (This theme is carried across the site, with other pages linking to their sections. Such as the Locations page links to the "Locations" about section).

The first page the users are directed to is the simplier of the two visualisations. This page simply requires users to upload their 'Profile' data (a json file).



Once the file is uploaded, a graphic loads on the page, determined by the information within the data. This page simply presents back to the user their identity information linked to their Google account *(see right for a scaled version of this graphic)*





FUTURE

While this project is currently restricted by the small number of interactions, future development of this project would see a lot more functionality. The main future aims for this project include:

1. More data types implemented
2. More advanced visualisations that encourage more interaction from the user(s)
3. Simplified data retrieval methods (this would enable less technically literate users to engage with the website)
4. Compatibility with other companies, not just Google (this would include a switch in project aesthetics)
5. Refinement on page aesthetics before visualisation are loaded

REFERENCES

Kielczewski, B. (n.d.). Apply Magic Sauce - Prediction API -. Retrieved June 21, 2017, from https://applymagicsauce.com

Evolving the Google Identity. (n.d.). Retrieved June 22, 2017, from https://design.google/library/evolving-google-identity/

Google Design. (n.d.). Retrieved June 22, 2017, from https://design.google/

About - Materialize. (n.d.). Retrieved June 22, 2017, from http://materializecss.com/about.html

Firms Are Buying, Sharing Your Online Info. What Can You Do About It? (n.d.). Retrieved June 22, 2017, from http://www.npr.org/sections/alltechconsidered/2016/07/11/485571291/firms-are-buying-sharing-your-online-info-what-can-you-do-about-it

April 16, W. T. on, & business, 2015 WebpageFX is a full-service Internet marketing company that offers the perfect solution for your. (2015, April 16). What Are Data Brokers - And What Is Your Data Worth? [Infographic]. Retrieved June 22, 2017, from https://www.webpagefx.com/blog/general/what-are-data-brokers-and-what-is-your-data-worth-infographic/

Types of Dark Pattern | Dark Patterns. (n.d.). Retrieved April 20, 2017, from https://darkpatterns.org/types-of-dark-pattern