User Experience Optimization in Geo-portals for Data Discovery

Date: April, 2018

Student: Han Liu

Supervisor: Professor A-Xing Zhu

Committee: Professor Qunying Huang, Professor Song Gao

University: University of Wisconsin - Madison
Abstract

This research focuses on user experience of Geo-portals for citizens to discover geospatial data online. The misguided design ideas in the last decades encourage engineers to focus on their own behaviors rather than citizens’ experience. Therefore, most geo-portals are imitating desktop GIS data management software, such as ArcGIS, which is not designed for data discovery, resulting in that current geo-portals with cluttered information are too technical for citizens to discover one piece of geospatial data.

According to previous studies, citizens want a single piece of geospatial data instead of interacting with the geo-portal. As a result, a better content architect was proposed to satisfy citizens’ goals. The question is how to offer an overview of available geospatial data across political boundaries with one click of a button for the absolute beginners.

With new technologies and user behavioral studies, a new method was proposed in this thesis to involve more citizens in the open data movement. The new approach integrates existing geo-portals into one platform which has the ability of interacting with the APIs of the existing portals. The design of the interface to this platform is centered on the easiness of data discovery for non-specialist users of geospatial data.
Analysis shows that the developed platform is easy, fast, and intuitive for users to discover geospatial data. Future study should focus on the automation in adding new data discovery gateways or data depositories into the new platform.
Acknowledgement

This master thesis is the final product of my Masters in Geographical Information & Cartography. During my second year, I had the chance to design this geo-portal, and apply understandable intuitive design and user experience principle for people without geographical knowledge. I want to thank Professor A-Xing Zhu for this opportunity.

During this study, I became more aware of the importance of geospatial data to citizens. Therefore, I focused my thesis on UI/UX of a geo-portal that should focus on simplifying data discovery, especially to citizens. Thanks to my supervisor Professor A-Xing Zhu for the opportunity to work on this project. I want to thank him for all the help and advice during this period. Also, my thanks go to the colleague, Guiming Zhang and YuYing Chen, who gave up their time to help me during this research. Without their help this research wouldn't be possible. Finally, I want to thank my committees, Professor QunYing Huang and Professor Song Gao for supporting me during this period.

April, 2018
# Table of Contents

1. Introduction .................................................................................................................. 8
   1.1 Objectives: Simplification is the priority ................................................................. 8
   1.2 On geo-portal ........................................................................................................... 10
      1.2.1 The definition of geo-portal .............................................................................. 10
      1.2.2 The history of geo-portal ................................................................................ 11
      1.2.3 The importance of geo-portal ......................................................................... 12
   1.3 Organization of the thesis ......................................................................................... 12

2. User experience issues of Geo-portals ......................................................................... 13
   2.1 Accessibility issues: heterogeneous ........................................................................ 13
   2.2 API issue: lacking machine-readable formats for automation .............................. 14
   2.3 Usability issue: cluttered user interface ................................................................. 16

3. Current solution and research question ....................................................................... 20
   3.1 Lacking data discovery across political boundaries ............................................... 20
   3.2 Lack of APIs integration focus on discovery ........................................................... 21
   3.3 Unfriendly UI design .............................................................................................. 22

4. Methodology .................................................................................................................. 23
   4.1 Define users’ goal ..................................................................................................... 23
   4.2 Integrating available APIs ..................................................................................... 25
   4.3 Intuitive interface design ....................................................................................... 26
      4.3.1 The minimal information required to conduct a search ................................. 27
      4.3.2 Search geospatial data across political boundaries fast ............................... 28
      4.3.3 Make Intuitive UI that does not require “Help” .............................................. 28

5. Results ........................................................................................................................... 35
   5.1 Easy, fast, intuitive, informative access ................................................................. 35
   5.2 Low-cost open source application ......................................................................... 41

6. Discussion ....................................................................................................................... 42
   6.1. The uniqueness of thesis ...................................................................................... 42
   6.2 Thesis evaluation ..................................................................................................... 43

7. Future research .............................................................................................................. 44
7.1 Multiple formats, open standards ........................................................................... 44
7.2 Put spatial first ........................................................................................................ 45
7.3 Recommender system ............................................................................................... 45

Reference ..................................................................................................................... 47
Table of Figures

Figure 1.1 The landing page of City of Edmonton's geo-portal. ........................................9
Figure 2.1 The list of available format in Data.gov when user search “soil” as the keyword. ................................................................. 15
Figure 2.2 The Washington State Department of Transportation geo-portal default layout. .................................................................................................................. 17
Figure 2.3 The interface of ArcGIS ................................................................................................................................. 18
Figure 2.4 The toolbox dropdown menu in ArcGIS. ... Error! Bookmark not defined.
Figure 2.5 The "spatial Join" tool in ArcGIS ..................................................................................................................... 19
Figure 3.1 The landing page of ScienceBase.com ........................................................................................................ 23
Figure 4.1 Parsing JSON data from ScienceBase.com with Ajax .............................................................................................. 26
Figure 4.2 The Z-pattern of human eye path. .................................................................................................................. 29
Figure 4.3 The wireframe layout of proposed landing page .................................................................................................. 30
Figure 4.4 The logo in blue and green .............................................................................................................................. 31
Figure 4.5 User interface of landing page .......................................................................................................................... 32
Figure 4.6 User interface of landing page with hidden elements .......................................................................................... 32
Figure 4.7 The alert message when users forgot to define the bounding box ........................................................................ 33
Figure 4.8 Auto-complete search box dropdown menu .................................................................................................... 34
Figure 4.9 The presentation of the search results ............................................................................................................. 35
Figure 5.1 Figure before search ........................................................................................................................................ 36
Figure 5.2 Figure during search ..................................................................................................................................... 37
Figure 5.3 Figure after search ........................................................................................................................................ 38
Figure 5.4 Popup box of summary on discovered data ..................................................................................................... 38
Figure 5.5 Hidden data repositories and sharing buttons in the lower left corner .................................................................. 39
Figure 5.6 The map on the maximal zoom level ............................................................................................................. 40
Figure 5.7 The button icon for social media sharing and data repositories ........................................................................ 41
1. Introduction

1.1 Objectives: Simplification is the priority

The aim of this research is to simplify geo-portals by integrating APIs and redesign user interface to minimize the workload of data discovery for beginners. Currently, the effort required from users who are new to the use and application of geospatial data to discover a geospatial data online is prohibitively high (Guralnick and Neufeld, 2005). First, there are large numbers of geo-portals available from federal and local governments, universities, NGOs and commercial vendors (Budhathoki and Bruce, 2008). With different publishing policies, every geo-portal has different interfaces. Going through them is not only time consuming but also frustrating.

Secondly, some geo-portals are too technical and complicated from a usability perspective. They focus on adding more map layers and they pay less attention to user experience and data discovery. Take the City of Edmonton as an example. In the below picture, its geo-portal has 8 map layers, one drop-down menu for selecting one out of many neighborhoods. There is no search keywords option, so finding a piece of geospatial data from 8 map layers requires users to go through every layers which is time-consuming. According to Allen Glen's user research, as cited by Timoney (2012), the average local government geo-portal visit time is one minute and forty-three seconds. Therefore, once the page is
loaded, if geo-portal cannot deliver results in 2 minutes, users will most likely leave.

![Image of City of Edmonton's geo-portal landing page]

*Figure 1.1* The landing page of City of Edmonton's geo-portal.

The priority of geo-portals should be making data discovery easier, but geo-portals are more focused on building tools for sorting, grouping, graphing geospatial data. Most citizens do not conduct any geo-spatial data analysis online. They simply wish to accomplish the task of discovering geospatial data of their interests. According to researchers from Denver City's Web GIS team, 65% of users land on geo-portal from Google with a single topic in their mind which is finding the data they need (Timoney, 2012). Therefore, even analyzing data is important, the priority of this work is to help users find the data they need in the first place.

From a user experience perspective, having a single point of access to see what data is available is better than giving citizens too many details and options
at first. For that reason, this thesis suggests integrating geo-portal APIs into one place to reduce the amount of time that users need to spend on each portal and the unnecessary interactions on each interface. Therefore, any functions, features, icons, and text messages that distract or confuse users should be avoided. Hence, unless a function can benefit most users and fit the core purpose of data discovery, taking time to build one is unnecessary.

To further simplify geo-portals, redesigning the user interface is necessary. Less detailed and information-rich interfaces allow users to retrieve geospatial faster. Without help, citizens will be frustrated by features on the portals they do not understand. Since data discovery is displaying the most relevant information in the least amount of time, spending time to understand features degrade user experience by watering down the true value of geo-portals. In addition, users simply leave if they need to figure out what part of the page is most valuable from the sea of information provided by geo-portals. Therefore, the goal here is to build an intuitive interface that is familiar, comfortable and predictable with minimal elements.

1.2 On geo-portal

1.2.1 The definition of geo-portal

The word portal comes from Latin word porta, which means an entrance point (Maguire and Longley, 2005). There is a general portal like Yahoo and
specialized one like geo-portal, such as ScienceBase.com, which is a website where the discovery of geographic content is a primary focus (Tait, 2005).

Geo-portal as a website is implemented using three components: a front-end website to demonstrate data; a back-end web service to publish functionality (like map rendering, address geocoding, name searching); a data managing software to hold the spatial data (both raster and vector data). Taken together, geo-portal provides the ability for searching, mapping, publishing geographic information (Tait, 2005).

1.2.2 The history of geo-portal

There are hundreds of geo-portals on the Internet at present, which starts with a fuzzy ever-evolving concept, spatial data infrastructure (SDI), that was created in 1993 by the US National Research Council (Maguire, 2005). The key idea of SDI is a catalog of metadata that can manage data and other resources using time, space and other attributes.

In 2002, the second generation of SDI, Geospatial One-Step (GOS), was initiated to create a web-based one-stop access to inventories of all available maps, data, and spatial services. As of April 2004, GOS had 5622 users with 100 page views each person per week; there are 305 publishers and 9672 publications. In the first six months, most viewed catalogs are imagery base maps (15.5%), administrative and political boundaries (10.6%), and geological and geophysical data (10.2%). In 2011, GOS was closed and data was moved to a
broader open data site Data.gov where API (application program interface) is available for the public (Maguire, 2005).

1.2.3 The importance of geo-portal

According to a McKinsey’s study, open data could allow for a combined $3 trillion a year in savings for government agencies (Manika, 2013). In addition, citizens and volunteer geospatial information (VGI) can also deliver value and insight to non-GIS beginners. There is no shortage of geospatial data generated, the problem currently is that no method is efficiently connecting those who with data to those who need it, particularly the beginners.

A significant purpose of geo-portal is to facilitate collaboration among public who are interested in the transparency of government and market. Transparency requires data to be open and usable. However, one of the largest barriers to open geo-data is that portals are not very usable for beginners (Tait, 2005). Most geo-portals focus on the implementation of complicated functionalities which are oriented for researchers and experienced users of geospatial data (Resch, 2013). As the need has grown for non-GIS users, functionality alone will not ensure the success.

1.3 Organization of the thesis

The first chapter is the background: the basic principles, the issues, and the implementations. Chapters 2 - 4 explain and justify why a new method is needed.
In chapter 2, specific user experience problems were outlined. Chapter 3 provides an overview of available methods. Chapter 4 lists research question and sub-questions. Chapter 5 describes the implementation that improves the user experience. Chapter 6 evaluate the results of the new method. Chapter 7 give a summary and discussion. Chapter 8 offers recommendations for improvements in the future.

2. User experience issues of Geo-portals

2.1 Accessibility issues: heterogeneous

Geospatial data repositories are packed in different geo-portals all over the Internet. More agencies are publishing the data in different formats across all levels and branches. However, there is no central place to find those data repositories, because most geo-portals were designed for governments and large institutions to manage data, not for the public to find data. According to the deputy director of digital publishing at the office for National Statistics, Laura Dewis, who believes that portals address the problems when they hold a specific dataset for a very specific group of users (Whitfield, 2017). But portals do not address the problems for the general public to discover data across a lot of domains. Because when people need to have an overview of available data repositories, they always have to start with Google instead of portals (Whitfield, 2017).
Since there are a very large numbers of portals on Google, users who do not have enough GIS knowledge will most likely not find the right portal for themselves in a short time. The most frequent questions GIS professionals get from non-GIS users are: “What data do you have?” and “Where can I find them?” Since geo-portals are important and not cheap to make, the future of Geospatial data should not be in the hand of GIS professionals only. Citizens should have a central place where they can have an overview of available data in short time.

2.2 API issue: lacking machine-readable formats for automation

To build a central place for citizens to search geospatial data, API (Application Programming Interface) integration is the correct solution. API integration can be imaged as a communication system where each data repository is linked to the central place. When users search for a keyword in the central place, the central place asks each repository for an answer. In this way, users will save time by not going through each repository. The problem is that currently, most data repositories are not possible to be linked because their data formats are not machine-readable.

Most geo-portals publish their data only in human-readable formats, such as HTML and PDF. For example, Data.gov is the clearinghouse for data including geospatial data. It is the largest governmental data portal in the US. However, the director of the US Open Data Institute, Waldo Jaquith says, as cited by Worrall...
(2014), almost no data in Data.gov is in a format that is genuinely usable by web developers, businesses or interested parties, and he believes “data.gov is a haphazard collection of data”. For example, if users search for “soil” as the keyword in Data.gov, the return results are mostly in HTML, BIN (a format for webpage) and Text format. In the below picture, the available formats of the return results are listed. Inside the parentheses is the number of datasets that were published in each format. As can be seen, only very small amount of data is published in JSON format.

![Available Formats in Data.gov](catalog.data.gov/dataset?q=soil&sort=score)

*Figure 2.1 The list of available formats in Data.gov when user search “soil” as the keyword*
As mentioned in the above, only few geo-portals publish data in machine-readable formats. This is a major barrier to API integration because only machine-readable formats can be linked automatically. For human-readable formats, users will need to open each file and read it all to find that piece of information they are searching for. In other words, without machine-readable formats, automatically linking all the repositories into one central place is not possible.

2.3 Usability issue: cluttered user interface

To improve discoverability for the general public, the interface of geo-portals should be simple. However, geo-portals provide too many map layers, leaving users unhappy. Taking the geo-portal, in the below picture, from Washington State Department of Transportation, as an example, this kind of default layout is designed in web 1.0 stage, where geo-portal is trying to imitate desktop GIS software, ArcGIS, which worked fine only if users have worked with GIS interfaces on a daily basis. For beginners, multiple map layers make it difficult to elicit a user's intent and overestimate their interests in interaction with the map.
In addition, Geo-portals need to be more minimalistic because users are not interested in interaction with the map. Denver's geo-portal had over 7000 visits in a month, but only 1% of keyword searches that led users to the geo-portal included the words “map,” “location,” or “find” (Timoney, 2012). According to Dominikus Baur's article, "The death of interactive infographics," only 10-15% of users click on buttons of maps (Baur, 2017). Those two examples in the above show that most users do not want to interact with the map. They just want a single piece of information as they want a needle in the haystack. Most geo-portals are preoccupied with toolbars and layers which only make users want to leave.

Geo-portal was designed with the same principle of desktop GIS software such as ArcGIS, which is designed for data management rather than data discovery. The workflow of data managing or analysis on ArcGIS is a flow diagram. For example, if the goal is to calculate the number of properties within a flood area, two map layers are needed. One map layer shows the flood area as a
polygon and one map layer shows the properties as points. The ArcGIS interface for this task is in the below picture:

![Image of ArcGIS interface](image1.jpg)

**Figure 2.3 The interface of ArcGIS.**

In here, users click the “Spatial Join” tool from the toolbox drop-down menu (in the picture on the right), then a white dialog will pop out (in the below picture). After filling out the dialog, those two layers will be joined based on spatial location. Then the result can be saved in a new place.

![Image of toolbox drop-down menu](image2.jpg)

**Figure 2.4 The toolbox drop-down menu in ArcGIS.**
As can be seen, the workflow of ArcGIS is data flow through tool after tool as water through pipes, then the result is collected at the bottom. For GIS professionals, the more layers they have, the more likely they can reveal a multivariate relationship. On the other hand, public users do not understand this workflow or the logic behind those tools. Their goal is only to retrieve a piece of information that they are interested in, not decoding the buttons on the interfaces. Therefore, they need a much simpler, clear and meaningful interface for data discovery.
3. Current solution and research question

3.1 Lacking data discovery across political boundaries

Currently, instead of integrating data into one place, cities and counties are keeping on building their own portals. For example, Culver City and West Hollywood are next to each other with similar populations. They both hire Socrata (a startup company for government data) to build their portals. Culver City pays $28,000 for 15 datasets, and West Hollywood pays $15,983 for 50 datasets. According to Culver City's Chief Financial Officer, Jeff Muir, as cited by Mendel (2015), their portal does not garner much traffic (Mendel, 2015). Clearly, this is not a cost-efficient method to publish data for the city budget, and not an effective method for data discovery across boundaries.

In addition, cities are still focused on data analysis for researchers instead of data discovery for beginners. For example, Boston developed a new portal in 2016 called Analyze Boston. It targeted researchers at large institution, such as the Boston University and the MIT. According to Boston Chief Data Officer, Andrew Therriault, by using OpenGov (an open source platform) instead of a private company, their portal will improve researchers’ access to datasets better because they can keep adding new features that will yield more options for data
analysis over time (Nyczepir, 2017). Boston is the first city keeping researchers as target users and wants user feedback to improve themselves. However, what they are doing is not what most public users need. Developers from Boston stated that they want the flexibility to use data however they want outside of the portal (Nyczepir, 2017). Therefore, to truly make data more engaging to the public, geo-portals should first focus on data discovery rather than data analysis.

3.2 Lack of APIs integration focus on discovery

Rather than data discovery for citizens, most studies of geo-portal API integrations focus on data analysis such as developing a specification to encapsulate data and functions as Web service. For example, a geo-portal for urban research integrates services such as thematic mapping, non-spatial or time-series attributes (Zhu et al., 2014). This kind of API integrations may facilitate a small group of expert, but not enough for data discovery for public users.

Current geo-portal API integrations only focus on few specific topics instead of general topics. For example, GovWild.com integrated few other geo-portals from the US and the EU. It only focused on politicians, companies and government spending (Shvaiko et al., 2012). This kind of API integration limited the potential value of geo-portal as a central place for the public. Since a central place can be imagined as a shopping center which prevents the hard work of
going through each shop. Therefore, geo-portal should include as many topics as possible instead of just a few topics.

3.3 Unfriendly UI design

One major issue of geo-portals is squeezing a variety of information into a single interface. Think of search engine such as Google, when most users expect a simple keyword search box, geo-portals have a list of features that distract users from carrying out their primary intention. Making content competing with each other for attention only increase the cognitive barriers. Taking ScieneBase.com (in the below picture) as an example, irrelevant or rarely needed information is displayed on the landing page. This information was cut into several dialogs with the same background color, shape, and size. Which indicates that every feature on this landing page is equally important. But in reality, only the search box and the “advanced search” option are the priorities. However, they are very small and hard to see compared to other elements.

Since making the important components visible means making the less important components invisible (Norman, 2002). A geo-portal’s interface should be less visually busy. To improve the user experience for public and help them to concentrate on performing, the search box should stand out more.
To remedy the user experience issues mentioned above, the research question for this thesis is how to offer an overview of available geo-data across political boundaries with one click of a button for the absolute beginners. The basic principles for the design are: 1) does not require any training 2) find data fast 3) gives users links for more detailed information and 4) requires minimal coding.

4. Methodology

4.1 Define users’ goal

The first step for customizing geo-portal for beginners is to create a user persona. User persona is one of the most effective methods that is promoted by many user experience experts. The idea is following a logical order that ends
with refined characteristics of users to guide decisions about features, interaction and visual design (Goodwin, 2018). According to Alan Copper, the creator of user persona method, the goal of users has flavors: end goal (what user get out of a product) and experience goal (how user want to feel when they use it) (Cooper et al., 2007).

In general, there are four types of users: 1) everyday people who need geospatial information such as parking locations, land use plans, population estimates; 2) commercial ventures who use data to boost sales; 3) researchers who use data to tell stories; 4) developers who build Apps.

Since the goal of open geo-portal data is to make information available to anyone, this thesis focuses on everyday people first. Historically, geo-portals focus more on technical users, in particular, GIS professionals as the early adopter. For beginners, it is hard for them to have an overview of available geospatial data. For example, to find a piece of geospatial data, users need to find the geo-portal which contains that data first. Most beginners do not know that they need to find the geo-portal first. They do not know the name or the address of that geo-portal. Therefore, they search the keywords on Google instead of geo-portals. Google will give them a large number of results, more than they could read. They do not understand the terminology or the logic in Geospatial data discovery. They can not identify top scientific papers, authors and agencies just by looking at the results list. They do not know where the data repositories are and where to start. Users need a place to start, where they can easily
determine the search areas and instantly get a list of result from multiple repositories, where they can also click the hyperlinks of the result to open a new page with more detailed information.

In summary, Users’ end goal is to see an overview of available data repositories, to know which data repositories are the most relevant to their topic, to open those data repositories for more detail information, to feel positive emotion in the whole process.

4.2 Integrating available APIs

API integration is the top priority and main barrier to building a central place. Since most geo-portals are not focused on building efficient APIs, linking them together is difficult. Most geo-portals’ APIs only support flat data (plain text format) that includes a table with one record per line and the columns are separated by a comma. This type of API is good enough for building a specific feature, but not good enough for powering a search platform. For example, the API endpoint can tell us the current temperature of an area, but can not give a hyperlink to datasets which record all the temperature through the years. For many developers, showing the current temperature is enough for their application, but it is not enough for finding an overview of available data.

On the other hand, USGS has developed an innovative API-driven application, ScienceBase.com, which is a great example of API-first mentality. Searching for items is as simple as sending a URL with correct parameters. For example, a URL
that starts with:

"http://www.sciencebase.gov/catalog/items?s=Search&q=water" can define the data format and spatial extent for the return results. The returned result is in JSON format which can be displayed as a hyperlink for users to click. The JSON data from ScienceBase.com was parsed with Ajax “json.parse” function. The code is in below picture.

```javascript
var xhr = new XMLHttpRequest();
xhr.onreadystatechange = function() {
    if (this.readyState == 4 & this.status == 200) {
        var i;
        var obj = JSON.parse(this.responseText);
        for (var i = 0; i < obj.items.length; i++) {
        }
    }
}
```

*Figure 4.1 Parsing JSON data from ScienceBase.com with Ajax.*

For APIs that only support flat data, its URL was linked with the central place for user to click to open the webpage of the search result on another window. Since those APIs do not support search function, we think this method can at least save users’ time by not typing a keyword again and again every time they open a local portal.

### 4.3 Intuitive interface design

The presentation of the results, the display of metadata, the order of datasets, the use of font size was all taken into consideration to make the interface intuitive. The most intuitive interfaces are almost invisible to users because they
are very predictable. They work as users expect them to work. They avoid any unnecessary elements to reduce confusion (Rosenfeld, 2007). Therefore, geo-portal should help users to be able to foretell the workflow. It should act like the rest of the major search engines, such as Google, to make users feel more comfortable.

4.3.1 The minimal information required to conduct a search

The idea of minimum viable product from Lean Startup movement is used here. Which is instead of spending years and years building the perfect product, build it fast that has just enough core feature to be useful and polish later with real user feedback (Ries, 2011). The idea is adjusting strategy based on user behavioral instead of whiteboard theoretical features.

According to user behavioral studies, the process to a clean and simple interface start with a simple set of keyword. 65% of all geo-portal users retrieve a single keyword and leave, 12% of users search for more than three features during a visit (Timoney, 2013). Since the vast majorities do not want to spend time to interact, user inputs should be as few as possible. The most important information that a geo-portal need from users are: the keyword and the geographic area.
4.3.2 Search geospatial data across political boundaries fast

There are three general ways to define a geographic area: 1) by specifying place name; 2) by drawing bounding box (spatial extent); 3) by entering coordinates (Resch, 2013). The first method is time-consuming. When users need to search multiple cities next to each other, they need to type many places names. The second method, drawing a bounding box, is clearly faster. Users only need to click a button and drag a polygon on the map. However, the area defined by the bounding box is not as accurate as the area defined by coordinates or place names. But considering manually entering coordinates require a lot of typing, drawing a bounding box is a better way to define a geographic area for geo-portal.

In addition, another interaction that facilitates drawing box, zooming in and out, was also taken into consideration. In research, users enjoy double-clicking on map more than selecting discrete zoom levels (Haklay, 2003). Users should be less occupied with interaction itself and more occupied with the actual task and workflow. Any tool that is not related to zooming or drawing the bounding box was avoided.

4.3.3 Make Intuitive UI that does not require “Help”

Intuitive interface emphasizes the low mental effort required while interacting with a product (Loeffler, 2013). In other words, geo-portal's
interfaces should work as how users expect them to work. The first step is setting up the layout of the landing page. What size and where should each element be? The spatial relationships between each element were structured based on their importance. The most important elements are the keywords search box and the map. Therefore, they should occupy most parts of the page. A logo that is used as a headline is also necessary for grabbing attention and showing the purpose of this page.

According to study, human eye path reads in a “Z-pattern” which starts from the upper left corner and move onto the right (in the below picture). Then move on to the lower left corner and ends on the lower right corner (Stelzer, 2018).

![Figure 4.2 The Z-pattern of human eye path.](image)

Considering that the logo is the first thing users should see, it is located on the left upper corner to be in harmony with natural reading gravity. All elements were arranged in a sequence that follows the order from 1 to 8 in the below picture.
Moving on to the lower left corner, there is a list of data repositories. Since data repositories are not the first thing users need to see, it should be hidden to reduce visual busyness. With a click of a small button, users can open the list up or hide it again. This will place users' attention only on the most important elements and leave white space for users to stay relaxed. Below the list of data repositories, the last element in the landing page is a group of social media sharing buttons, which will also be hidden behind a small button to aid a sense of order.

The logo and the map are the first areas users will focus on when they land on the page. Since they are the most colorful area, its size, resolution, information density, interaction possibilities should all follow user's need. Considering that most users are familiar with Google Map which always occupy
the whole webpage, the map of geo-portal accordingly will cover 100% of the page vertically from top to bottom. In addition, to prevent users seeing the landing page as a navigation website which takes them from A to B, the map covers only 70% of the page horizontally to leave enough space for other elements.

How many colors should be used? Considering that the content in the map is complex, the visual busyness needs to be toned down by using only two main colors. The main colors should represent the real world content to reduce cognitive demand and help users focus more on the actual task. Considering that the main surface of the earth is covered by water and plants, and water is typically understood to be blue, while plants are typically understood to be green, the two main colors should be blue and green.

Also, blue and green give a first impression of trustworthiness and nature. In North-American culture, blue gives a feeling of trust, and green means freshness and nature which also associate with positive emotion such as peaceful and relaxing (Niggulis, 2017). To keep visual consistency, the logo (in the below picture) is also in blue and green.

Figure 4.4 The logo in blue and green.
Strategically, the entire portal interface (in the below picture) has only one typeface for clarity. Different font sizes were used to help increase scalability, legibility, and readability. Carefully, "Arial, sans-serif," the same typeface that Google is using for their interface, is used here to create familiarity.

Figure 4.5 User interface of landing page.

Figure 4.6 User interface of landing page with hidden elements.
To make sure that the system communicates what’s happening, the system will pop-up an error message (in the below picture) if users forgot to define the bounding box. This use of message can reduce frustration for users.

![Image](image_url)

*Figure 4.7 The alert message when users forgot to define the bounding box.*

Thinking about the defaults, auto-complete entry box (in the below picture) and a placeholder was provided to reduce the burden on users. This is very important because auto-complete drives clean user queries. As the first generation assistant, it also saves users’ time by not typing the entire phrase.
According to a study, the objects that GIS professionals are searching for on geo-portals are: 52.4% for thematic geo-data (land parcels, land use, property owners, zoning plans), 23.8% for locations or addresses, 9.5% for real properties, 1.9% for geo-data services, and 12.4% for other objects (Points of Interest (POIs), persons, buildings, metadata, city maps) (Resch, 2013). As a result, 300 of the most popular objects of geo-portal was collected to create the auto-complete search box in the above picture.

Apart from search capabilities, the presentation of the search results is also a central part of the design. Most users would like the search results to be structured (Resch, 2013). In terms of the visual presentation, displaying results on multiple pages makes users frustrated. Therefore, a list of the results is on one page, in a structured list. Clickable blue text hyperlink is used to direct users.
Studies show users engage more with a blue text hyperlink. The user experience manager of IBM, Paul Ray, as cited by Fried (2009), said they test a number of colors and the specific blue (#0044CC) make users click more advertisements, which give them additional 80 million in annual revenues. Further, due to the Google influence, users are familiar with a clickable hyperlink. With this insight, to drive more engagement, the result is listed as hyperlinks (in the below picture).

![Image](image.png)

**Figure 4.9** The presentation of the search results.

## 5. Results

### 5.1 Easy, fast, intuitive, informative access

The geo-portal proposed in this thesis is easy enough to operate without any pre-knowledge. For example, if users want to find water temperature data in the Great Lakes area, they will first land on the page in the below picture.
Once their mouse hovers over “the keyword input box” element, they will see shadows under that element. When they click that element, the placeholder, “Type a keyword...” will disappear. Then they will need to type one character, “w”, on their keyboard. Now, the auto-complete drop-down menu (in the below picture) will be able to be seen. When they select “water temperature,” the box will be filled automatically. By clicking the little black box on the upper left corner of the map, they can draw a pink bounding box to define the Great Lakes area by dragging on the map.
After users click the search button, the map will disappear, and a list of result will be able to seen in the same place (in the below picture). Users will open a new window every time they click a hyperlink in the list of result. In addition, if users hover over a hyperlink, a pop-up text box will appear to provide the summary on the discovered data. By clicking the “search again” button, they will refresh the whole page.
After the search, if users want to see the data repositories or share this portal on social media, they can click the small buttons on the lower left corner of the page to see those two hidden elements (in the below picture). Clicking those buttons again, those elements will be hidden again.
Figure 5.5 Hidden data repositories and sharing buttons in the lower left corner.

In summary, the map is a page canvas rather than a tiny sidebar window. Very few visible distractions that narrow workflows down without cluttered information and list of toolbars. A message is hidden in a popup and the map is easy to read from cartography point of view. Users can see a list of results in less than one minute is quite fast.

As City of Denver’s geo-portal usage statistics showed, the primary use of geo-portals is quick information retrieval, not complex interaction with the interfaces. Insights found through this research break a decade of misguided professionals’ assumption about the public, their behaviors, and the geo-portals they need. Pushing prolonged loading of little interest to the vast majority of users is not useful, which puts their priorities backward.

On the contrary, the proposed geo-portal is intuitive for beginners because every piece of information is arranged to serve their specific purpose instead of
incorporate a number of different features mashed together to make the whole. For example, the maximal zoom level of the map is set to 9 to prevent users from looking into the details of the map to get distracted (in the below picture).

![Map on maximal zoom level](image)

*Figure 5.6 the map on the maximal zoom level.*

In addition, one of the earliest, most immediate takeaways from user research was that most open portal users are young, ages from 25 to 45. In addition to simplification, the visual language has to make geo-portal feel modern. In order to be modern, the geo-portal is responsive and optimized for mobile. Flat design idea is used from minimalism to design button icons (in the below picture) representing their true functionalities.
In conclusion, the interface minimizes the knowledge gap between what users know before coming to and what they have known in order to use it.

5.2 Low-cost open source application

This geo-portal is lightweight and JavaScript-based using shareable URLs and open-source libraries such as Bootstrap, Leaflet, jQuery and Open Street Map. The base map was built with Open Street Map and Leaflet library for drawing bounding box. jQuery and Bootstrap were used to make it responsive to screen size. Adobe Illustrator was used to draw wireframe and icon.

Comparing to most cities’ geo-portals, the budget of this portal is very low. All thanks to ScienceBase.com and Data.gov for opening their data through well-designed API. According to the “Open Data in California,” building cities’ geo-portal is not cheap: Los Angeles city Pays $25,000 annually to OPENGOV; L.A County pays $892,824 in three years to Socrata to host their datasets. The annual salary for open portal labors cost the state of New York and California 4 to 5 million dollars (Boyle, 2015). Comparing with above portals, the technological achievements of our project is highly cost-effective.
6. Discussion

6.1. The uniqueness of thesis

Most user experience research for geo-portal focus on experts who have experienced GIS software, which makes them miss half of the audiences. The uniqueness of this thesis is putting the beginners first. Identifying beginners’ unique goal then customizing the content structure is very important because every design decision starts with the question, “who is your audience?”. While geo-portals have served the scientific community well for many years, it still suffers from user experience limitations.

Beginners are very valuable for GIS community. When data is shared and uploaded by more beginners, they will help the community to visualize and improve data. For example, since 2015, the City of Cambridge has opened its portal to citizens, and data collaborations have increased in large number (City of Cambridge, 2016). Their “Open Data Report in 2016” shows public prefer tool such as “Crowd-sourced Problem Inventory.” (City of Cambridge, 2016) As can be seen, public are interested in further exploration and research and geo-portals could support them.
6.2 Thesis evaluation

This thesis provides a new method of getting data from multiple repositories into a central place which dramatically improve the discoverability. This thesis answered the question of how to offer an overview of available geo-data across political boundaries with one click of a button for the absolute beginners.

However, there are several factors to improve: First, only 2 data repositories were linked with the central place so far. Most geo-portals do not have well-structured APIs that can be linked. Governments have limited budget for geo-portals while API development is not their priority for now. As a result, for the 26 data repositories in the front page, users need to open them in a new window and push the search buttons on those pages manually. On the bright side, users do not need to type the keywords repeatedly every time they open a page. When they opened those repositories from the central place, the keywords were already filled out automatically.

Secondly, the items in the list of result cannot be sorted by name, popularity or relevance. This means users will need to find the data out by reading the whole list. Clearly, this is not an ideal situation. In the future, users should have more options to structure the list of results.

Thirdly, although the validity of the new platform is apparent through the above analysis, it would be desirable to conduct a user test to gain more insights into the validity and potential improvements.
7. Future research

In addition to the potential future researches as outlined in the discussion section above, a few further suggestions are provided below.

7.1 Multiple formats, open standards.

Formats and open standards of metadata are essential to geo-portals. ‘Formats’ were designed to fit the unique goals for different groups of users. This means that more formats geo-portals support, the more users it will have. In general, non-GIS users including journalists, designers, engineers and many other professionals, have been overlooked by geo-portals which focus more on formats for GIS software. To make sure that users can upload data into their software of choice, geo-portal should provide data format conversion on the fly.

Here are some most popular formats: 1) GeoJSON, is versatile, machine-readable, viewable in a web browser; 2) XML, is machine-readable and it offers a lot of power for tabular data; 3) Esri Shapefile, is the most popular spatial data format in the GIS community; 4) KML, is instantly viewable in a web browser and it is the format for Google Maps and Google Earth; 5) PDF, looks clean and is easily shareable; 6) CSV, as a tabular format, is the easiest to scan by humans.
7.2 Put spatial first

Since “80% of data is geospatial,” and many of the most valuable datasets from government agencies are geospatial (Kuper, 2000), for Geospatial data to have bigger influence, publisher need to treat geospatial as the most important attributes. For now, geospatial data is served as secondary attribute. Most cities’ data portals do not tag data with coordinates, not to mention residential address.

Geo-portals have not utilized the powerful search engine optimization (SEO). Typing in any residential address into Google and users will sort through tons of real estate advertisements before they can see any links to geospatial data repositories. If it is not in the top 5 pages of Google, it will most likely not be seen.

With the REST-API, data repositories can be shared as a unique URL, which is suitable for prioritizing SEO. Since Google may be a beginner’s first stop when looking for information, geo-portals should provide a better user experience for the Google generation.

7.3 Recommender system

Since the aim to built a one-stop-shop for geospatial data shoppers, the ability to sort, filter and auto suggest items is crucial. This enables the possibility to present user with additional data that is related to their research.

Amazon.com is a great example of a recommender system. They combined the collaborative filtering and the content-based filtering approach (Schneider, 2005).
Collaborative filtering analyzes users’ previous interaction, which requires huge number of users and predicts their preference by comparing with other users. Content-based filtering focus on user’s profiles, which is based on their own previous action (Vockner, 2014).

To build recommender system in the future to improve geographic resources discovery and user experience, tracking the user’s interactions is a must. In geo-portal, users’ location, search area, keyword and clicks should be recorded. Those actions themselves can be divided into two categories: “view” and “buy.” View actions is entering a keyword. A click on the details could be thought of a buy action. For now, only geo-portal API implemented on ArcGIS online cloud-based platform has item ratings capability.
Reference


2018, from https://conversionxl.com/blog/which-color-converts-the-best/


