

Data Visualization of Geospatial Data for Future Business Investments

by

Naga Mahathi Meruga

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Arts

in

Human – Computer Interaction

Carleton University
Ottawa, Ontario

© 2019, Naga Mahathi Meruga

Abstract

Economic developers make elective decisions and solve problems in complex or ambiguous situations by gathering, diagnosing, and judiciously analyzing the information about the situation and environment in order to identify and evaluate options and select best course of action (Leigh, 2017). With increasing globalization and corresponding shifts in market and production infrastructures, the demand for a web-based tool to efficiently address these types of problem solving is increasing. The research undertaken for this study leverages recent developments in Asset Information Modelling (AIM) — considered here as a framework for managing inter-related geometric, graphic and text data for asset management and customer engagement. Our study considers the qualities of user experience — including statistically significant changes in the workload and accuracy of decision making — by comparing text search and data visualizations.

Keywords: data visualization, geospatial data, decision making, cognitive load, decision support systems, NASA TLX test,

Focus: Promoting informed decision-making using geospatial data through data visualizations and minimizing cognitive load.

Acknowledgement

It is a great pleasure to acknowledge my deepest thanks and gratitude to my thesis advisor Dr. Stephen Fai, School of Architecture / Engineering & Design at Carleton University, for allowing me to work at Carleton Immersive Media Studio (CIMS) and develop my thesis topic under his kind supervision. I would also like to express my sincere thanks to CIMS and the entire team of Ontario East Economic Development for providing me with all the necessary facilities.

I am very much thankful to Dr. Robert Biddle, School of Computer Science at Carleton University for his unending support, constructive comments, and suggestions to make this study possible.

Finally, this accomplishment would not have been possible without the support of my parents and little sister. Thank you for helping me survive all the stress and not letting me give up in the toughest times. To all my friends scattered around the world and

roommates, thank you for being a family away from family, helping, listening and being there whenever I needed.

Table of Contents

Abstract.....	ii
Acknowledgements	iii
Table of Contents	v
List of Tables	viii
List of Illustrations.....	x
List of Appendices.....	xii
1 Chapter: Introduction	1
1.1 Background.....	1
1.2 Asset Information Model.....	2
2 Chapter: Literature Review.....	6
2.1 Urban Planning.....	6
2.2 Evaluation Methodology for Visualizations.....	9
2.3 Cognitive Load Theory.....	15
2.3.1 Sternberg Memory Test.....	16
2.3.2 Tapping Test.....	17
2.4 Commercially Existing Data Viz and Analytical Tools.....	18
2.4.1 Google Analytics.....	18
2.4.2 Tableau Software.....	19
3 Chapter: Design	22
3.1 Defining the user and the problem	22
3.2 Defining the data to be represented	28
3.3 Defining the dimensions required to represent the data	30

3.4	Defining the structure of the data	31
3.5	Defining the interaction required from visualizations	34
3.6	User interface design of AIM	36
3.7	Technology used for building UI.....	40
4	Chapter: Methodology.....	42
4.1	Participants	43
4.2	Procedure.....	44
4.3	Material	45
4.4	Data Analysis.....	46
4.4.1	Pre-test questionnaire (Qualitative data)	46
4.4.2	Post-test questionnaire (Quantitative data)	46
5	Chapter: Results.....	47
5.1	Pre - Test Questionnaire	47
5.2	Task 1 - Text Search.....	55
5.3	Task 2 - Data Visualization	56
5.4	Task Load Index of Task 1 & Task 2	59
5.4.1	Mental demand.....	59
5.4.2	Physical demand.....	60
5.4.3	Temporal demand.....	61
5.4.4	Performance... ..	62
5.4.5	Effort.....	63
5.4.6	Frustration.....	64
5.4.7	Time.....	65
6	Conclusion	67
6.1	Limitations.....	68

6.2	Discussions	70
Appendices.....		72
Ethics Clearance		72
Recruitmnt Poster.....		74
Consent Form.....		75
Task Scenario		77
Pre - Test Questionnaire		78
Post - Test Questionnaire.....		84
NASA-TLX Mental Workload Rating Scale.....		88
Bibliography or References.....		90

List of Tables

Table 1	Participants age in Chapter 4	44
Table 2	Task 1 post-test questionnaire descriptive statistics in Chapter 5	55
Table 3	Task 1 NASA TLX descriptive statistics in Chapter 5	56
Table 4	Task 2 post-test questionnaire descriptive statistics in Chapter 5	56
Table 5	Task 2 NASA TLX descriptive statistics in Chapter 5	58
Table 6	Mental Demand Paired Samples Statistics in Chapter 5	59
Table 7	Mental Demand Paired Samples Correlations in Chapter 5	59
Table 8	Mental Demand Paired Samples Test in Chapter 5	59
Table 9	Physical Demand Paired Samples Statistics in Chapter 5	60
Table 10	Physical Demand Paired Samples Correlations in Chapter 5	60
Table 11	Physical Demand Paired Samples Test in Chapter 5	60
Table 12	Temporal Demand Paired Samples Statistics in Chapter 5	61
Table 13	Temporal Demand Paired Samples Correlations in Chapter 5	61
Table 14	Temporal Demand Paired Samples Test in Chapter 5	61
Table 15	Performance Paired Samples Statistics in Chapter 5	62
Table 16	Performance Paired Samples Correlations in Chapter 5	62
Table 17	Performance Paired Samples Test in Chapter 5	62
Table 18	Effort Paired Samples Statistics in Chapter 5	63
Table 19	Effort Paired Samples Correlations in Chapter 5	63
Table 20	Effort Paired Samples Test in Chapter 5	63
Table 21	Frustration Paired Samples Statistics in Chapter 5	64
Table 22	Frustration Paired Samples Correlations in Chapter 5	64

Table 23 Frustration Paired Samples Test in Chapter 5	64
Table 24 Time Paired Samples Statistics in Chapter 5	65
Table 25 Time Paired Samples Correlations in Chapter 5.....	65
Table 26 Time Paired Samples Test in Chapter 5.....	65

List of Illustrations

Illustration 1	Search for businesses (Text Search) in Chapter 1.....	5
Illustration 2	Data Visualization in Chapter 1.....	5
Illustration 3	Taxonomic group for heuristics selection in Chapter 2.....	11
Illustration 4	Overview of Google Analytics dashboard in Chapter 2.....	19
Illustration 5	Visualization created using tableau software in Chapter 2	21
Illustration 6	Persona of an Investor in Chapter 3.....	27
Illustration 7	Persona of an Economic Development Officer in Chapter 3.....	27
Illustration 8	Summary of counts of subsectors, industry groups, industries, and Canadian industries for each of the NAICS sectors in Chapter 3.....	29
Illustration 9	Representation of counties and municipalities of Ontario East with their geographical boundaries in Chapter 3.....	29
Illustration 10	North American Industry Classification System’s structure for organization of data in Chapter 3.....	31
Illustration 11	Code classification for sectors, sub-sectors, industry groups, industries and Canadian industries in Chapter 3.....	33
Illustration 12	Infographics and visualizations exist on a continuum in Chapter 3.....	34
Illustration 13	Pie-charts plotting Industry group level data versus Industry level data in Chapter 3.....	36
Illustration 14	Landing page Chapter 3	37
Illustration 15	Infrastructure map style Chapter 3	37
Illustration 16	Agriculture map style Chapter 3	38
Illustration 17	Real estate map style Chapter 3	38

Illustration 18	Side menu Chapter 3	39
Illustration 19	Location dropdown in text search feature Chapter 3	39
Illustration 20	Industry dropdown in data visualization feature Chapter 3	40
Illustration 21	Simple bar of percentage by previous investment experience in Chapter 5.....	48
Illustration 22	Simple bar of percentage by investment familiarity in Chapter 5.....	48
Illustration 23	Simple bar of participants interest in properties and business types by percentage in Chapter 5	49
Illustration 24	Simple bar of participants investment goals by percentage in Chapter 5.....	49
Illustration 25	Simple scale of Participants agreement on investing in an unfamiliar location in Chapter 5	51
Illustration 26	Simple scale of Participants agreement on considering property's location/neighborhood and accessibility in Chapter 5	52

List of Appendices

Appendix A Ethics Clearance.....	72
Appendix B Recruitment Poster.....	74
B.1 Consent Form.....	75
B.2 Task Scenario.....	77
B.3 Pre-test questionnaire.....	78
B.4 Post-test questionnaire.....	84
B.5 NASA-TLX Mental Workload Rating Scale.....	88

Chapter 1: Introduction

In this paper, we present a comparison between full-text search and data visualization of geospatial data related to the agri-food industry in Eastern Ontario. Adopting a relatively new paradigm developed for the Architecture, Engineering, Construction, and Operations industry (AECO) — Asset Information Modelling (AIM) — we evaluate the most effective interface for remote access to large and diverse datasets utilized for the evaluations and promotion of economic development opportunities. Our AIM provides users with a data visualization feature using a multilevel search. We discuss how users processed visual information, communicated findings in constructive ways, derived connections and trends between search results from the data with the help of visualizations.

1.1 Background

Cities, towns, countries, and all local entities in a global economy have the challenge and opportunity of crafting their own economic destinies (Leigh, 2017). The Ontario East Economic Development Commission (OEED) promotes Eastern Ontario as a center for investment and business growth ("Choosing Ontario East Is Easy..."). Ontario East has an abundance of high-quality investment ready infrastructure and lands with close proximity to major markets in both Canada and Northern America. Expressways, border crossing, rail, airports, deep water ports, etc., can offer growth to existing and future economy. In order to foster this development, economic development officers (EDO) of OEED are making great efforts to connect and market to entrepreneurs, businesses and investors across the globe. During this process an EDO is constantly collecting and analyzing

geographic, demographic and infrastructural information from different sources. These data sets are large and are not often adequately structured for the purpose. Therefore, there is a need for refined data management and query. Determining what the data means from different sources is not an easy task and requires applying a new approach to old problems.

Carleton Immersive Media Studio (CIMS) in collaboration with Ontario East is working to get the most value from the assets by building an Asset Information Model (AIM). Bringing all the data together in a way that allows users to make informed decisions is the goal. The model displays 3D representations of assets and their associated asset information in a map-based format and data visualizations. It is a virtual model that is easily accessible by people in different locations and via different interfaces. AIM is designed by identifying, assessing and comparing the data at EDO disposal into a single information resource allowing staff, investors and businesses to collaborate and make more informed decisions.

1.2 Asset Information Model

The Asset Information Model or AIM is a term used to describe the collated set of information gathered from all sources that supports the ongoing management of assets (McPartland, 2017). AIM is a model that consists of all the resources, data and information that is required to assist asset management and asset operation. As specified in PAS 1192-2:2013 by British Standards Institution, Building Information Modelling is “the process of designing, constructing or operating a building or infrastructure asset

using electronic object-oriented information.” (Building Information Modelling Scottish Futures Trust, n.d).

In the AIM developed for OEED, the model allows users to explore building interiors, exteriors and surroundings in a 3Dimensional environment. The 3D models of buildings are modelled using Building Information Modelling (BIM) – developed from laser scanning – to record building geometry and to facilitate geo-referencing to a universal co-ordinate system. The 3D models are digital representations of physical and functional characteristics of the Ontario East region.

Collaborative use of BIM, geospatial data, and farm business registration data from the Ontario Ministry of Agriculture, Food and Rural Affairs and North American Industry Classification System (NAICS) through the AIM brings forth a new kind of application that can be used by EDO for urban planning, municipal management, and asset management. Business investors can use this tool for market research, competitive research, business planning and so on. Users can make useful suggestions or decisions for existing and future business investments. The AIM can be used on a wide range of devices such as personal computers, laptops, and tablets when connected to the WWW. The OEED AIM also uses open source geospatial data for temperature, precipitation, wind speed, soil, exclusive – with datasets registered by latitude and longitude. It holds both geographical and graphical information along with metadata specific to properties and portfolios. These can be updated as required by individuals with administrator credentials. This tool empowers economic developers and investors to plan, execute and

monitor activities required to attract investors around the globe, investigate potential outcomes of economic development and manage moves in a region that stretches from Port Hope to the Québec border – an area of 46,200 square kilometers.

One of the unique features of this AIM is that the model is not just overlaid with topography, but it can also be customized to include different types of information such as infrastructure, agriculture, soil conditions, landcovers, real-estate pertaining to land parcels and so on allowing for simplification of the data. Our research on the interface is concerned primarily with two main features: a) search for businesses and b) data Visualizations. The “search for businesses” feature allows users to filter location and industry type as needed and draw spatial relations between results when they appear on maps. Users can further pan and zoom into specific areas of the map or a building to view a 3D model of an asset generated using BIM. Users can delve further into details like underground water and gas pipe lines, nearest highways, accessibility – allowing extensive parameters for “virtually” visiting an asset. The “data visualization” feature on the other hand, facilitates market and competitive research by providing an overview of data through visualizations. These visualizations can be generated at different levels and

can be customized to draw comparisons between industries, draw conclusions on supply and demand and so on.

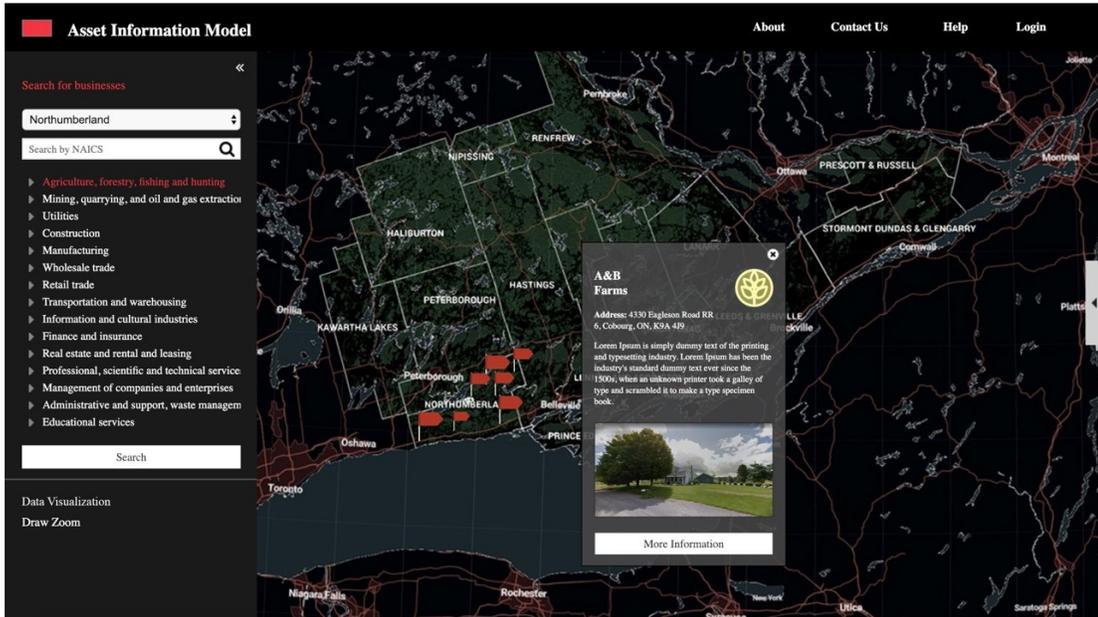


Illustration 1: Search for businesses (Text Search)

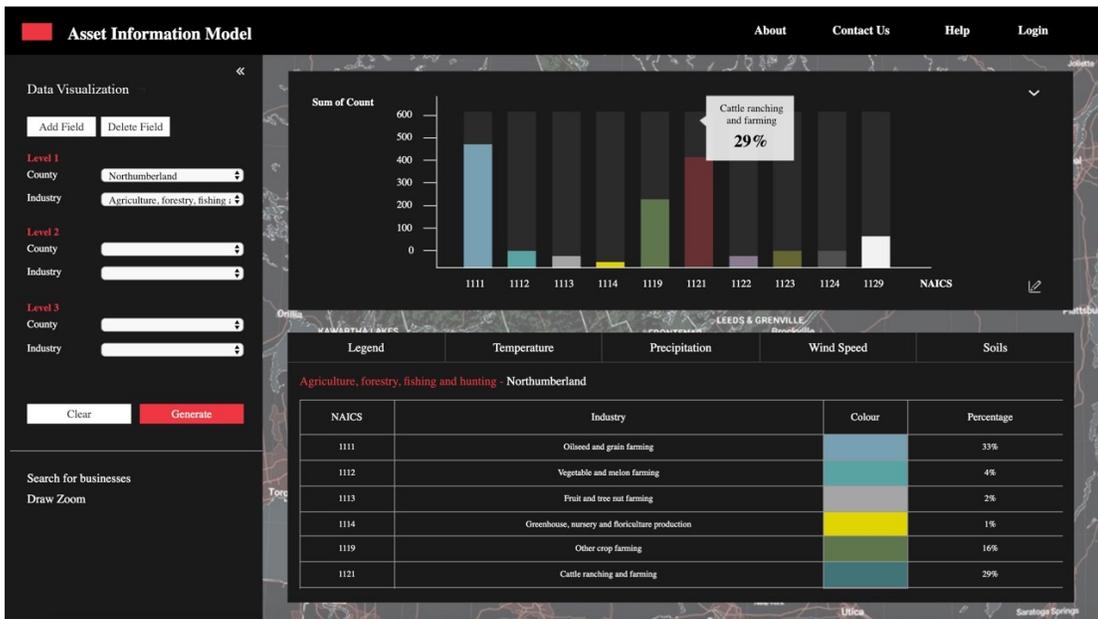


Illustration 2: Data Visualization

Chapter 2: Literature Review

This literature review intends to give a synopsis appropriate for this study's investigation question. Pertinent research papers were sought utilizing the keywords such as geospatial data, informed decision making for returns on investment, data visualization, and cognitive load theory. More than 40 relevant papers were reviewed, and 21 papers were selected for further discussion. Urban planning, usability evaluation methods, cognitive load theory, commercially existing data visualization and analytical tools, and design, are discussed below.

2.1 Urban Planning

Implications of identifying priority areas by economic developers to redesign agriculture and real-estate sectors are huge. Information on industry and business variants, supply chain processes, parcels of lands and real-estate are all open data available through NAICS. NAICS is a comprehensive system encompassing all economic activities (Statistics Canada, 2012), such data holds tremendous potential and opportunities for urban planning and in attracting investors. OIED aims to use both fundamental and strategic layers of this data to provide valuable assistance for decision making tasks. But as fascinating as it sounds the complexity in the data can misrepresent the enterprise activities. In a business environment, where quick and reliable access to knowledge is a crucial factor of success, efficient processing of data and information resulting in acquiring new knowledge concerning enterprise becomes essential (Dudycz, 2011). Economic developers aim for urban development and to improve market productivity. They act as middle men between businesses/properties and investors. Developing

regional scenarios enables planners to engage land managers in discussions about the future, especially in contexts that are complex, uncertain and difficult to control (Carter et al., 2018). Many decisions are influenced by location (Sedlák, Komárková, Hub, Struška, & Pásler, 2015). Spatial information is very important for all human beings because things happen somewhere and objects are located somewhere (Sedlák et al., 2015). Investment process is exciting and scary at the same time. Return on investments are uncertain, markets are fluctuating. Effective urban planning in a complex and volatile market environment, is a daunting task for both economic developers and investors. And in such a world people often use both emotions and logic to make decisions. But is that enough to plan cities or invest well? If we need to work in such an environment, visualization of the semantic network in topic maps can provide valuable assistance for the economic data analysis and decision making tasks (Dudycz, 2011).

In terms of urban data visualization, it extends its role in communicating and interpreting the economic assets in a relevant context providing valuable feedback. This can prove to have an impact on processing data to a level where policymakers and investors can base their marketing pitch and investment decisions. A well designed/built data visualization tool can intervene with users to detect and eliminate ambiguity in data analysis. Quality of scalable and searchable information can be enhanced with legitimacy and credibility. Information on land usage, weather, climate, air, water, infrastructure assets such as transportation, communication and distribution channels, crops and natural vegetation, etc. can be digitalized and automated. One may argue that Google street maps provide all such details, but users do not always have the developed spatial abilities required to

understand three-dimensional environments (Carvajal, 2005). This kind of environment lacks in summarizing vast data according to user's needs. Data visualizations not just indicate geographic location but can translate information into statistical graphs to mitigate future impacts and concerns. Incorporating spatial analysis have been undertaken to imagine future possibilities not otherwise anticipated, to raise awareness about plausible landscape futures (Carter et al., 2018). Visualization is an essential ingredient for understanding. According to research conducted into learning processes, Treichler concluded that people learn up to 83% of new information through sight (Carvajal, 2005).

The other major aspect of urban planning is the competition- competition between industries manufacturing or businesses selling similar products and services in an intention to achieve profits. The four P's of marketing mix Product, Price, Place and Promotion as introduced by Neil H. Borden (Banting & Ross, 1973) are critical to forecast demand and supply using the existing data to design successful marketing strategies or to make minimum risk investment decisions. Developing a systematic plan can be divided into three main steps:

- a) Research: Identifying gaps, paying attention on existing infrastructure and choosing the right scale for the plan.
- b) Analysis: Conducting a quantitative assessment on how much of each business or industry is required for economic development.
- c) Planning: Transforming a static document of a business plan to a dynamic model by considering things like proximity to other businesses and services,

competition, ordinances, public transport hubs, logistics, commercial centres, water and power supply, access to required services or raw materials, etc.

Portfolio selection problems, in particular those involving multiple evaluation criteria, place a significant cognitive burden on the decision maker. Decision support systems employing visualization methods that allow the decision maker to progress iteratively towards a most preferred solution can alleviate this burden (Kiesling, Gettinger, & Stummer, 2011). Real-time data visualizations can eliminate the need to maintain prior knowledge, which can only be achieved by years of extensive research. They facilitate the exchange of information between visualizations and the decision maker by unfolding layers of information that cannot be viewed otherwise.

2.2 Evaluation methodology for visualizations

While building a tool, it is important to understand how users perceive the role of the tool and how successfully the tool can fill gaps in their lives. Usability evaluation is a means to gauge and gather this information. Though there are many usability evaluation techniques proposed, so far Data Visualizations are different from other Human and Computer Interaction disciplines. The prime focus of users while reviewing Data Visualizations is to look at new information with a purpose to discover connections between already existing knowledge, foreseen or unknown concepts and content. The set of tasks users want to perform might not be known. Evaluating such systems is problematic because controlled studies might not effectively represent analytical tasks (Perer & Shneiderman, 2009). According to Tufte, a good visualization is defined as a well-designed presentation of interesting data, aiming to communicate ideas with clarity,

precision and efficiency. The field of information visualization is influenced by many different research domains including psychology, semiotics, graphic design, and art (Zuk, Schlesier, Neumann, Hancock, & Carpendale, 2006). The objective of Information visualization is to gain knowledge and understanding of data to intensify perception and cognition. A combination of interactive and statistical techniques is necessary for a successful information visualization system (Perer & Shneiderman, 2009). Heuristic evaluation is a well-known instrument of evaluation in the area of Human-Computer Interaction (HCI), but in the Information Visualization area it is not used to the same extent (Barcellos, Viterbo, Bernardini, & Trevisan, 2018). The word “Heuristic evaluation” here is used to refer the process of evaluating user interface design and their usability problems. While usability heuristics, as known from HCI, encompass a wide variety of issues pertaining to visualizations and the interaction with them, we believe that more specific heuristics are of value, in particular since a wide variety of research fields are concerned. These higher-level issues also require a holistic evaluation of entire systems, and so do not lend themselves to a strategy of divide and conquer (Zuk, Schlesier, Neumann, Hancock, & Carpendale, 2006). Considering Economic Developers and business stakeholders as knowledgeable domain specialists, their research is conducted with predefined objectives, their interactions with AIM can be regarded as social actions. Understanding the past, present, and future implications of the actions in the research ties into social action theory by Max Weber. Social action’s interactive statistics empower users to dig deeper, without forcing them to choose an arbitrary cut-off before analysis begins (Perer & Shneiderman, 2009). Therefore, our evaluation

methodology should assess the activities of users with regards to the meanings assigned to them and the relationship these activities have with the other existing activities.

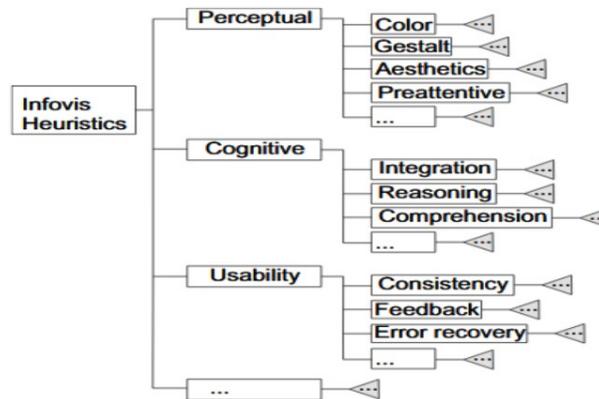


Illustration 3: Taxonomic group for heuristics selection (Barcellos, Viterbo, Bernardini, & Trevisan, 2018).

Illustration 3 Shows different categories of visualization’s heuristics based on perception, cognition and usability. In the current study as I am interested in finding user’s reliability on the tool. I want to investigate how the tool helped in maximizing confidence levels while making decisions, if users were able to comprehend the information that was presented by the prototype and how accurate they were with their reasoning.

Visualizations should not only be well presented but should also help in better interpretation of the data in the shortest possible time. We must identify a specific audience and what they need to know or do and determine the data we’ll use to illustrate our case (Knafllic, 2015). Evaluation methodologies can be broadly classified into inspection, testing and enquiry. Inspection methods focus on a user interface, testing methods focus on task performance, and enquiry methods focus on user data (Sugandha Gupta, 2015). Therefore, our results should include cognitive and usability heuristics

along with task performance and collection of data (qualitative & quantitative) about the user's previous experiences and preferences. "Time" can be one of the important metrics in our usability study while evaluating visualizations. Data Visualization should provide to the viewer the best possible interpretation in the shortest possible time (Barcellos, Viterbo, Bernardini, & Trevisan, 2018). Testing with Questionnaires help summarizing results. Representative test users use one or more interfaces for some typical tasks and are asked to rate the interfaces on a series of scales (Keith, 2008).

Now not all methods are tailored to perfection to collect the information aimed for. They may fail when not integrated well. At this current level, above mentioned methods (testing and enquiry) can collect all the intended data, but does the participant understand the proficiencies and limitations of a prototype? Are they able to put themselves in a situation that is close to reality and imagine themselves doing the task? Maybe not. Our goal is to encourage an approach to evaluation that is based on evaluation goals and questions instead of methods, and to encourage our community to adopt a wider view on the possibilities for evaluation in information visualization (Lam et al., 2012). The best results can often be achieved by combining several methods (Jakob, 1994). The ideal way to collect accurate, unbiased information is to encourage participants by giving as many details as possible and creating a realistic situation. This helps them in responding with realism, concreteness and precision. Adopting an expressive strategy as opposed to a

prescriptive methodology. Mocking a real situation in a lab setting is very important in order to allow the participant to live the situation.

Usability evaluation can occur at different stages of project life-cycle such as pre-design phase, design phase, prototyping phase, deployment phase and redesigning phase. AIM is a high-fidelity prototype stimulating main functionalities (Full-text search and data visualization) and design details. The full-text search feature allows the user to search industry and business variants, supply chain processes, parcels of lands and real-estate data accompanied with geographic components in the form of latitude & longitude coordinates, addresses, cities, and postal codes. The prototype mocks full-text search capabilities inspecting all the words and numbers of every document and tries to match the search criteria. Data visualization feature allows users to convert data into charts and graphs making users understand, interpret and visually illustrate relationships in the data filtered based on search criteria. The scope of evaluating an AIM prototype is to deploy it in a relevant scenario to see how visualizations influence workflows and support processes to assess the visualization's effectiveness and its use in the field (Lam et al., 2012). The scenarios for understanding data analysis are: Understanding environments and work practices (UWP), Evaluating visual data analysis and reasoning (VDAR), Evaluating communication through visualization (CTV), and Evaluating collaborative data analysis (CDA) (Lam et al., 2012). The approach of combining different evaluation methods and reviewing the variety of methodological alternatives can help in creating a realistic event during usability testing. Understanding the environment where the final tool will be deployed can extract formal requirements. Therefore, the scenario based

visually represented spatial data can articulate urban planning in simple, easy and transparent ways. Visualization techniques can help to convey implications of scenarios, thereby providing a useful vehicle for engaging stakeholders in discussions about alternative futures (Carter et al., 2018).

After reviewing different methods and scenarios to find the relevant evaluation methodology, Visual Data Analysis and Reasoning (VDAR) is chosen. VDAR studies how visualizations support in generating actionable and relevant knowledge in a domain (Lam et al., 2012). The main ability of VDAR is to examine how visualizations persuade users to agree with facts and provide evidence, and opportunities to interpret information. Results will be both quantitative metrics such as numbers their significance and qualitative feedback on experiences and opinions. This method supports evaluation of a tool as a whole unlike heuristic evaluation. It is more inclined towards process orientation than identifying issues in visual and interface elements such as colors, shapes, images, fonts, buttons, menus, icons, etc., It looks at how an integrated visualization tool as a whole supports the data exploration, knowledge discovery, hypothesis generation and decision making, rather than studying an interactive or visual aspect of the tool in isolation. Context in which visualizations will be used, task accomplishments, target users, their characteristics, environment the tool will be deployed in, and how they are adding value to solve current challenges are a few more to add on. Also, since we are evaluating a high-fidelity functioning prototype and VDAR requires fairly well developed and reliable software (Lam et al., 2012) after thorough evaluation of different methods and scenarios to find relevant methodologies an amalgamation of testing – task

performance, inquiring – user data, cognitive and usability heuristics – time spent on research and decision making and VDAR scenario for understanding data visualization will be used as a plan of action.

2.3 Cognitive Load theory

Given that the human brain can process only a certain amount of information at a given time, it is critical that a designer ought to be deliberate about what the tool is asking users to do. Investment decisions are highly uncertain and need clarity on numerous areas to clear ambiguities. For example, let's assume you are planning to start a brewery in Ontario East. You aren't aware that in specific territories of Ontario East hops are difficult to find and brewers need to acquire them from far away. Now this can be a deciding factor of whether or not you can make a profit. While I can argue that Asset Information Model does the job by facilitating this research and by providing the information, as a user experience designer it's also my job to design an efficient tool/artifact that is quick in fetching the data. The impact of cognitive load on a user can make or break the future of this artifact. Since we are proposing an artifact to minimize this load, it is also important to identify and evaluate if this very proposed artifact impacts users cognitive load. The tool should allow task completion with minimal cognitive load. It also needs to consider the amount of information working memory can hold to complete a search and comprehend the results. In a traditional scenario where there is no artifact to assist decision making, the user's past experience and knowledge allows them to refer to heuristics. This knowledge is not stable, real-time and common among every individual and in-return the results turn-out to be less accurate. Regardless of whether one

contends for previous experience the high scale of vagueness and vulnerability in information leads to cognitive overload while filtering out random facts. Cognitive overload can be divided into two categories: saturation, when overload is too great to process; and pollution, when a user must process unnecessary flawed or poorly designed information (Tracy & Albers, 2006). A cognitive load evaluation of Asset Information Model adds to the learning in regard to utilization of visualizations to minimize uncertainty. Successful measurement will require a scale and there are various methods for measuring user's cognitive load.

2.3.1 Sternberg Memory Test

The sternberg memory test plans to gauge the course of events. It gauges how rapidly clients can look and recover information from their short-term memory. Before the actual test begins participants will be given a set of words or numbers in a sequence to memorize. Later, when the test on a website is on-going, participants are probed with previously given words or numbers, and will be asked if that particular word or number is a part of the sequence they memorized. If the participant takes time to confirm Yes/No while doing a task on the website, the researcher makes note of the component of the website they were using then. This means that a particular component of a website is

using high cognitive load and the designer needs to find reasons behind its poor performance and re-work on the design.

2.3.2 Tapping Test

The tapping test imposes secondary task along with the usability test. Users will be asked to tap with their non-dominant hand or a foot while performing tasks on the website. If a user misses a tap or pauses in between at certain areas, it means the user is applying extra effort and concentration to process information there, and that component of the website should be redesigned.

Both Sternberg Memory Test and Tapping Test are conducted simultaneously during a usability study and focus on inspecting user interface. As we are aiming for context evaluation of Asset Information Model, they are not the apt evaluation scales to measure cognitive load. We have adapted our usability study to data collection and task performance at a given situation, therefore, we then need to conduct a post-test to measure the overall work load required to perform the tasks. The NASA TLX test can be utilized to evaluate workload and working memory in laboratory tests. The user will be forced to rely on what they remember and provide their opinion based on memory (Tracy & Albers, 2006). The post-test analysis examines how users remember and feel about the interaction with the tool. The scale measures task load with six criteria namely, Mental demand, Physical demand, Temporal demand, Performance, Effort and Frustration level. Users are provided with a blurb associated with each criterion and a scale of 1 to 10. 1 being the lowest and 10 being the highest. The scale lets participants of the study decide

how much mental and perceptual activities such as thinking, deciding, calculating, remembering, and searching is required. How much physical activities such as clicking, closing, zooming, scrolling, shifting tabs is required. How much time pressure do they feel due to the pace at which the analysis occurs. How successful and satisfied are they in accomplishing the tasks. How hard did they work to perform the task and how frustrated they felt while completing the task.

2.4 Commercially Existing Data Visualization and Analytical Tools

2.4.1 Google Analytics

Google analytics is an open-source data analytical service offered by the Google Marketing Platform to monitor and track web traffic. It helps in understanding how users are using websites. It can be easily integrated to other platforms and can help in gathering relevant and actionable information to improve user experience. The tool can eliminate the need for analyzing data in Google docs or spreadsheets. To use this tool all you need is a Google account with admin rights for Google analytics to start collecting data from your website. It gives on the spot access to real-time reports about audience, advertising, acquisition, behavior, conversion, user flows, and so on. The data analysis and visualization feature allow users to filter, manipulate and segment the data to business needs. As seen in Illustration 4 (Klipfolio, "Google Analytics Daily Overview"), reports are generated in the form of dashboards which provides informing about your audience's demographics, number of site visits, web traffic channels, bounce rates, page views, session duration, ad campaign performance and conversion tracking, etc. Any company that has a web presence can greatly benefit from this tool. Businesses wanting to invest in

social media or online marketing can use google analytics to make informed decisions while planning their advertising budgets.

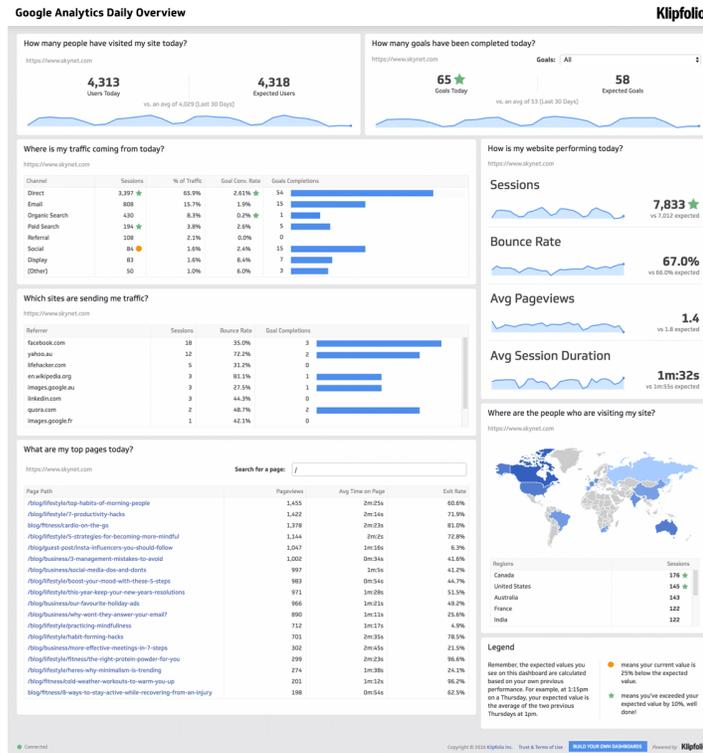


Illustration 4: Overview of Google Analytics dashboard (Klipfolio, "Google Analytics Daily Overview").

2.4.2 Tableau Software

Tableau first started in 2003 to redefine the process of analyzing data. It builds interactive data visualizations focusing on business intelligence and big data. It queries spreadsheets and cloud databases to generate reports and graphs on dashboards. Tableau also has the ability to store, extract and retrieve data from its in-memory data engine. It is a proprietary software, tailored to meet the needs of different user groups such as individuals, teams and organizations. Tableau can be used on desktops, online, servers or can be embedded into an existing infrastructure. The software has a mapping

functionality with built-in geocoding which allows to identify territorial units such as areas, cities, towns, counties, municipalities, state/provinces, postal codes, countries, etc. using latitudinal and longitudinal coordinates. It has the ability to connect to Esri Shapefiles, Keyhole Markup Language and GeoJSON to display custom geography. The software is designed to perform interactive, visual analysis to get insights in real-time and drive the business in the right direction. To have a closer look at how a dashboard in Tableau looks like refer to Illustration 5 (Anna, "Sales Summary", 2018). The platform allows collaboration with a share and comment feature where organized data can be presented to the stakeholders through interactive dashboards. The visual interface is customizable where users can play around with colors, shapes, labels, graph types, etc. Tableau provides visually pleasing data analytics to encourage businesses make better

decisions. People with little or no training can see and understand data faster than ever and in ways like never before (Tableau, 2018).

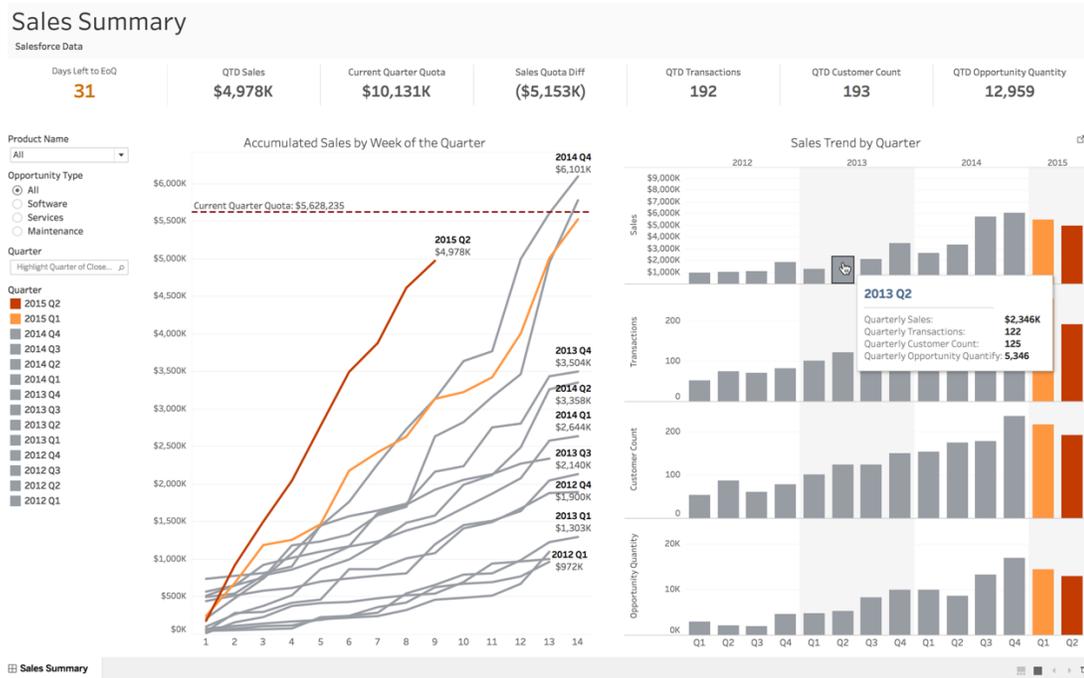


Illustration 5: Visualization created using tableau software (Anna, "Sales Summary", 2018).

Chapter 3: Design

Data representation can prompt extraordinarily wonderful pictures in addition to conveying information but as Ben Shneiderman said, “The purpose of visualization is insight and not pictures” (Yi, Kang, Stasko, & Jacko, 2008). We will discuss the thought process behind the information architecture of Data Visualization feature in Asset Information Model. The overall work flow of search initiation to completion is based on 5 steps to design data visualizations proposed by Riccardo Mazza (Mazza, 2009).

3.1 Defining the user and the problem

In the course of recent years, information and self-benefit examination has turned into a foundation of many ventures. Governments at each level have begun to grasp this new way to deal with business examination, with IT and authoritative pioneers to infer the greatest experiences and straightforwardness from their data. Our targeted end users of geospatial data are economic development officers and Business Investors (BI). Two decades prior, when you ask one hundred economic developers how they got into the field, the old adage goes, and ninety-five will tell you: by accident... (Waterhouse, 1997). Economic development practitioners and planners now belong to a recognized profession. However, as the quote suggests, this recognition is relatively recent (Leigh, 2017). Economic developer’s skills should be similar to town planners and city engineers (Waterhouse, 1997). Their profession has evolved from a loosely, often undefined occupation to its current recognizable status (Waterhouse, 1997). In most cases economic development jobs start on a contract basis and hence there is a huge possibility that not always the person doing the job has gained the spatial knowledge of the region he/she is

appointed to. The experience and working knowledge of a seasonal economic development manager is increasingly ineffective and, in many cases, even detrimental when applied to the new interconnected economic context (Leigh, 2017). Economic development, in very simple terms, is the process of businesses finding locations from which to generate marketable goods and services (Waterhouse, 1997). Two decades ago, Levy (1990), writing from an urban and regional planning perspective, characterized the profession as a mixture of rational planning and salesmanship. EDO are responsible for planning, developing, coordinating and implementing economic development policies, strategies, and initiatives that improve the community's socioeconomic, and business development. The largest portion of an economic development practitioner's efforts are spent on business attraction followed by business retention and small business development (Leigh, 2017). EDO strive for continuous development and maintain knowledge required to do their job. In order to do their job effectively they must be familiar with background materials and information related to their region's geography. While individuals of any academic background can become Certified Economic Developers, the economic development field does have academic roots in the applied disciplines of geography, business administration, public finance, political economics, and urban and regional planning (Leigh, 2017). Economic developers need to maintain an inventory of existing business sector's activities, current labor forces in the community and so on. Once the officer comes up with a project brief and a final report, initiatives and opportunities are communicated to key stakeholders or investors. EDO skills include computers and office technology, communication (both written and verbal), research and statistics, economics and economic geography, human relations and management,

education and workforce development, financing, real estate, international relations, marketing, law, government, planning (both community and strategic), business administration, and entrepreneurialism (Waterhouse, 1997).

From a different perspective, skills necessary for economic developers can be categorized as “leadership skills (communicating, planning, decision making, serving on boards, consensus building and group dynamics, and delegation); theoretical skills (globalization of the economy, general development theories, and specific economic development strategies); practical skills (computers, management, sales, and grant writing); political skills (understanding mission, funding sources, those who provide oversight and direction, competing agencies or groups, and allies, and knowing how to build and capitalize on networks and take credit when it is deserved); and change skills (becoming a lifelong learner, being willing to give up past ways of doing things, and keeping abreast of current events)” (Waterhouse, 1997). EDO are provided with computers, fax machines, internet, cellular phones, geographic information systems (GIS) and tons of unsorted data in the form of spread sheets or hard copies. Economic development data are not clear or clean (Leigh, 2017).

EDO have more data than they can handle at their table. Data on land types, business types, real-estate, main hubs, infrastructural details on main high ways, hospitals, train lines, airports, etc. The economic development practitioner constantly attempts to find good information on the local situation, ranging from demographics to institutional analysis (Leigh, 2017). All this information is not sorted well and is collected by different

departments belonging to different municipalities and counties. At times an individual responsible for collecting this information is not a part of the system as the job can be outsourced or was done by a person appointed on a contract basis. In most cases, economic development organizations are coordinating bodies that take few direct actions themselves. In such instances, the economic development practitioner is more resourceful person than an expert (Leigh, 2017).

Processes like these may lead to loss of specific expertise that's acquired while working on the ground for a long period of time and is hard to replace. This hands-on experience can also change from time to time as the economy is volatile and is dependent on several variables such as sudden change in climate, new highway construction, change in the local government power, and so on. This makes it humanly impossible to any professional to perform efficiently, make a concrete decision or market an asset to an investor with transparency. All in all, we can conclude that though people are putting a lot of effort to collect and document the data there is no substantial way to analyze the data to derive valuable insights for the end user. Our goal is to design information architecture of an Asset Information Model that can be compared to the human body describing the roles of each organ and how they are again interconnected and interdependent to one another to help economic developers of OEED function effectively.

Secondly, we must also pay attention to the fact that the AIM will not only be used by EDO but also by Business Investors. At one end there are EDO marketing their assets through AIM and on the other end there are the business investors pursuing their business

plans and strategies using the same artifact. The success of the artifact can only be justified when the EDO successfully convinces a Business Investor to invest in a property or a business. Therefore, the usability study is targeted towards Business Investors by providing them a real-life scenario.

By testing the prototype with business investors, I aim to collect insights on how AIM can be improved and what other types of information can be added based on investors current demands, experience and challenges. These insights will help in adding more intelligence to AIM. Business Investors can be anyone with the money and mindset to invest. Using this artifact, they must be able to develop familiarity with the place within the ecosystem. They must be able to evaluate specifications on industry sectors, impact on current market, business plan for risks and opportunities. Their main investment goals are to gain good information on several variables such as if the money invested will be

recovered and will fetch profits in the long run, knowing players/competitors to drive their decision.



Marcus Bernard
PASTRY CHEF

After graduating from an acclaimed culinary school and working as a Pastry chef at a five-star hotel, I decided to expand my family business. I couldn't continue working in the corporate empire when I have a huge opportunity here in my homeland.

CHALLENGES

- High risk than ever since his assistants have worked their entire life on building the brand name and the trust of customers
- Expand the business without losing the current business strategy and location

GOALS

- Develop a data base of more self-reliant and resilient local food producers network
- Invest in farming lands in-order to continue the tradition of using their own ingredients
- Wants to increase sales by setting up branches in 4 major cities

TECHNOLOGY



SKILL LEVELS

Adept organizer	80%
Business sense	70%
Outgoing	60%
Creative	50%

BIO

A 39 year old successful pastry chef working with Mandarin Oriental Las Vegas, USA comes from a family with different generations dedicated to feeding people. His father and brother run "The Bernardino" an ancestral family-style restaurant in Cowan's Bay, Ontario. The Bernardino makes food from ingredients grown on their farm and that are locally sourced. Their restaurant is well known locally and is a destination for many tourists driving to Emily Provincial Park.

Marcus wants to return home to expand his family's business along with his wife who is an MBA into a fine dining restaurant.

MOTIVATIONS

Honor	90%
Achievement	80%
Growth	70%
Power	60%
Social	50%

Illustration 6: Persona of an Investor



Helen Richardson
ECONOMIC DEVELOPMENT PRACTITIONER

I'm an Economic development practitioner in Belleville, ON with over seven years of professional experience. I promote economic growth by attracting new businesses and encouraging the expansion of existing businesses.

XLS
SPSS QGIS

CHALLENGES

- Creating new businesses/investment opportunities
- Staying up to date about local markets
- Constantly strapped for time and managing huge client base
- Most of the time goes in sorting information for marketing campaigns

GOALS

- To get a lot done in a short amount of time
- Have all the necessary information to complete tasks while working away from the office
- To stay organized and to keep track of everything

TECHNOLOGY



SKILL LEVELS

Communication	80%
Leadership	70%
Critical thinking	60%
Self-motivated	50%

TYPICAL DAY AT WORK

My day starts off by responding to phone calls, drafting documents and emails, filing and overseeing various advertising campaigns. My working time is divided between the office, meeting with community members, and traveling to outlying communities and through the province as necessary. I keep in touch with community members, to locate and meet with individuals, businesses, key contacts and government service providers.

MOTIVATIONS

Incentive	90%
Achievement	80%
Growth	70%
Power	60%
Social	50%

Illustration 7: Persona of an Economic Development Officer

3.2 Defining the data to be represented

NAICS consists of categorical data such as types of industries associated with industry classification code and locations with latitudinal and longitudinal coordinates, municipalities and counties. In the AIM prototype's data visualization, industry classification codes take on fixed variables, assigning a unit of observation on the basis of qualitative property. Illustration 8 is the summary of counts of subsectors, industry groups, industries, and Canadian industries for each of the NAICS sectors (Statistics Canada, 2012). The Ontario East Economic Development Commission represents more than 200 communities, and thousands of businesses ("Ontario East,"2017). Illustration 9 is a representation of counties and municipalities of Ontario East with their geographical boundaries. Visualizations also represent nominal categorical data that is industry

specific. For example, the agriculture sector accompanied by temperature, precipitation, wind speed and soils.

Sector code	Sector name	Sub-sectors	Industry groups	Industries	Canadian industries	Total
11	Agriculture, forestry, fishing and hunting	5	19	41	50	115
21	Mining, quarrying, and oil and gas extraction	3	5	10	29	47
22	Utilities	1	3	6	10	20
23	Construction	3	10	28	29	70
31-33	Manufacturing	21	86	181	251	539
41	Wholesale trade	9	26	72	72	179
44-45	Retail trade	12	27	58	74	171
48-49	Transportation and warehousing	11	29	42	58	140
51	Information and cultural industries	6	12	27	30	75
52	Finance and insurance	5	11	28	52	96
53	Real estate and rental and leasing	3	8	19	22	52
54	Professional, scientific and technical services	1	9	35	41	86
55	Management of companies and enterprises	1	1	1	2	5
56	Administrative and support, waste management and remediation services	2	11	29	34	76
61	Educational services	1	7	12	12	32
62	Health care and social assistance	4	18	30	37	89
71	Arts, entertainment and recreation	3	9	23	34	69
72	Accommodation and food services	2	6	10	18	36
81	Other services (except public administration)	4	14	30	38	86
91	Public administration	5	12	29	29	75
Total		102	323	711	922	2,058

Illustration 8: Summary of counts of subsectors, industry groups, industries, and Canadian industries for each of the NAICS sectors.



Illustration 9: Representation of counties and municipalities of Ontario East with their geographical boundaries.

3.3 Defining the dimensions required to represent the data

The quantity of measurements of data sets must be considered cautiously, as it will decide the conceivable data visualizations that can be utilized to represent the information. The AIM prototype represents independent dimensions. The patterns in data are not constant and are independent to location. Data can be filtered based on county/municipality and industry type to extract its contribution in the economy of that

county in percentages. This same analysis can be carried out in layers simultaneously by adding multiple fields with the same variables but in relation to each other.

3.4 Defining the structure of the data

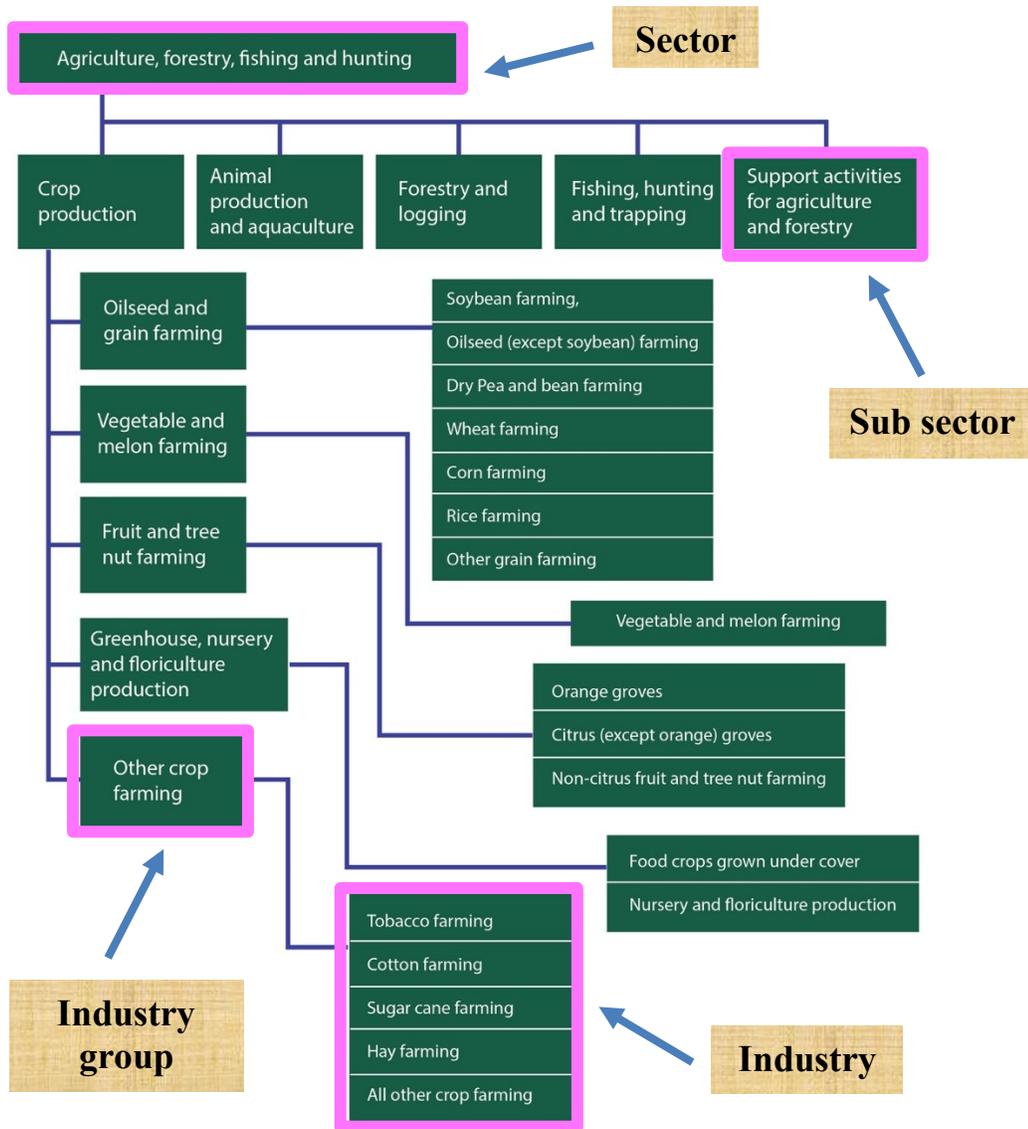


Illustration 10: North American Industry Classification System’s structure for organization of data

As seen in Illustration 10, the NIACS structure helps in understanding the existing structure of how the data is organized. The classification structure can be related to positions in a defined hierarchy. The flowchart with highlighted levels of hierarchy depicts sectors, sub-sectors, industry groups and industries. The pattern repeats throughout the North American Industry Classification System showing their order of connections between each level. The NAICS also has assigned codes to each level of hierarchy that extends taking the preceding code as they flow further. In Illustration 11,

you can see the highlighted numbers of code that are assigned and continued to extend further down in relation to classification hierarchy.

Sector	Sub-sector	Industry groups	Industry	Canadian Industry	
11	111	1111	11111	111110	
			11112	111120	
			11113	111130	
			11114	111140	
			11115	111150	
			11116	111160	
			11119	111190	
		1112	11121	11121	11121
				111219	111219
		1113	11131	11131	111310
				11132	111320
				11133	111330
		1114	11141	11141	111411
				111419	111419
			11142	111421	111421
				111422	111422
		1119	11191	11191	111910
				11192	111920
				11193	111930
				11194	111940
			11199	111993	111993
				111994	111994
				111999	111999

Illustration 11: Code classification for sectors, sub-sectors, industry groups, industries and Canadian industries

The data is also spatially, mutually exclusive. It can be classified based on county, municipality, latitude and longitude. But unfortunately, this specific data is not accessible to investors or coded in a way to quickly derive information for economic developers. The aim of our information visualization feature is to filter this data to derive geographical relationships.

3.5 Defining the interaction required from the visualization

The brain doesn't just process information that comes through the eyes. It also creates mental visual images that allow us to reason and plan actions that facilitate survival (Cairo, 2012). Representing data to communicate information efficiently by improving cognition with the help of graphics to see patterns and trends is vital. Therefore, it is also important to generate visualizations with interactions in order to provide agency to the user. The fine line between infographics and visualizations should be considered to design transformable and manipulatable models.



Illustration 12: Infographics and visualizations exist on a continuum

Graphics, charts and maps aren't just tools to be seen, but to be read and scrutinize (Cairo, 2012). In AIM the user may interact with the data in several ways, using bars, rectangles, sectors, lines and points that afford easier pointing and clicking. Rolling the

mouse cursor over any of these graphical charts, will pop up a label with a percentage contribution on the whole. In most cases, there is not just one way of encoding a particular set of data (Cairo, 2012). The prototype allows users to have a liberty to choose and shift between types of graphs (bar graphs, pie charts, line graphs, etc.) and drag and drop the positions to make close call comparisons. Visualizations are also location-centric and allows the user to filter data according to their location preference. Default interaction results are always bar charts. These are chosen because of the hierarchical structure of the data. Bar charts display only **Industry group level** information. This was chosen because the hierarchy starts to branch out drastically at **Industry level** and is difficult to produce clean and readable visualizations. Illustration 13, is an example of how a pie chart can look when generated at **Industry group level** versus **Industry level**. As mentioned earlier each bar in a bar graph is clickable and would open into a new tab generating a secondary detailed bar graph of Industry level. One of the reasons to choose this secondary interaction on bar chart is the hierarchically structured data suffering from occlusion. This interaction is required to see more than single level of information hidden within each level. Much mental integration is required to notice any patterns, and suffer from reduced spatial resolution (Shrey Gupta, Dumas, McGuffin, & Kapler, 2016). Bar charts also eliminate the over-plotting that would occur on 2D geographic maps, and also

solve the problem of widely-varying scales of distance between the location of interest (Shrey Gupta et al., 2016).

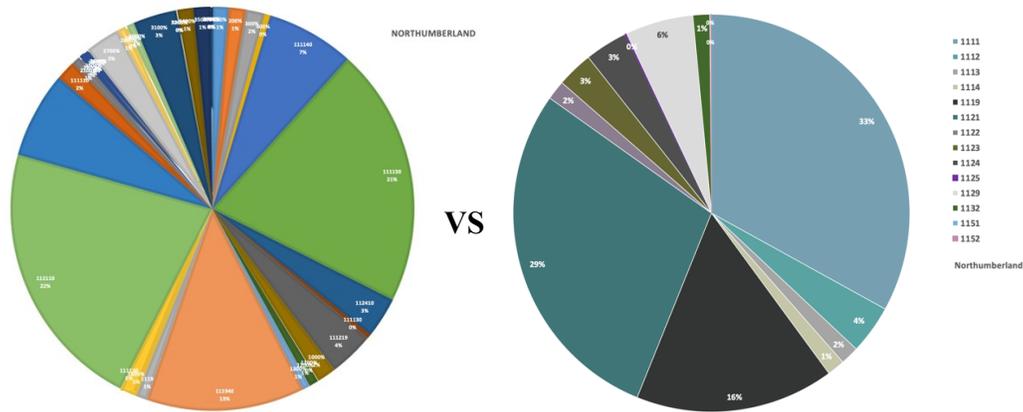


Illustration 13: Pie-charts plotting Industry group level data versus Industry level data

3.6 User interface design for AIM

The landing page of AIM displays geographical boundaries of Ontario East with the counties and municipalities (Illustration 14). It also displays different map styles overlaid with infrastructural, real-estate, and agricultural information allowing users to filter out unwanted or excess information (Illustration 15, 16, & 17). AIM has a vertical navigation on the left-hand side displaying all the features in a container that can show/hide the content with ease (Illustration 18). Every page uses this fixed vertical navigation with specific facets of location and county/municipality hidden in a dropdown feature. The side menu

with “Search for businesses” and “Data visualization” stay fixed while scrolling down the page allowing all links accessible at all times (Illustration 19 & 20).

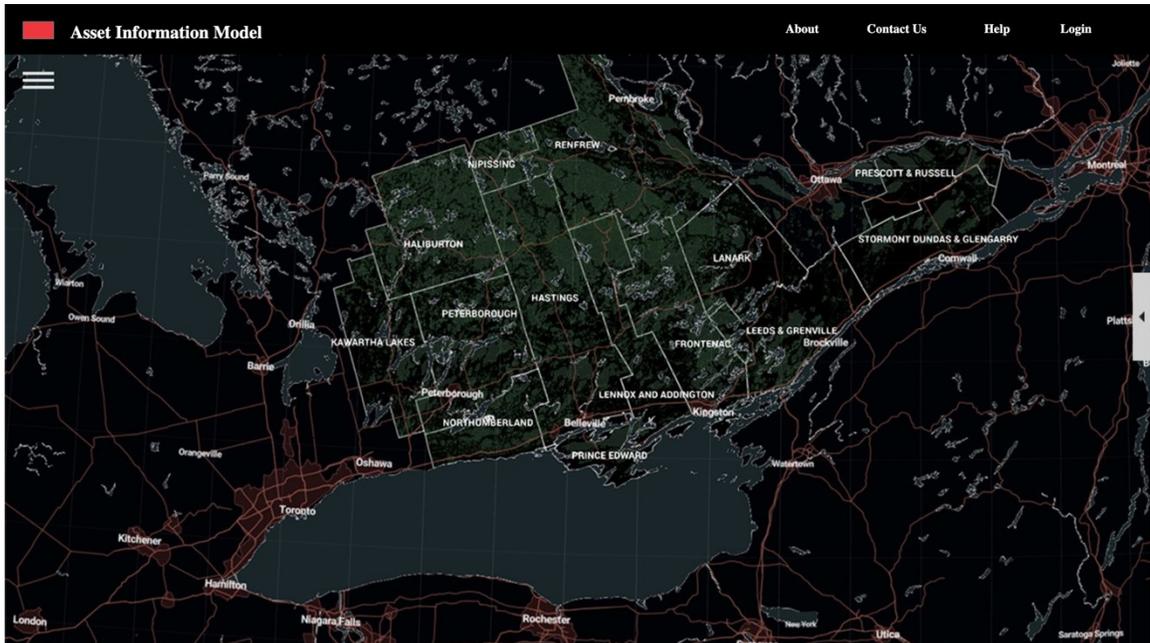


Illustration 14: Landing page.

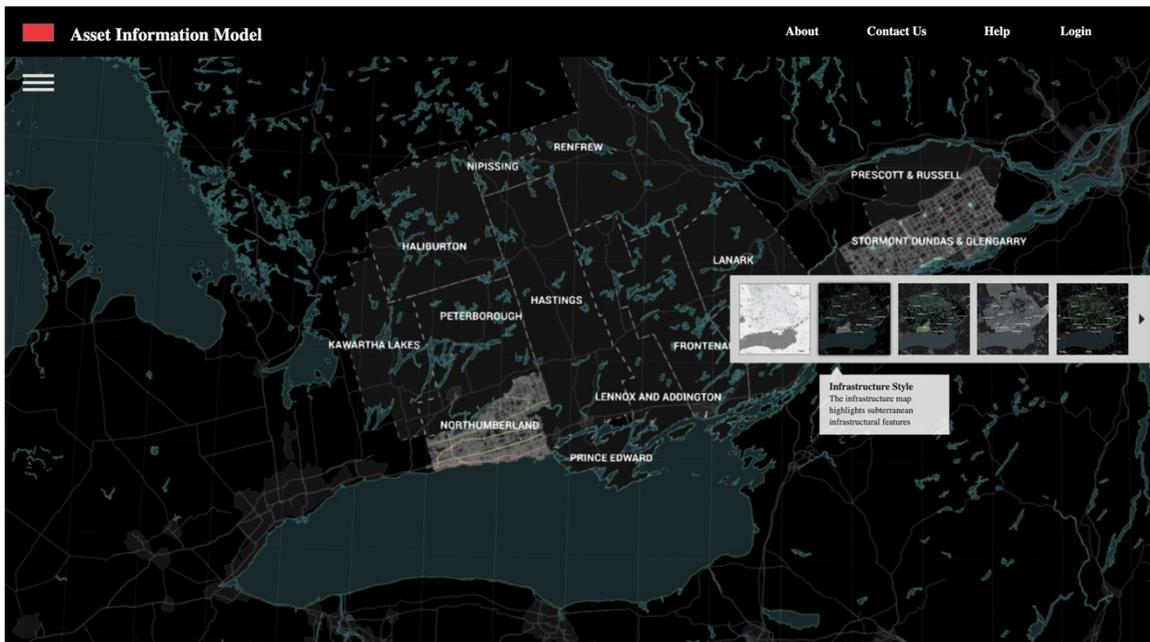


Illustration 15: Infrastructure map style.

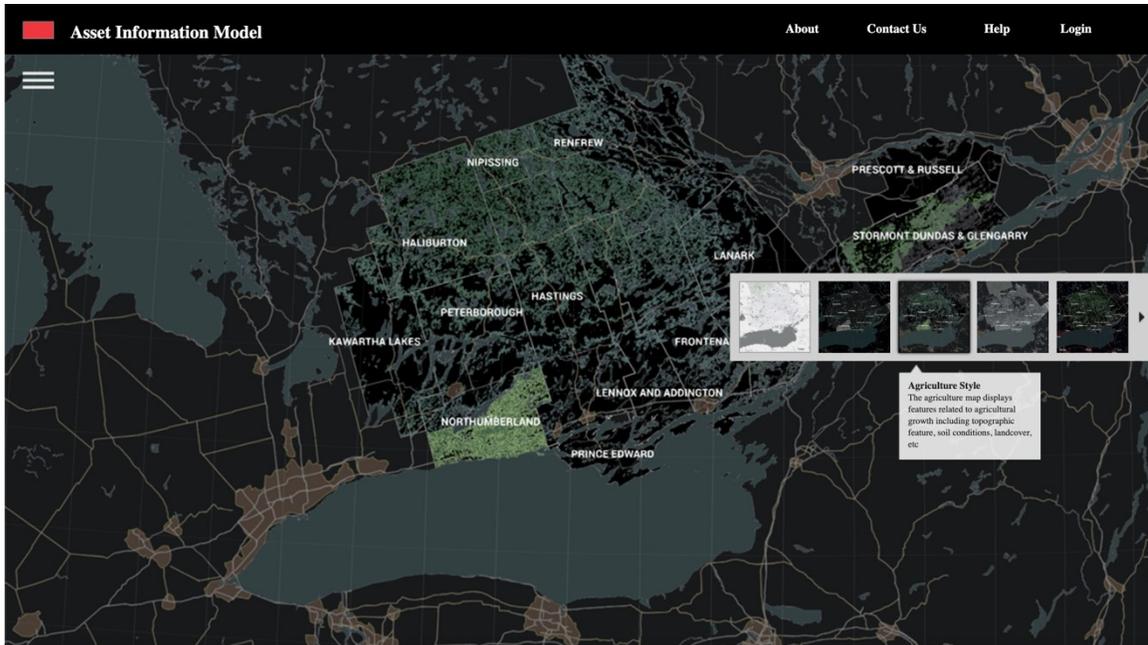


Illustration 16: Agriculture map style.

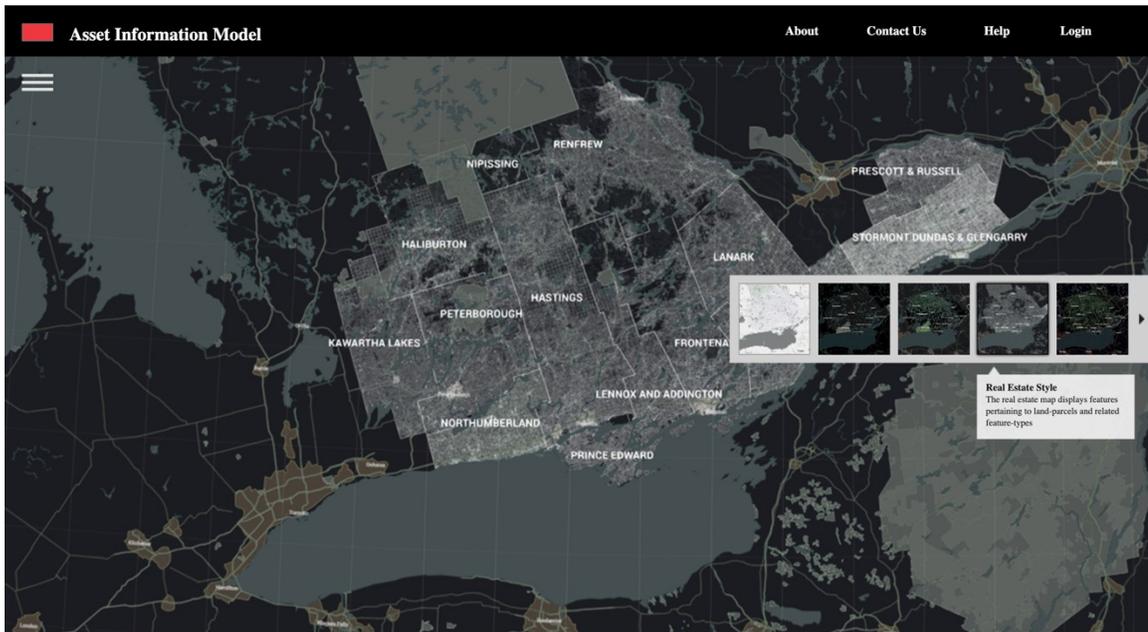


Illustration 17: Real estate map style.

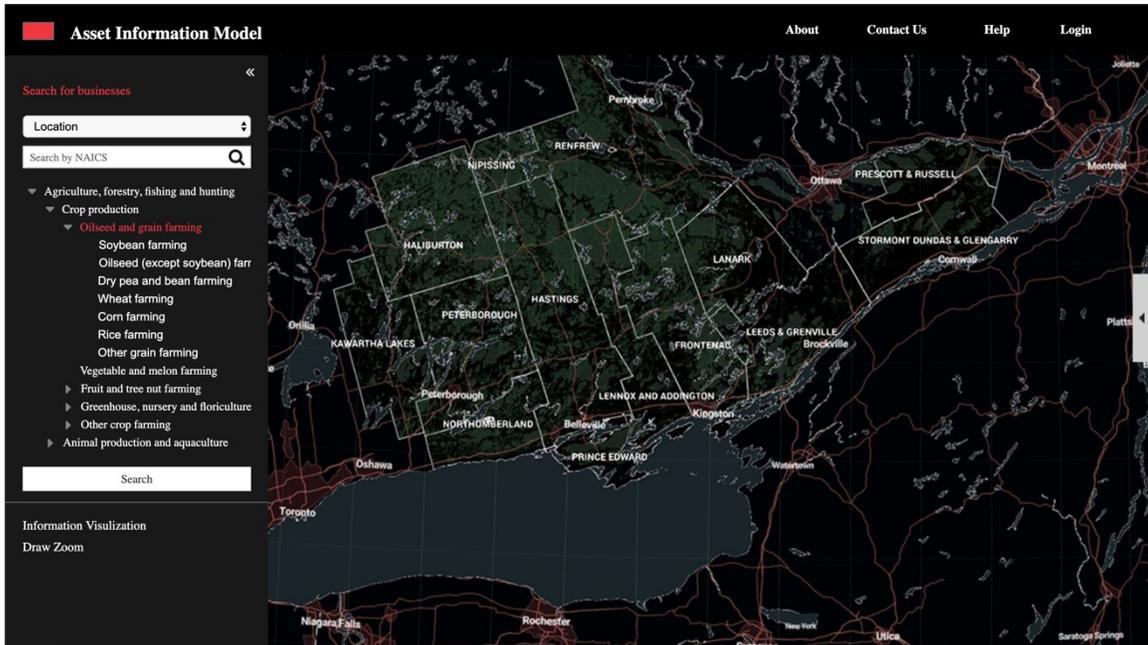


Illustration 18: Side menu.

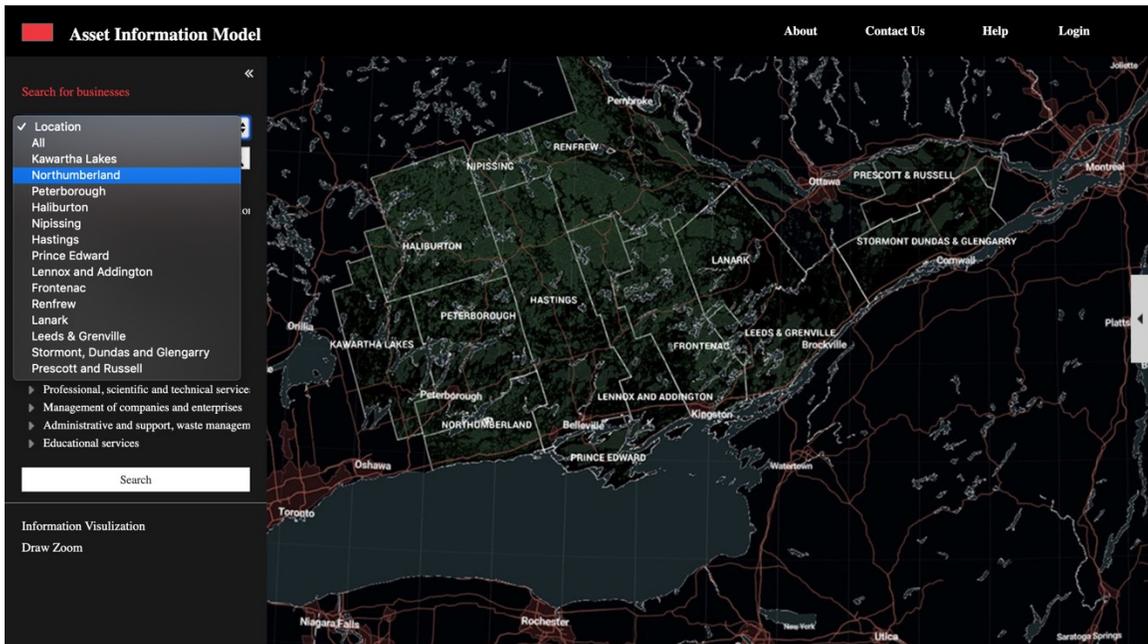


Illustration 19: Location dropdown in text search feature.

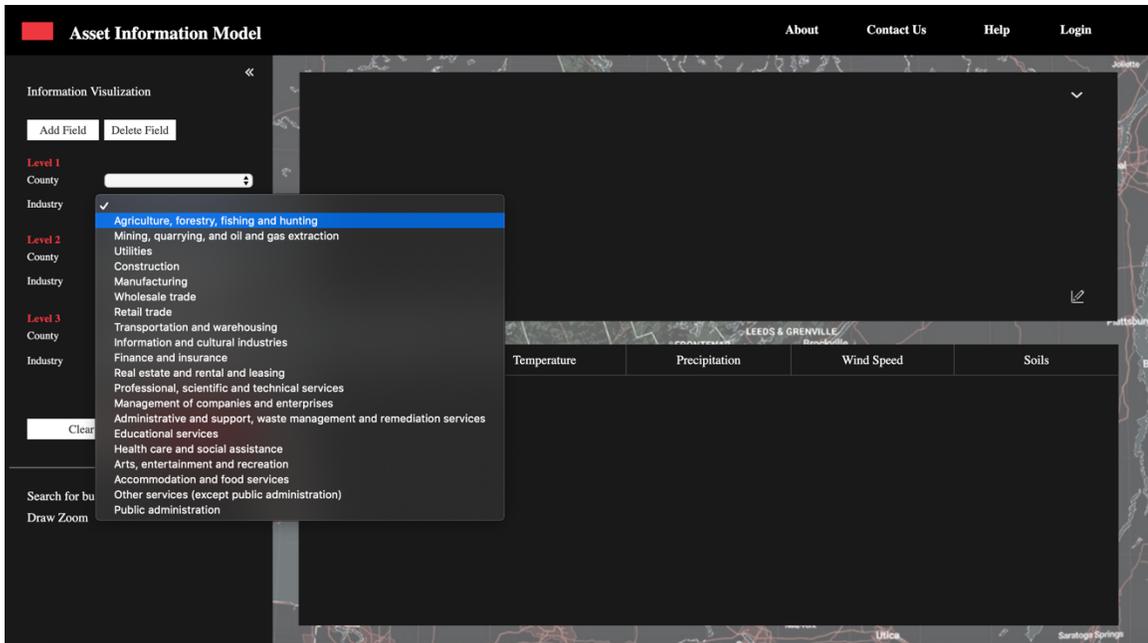


Illustration 20: Industry dropdown in data visualization feature.

3.7 Technology used for building UI

AIM prototype is built using Axure RP software. Axure RP is a wireframing, rapid prototyping software tool. It offers desired interface behavior allowing clicking, scrolling, hover effects, showing or hiding widgets to mouse interactions and touch gestures. It generates HTML links when published for preview, collaboration and review. These files can also be saved for further revisions with **.RP** extension. These data files contain project's images, text fields, buttons, etc., that were used for mockups. For the front-end

web development several java script libraries, state management tools, and database management systems are used. These are further elaborated below.

CesiumJS an open-source JavaScript library that permits the display and interaction of maps and 3D elements.

React a JavaScript library for building web interfaces. It splits the code-base into different components and is responsible for managing how the web application is displayed to the user. Every single element on the screen is a unique component. These components can be reused and edited without affecting the overall state of the application. Each new feature proposed can be added as a modular piece and inserted into the website.

Redux is a State Management Tool that focuses mainly on managing the state for the display of the application. Redux was chosen because we chose to prioritize the learnability, readability with maintainability, scalability with flexibility and database efficiency. There are 3 important things about Redux. Actions, Reducers and Store.

- Actions are dispatchers that move information from your application into the store.
- Reducers specify how the application's state changes in response to actions sent to the store. Actions only describe the user input; they do not keep track of how the overall state of the application changes.

- Store brings actions and reducers together. It holds the application state, manages the listeners, and moderates access.

PostgreSQL is used for database management along with **PuTTY** and **pGAdmin** to update and access the database. **Node.js** is used to build the front and back end. It allows the use of React, Redux and Express (a server framework for Node.js). **Gogs** a web-based host is chosen for dev modules (UI) and tiles (maps/visuals).

Chapter 4: Methodology

The goal of this research is to study how data visualizations of geospatial data facilitate and equip investors in their decision-making process. I was interested in answering the following research questions:

- How stakeholders **envisioned** Text search vs Data Visualizations?
- **Confidence levels** - How Text Search vs Data Visualizations helped stakeholders in building confidence while performing a competitive analysis?
- **Decision making** - How response times differed between Text search vs Data Visualizations.
- **Difference in mental demand, physical demand, temporal demand, performance, effort, and frustration** levels between the two navigation systems (Text search vs Data Visualizations)

To answer these questions, I have conducted a qualitative study consisting of a pre-test questionnaire to collect data from investors, who are one among the potential end users of

the Asset Information Model. This questionnaire had a total of 20 questions (multiple choice and open-ended) with a fair balance enquiring about Demographics, participants previous investment experiences and investment challenges. This qualitative study was followed by a **within -subjects scenario-based usability test** using two different prototypes accompanied by a post-test questionnaire with 6 questions (multiple choice and opinion-based). After completion of each test, participants also rated their performance with respect to the prototype on a scale of 1-10 using NASA Task Load Index.

4.1 Participants

Participants ($n = 10$) were working professionals with experience in investing on residential, rental and leasing, vacant lands, farm lands, agriculture, and hospitality. Table 1 represents Participants' age distribution. They describe themselves as Canadians (70%), Iranians (10%), Macedonians (10%), and Americans (10%). The mean number of years of participants residing in Canada was 25.15 years with standard deviation of 24.05 years. Because the Asset Information Model is proposed for attracting investors all over the world, ethnicity information was collected. We restricted eligibility to working professionals, they must be able to read, write and speak in English fluently and should have experience in using mobile phones, tablets, laptops and the like. They must also be

avid users of the internet and have previous investment experience or interest in physical properties or businesses.

Table 1

Participants Age

Number of Participants	Range
3	25 - 34
4	35 - 44
2	45 - 54
1	55 - 64
0	65 - 74

4.2 Procedure

The usability test session lasted from 45 minutes to 60 minutes and was conducted face-to-face or online based on participants preference. Participants were recruited through an invitation email sent to lead researcher's personal and professional contacts. Additionally, the snowballing technique was also used to recruit participants. Only people who responded to the email expressing interest were allowed. The invitation email also emphasized that participation is completely voluntary and non-compensated. All the participants were asked to read a consent form and provide written consent before starting the test. They were also reminded that they can choose to withdraw at any time

during the study or after completing the study. If a participant chose to withdraw during or after the study, they could inform the researcher to end the session and no justification is needed. In both cases all the data collected will be immediately destroyed.

The online study was conducted using Google Hangouts or Skype, while the face-to-face study was conducted in Carleton University or a coffee shop based on participants convenience. When interested participants contacted the researcher, their identifying information (email addresses) was protected on a password protected computer of the lead researcher. Survey responses were kept anonymous. Data provided by participants during pre and post-test questionnaires is coded, so no participant will be identifiable in any research reports.

4.3 Material

All the recruitment material made sure there was no obligation to participate in the study. All the responses of pre and post-test questionnaire and NASA TXL scale were submitted using Google forms. Google forms' responses are stored in a worksheet that can only be accessed through the lead researcher's google account. The transmission of data uses SSL to encrypt all the responses and Google's data centers/servers are located in United States. Data may be disclosed via a court order or data breach.

When participants opted for an online study, Google hangouts or Skype were used for communication. Google hangouts is an end-to-end encrypted communication application and Google's data centers/servers are located in the United States. Skype is a Microsoft

telecommunications application program and all of the Skype video files; audio files and messages are encrypted. Skype's data centers/servers are located in the United States.

4.4 Data Analysis

4.4.1 Pre-test Questionnaire (Qualitative data)

Data collected using multiple-choice questions, and dichotomous questions are reported using pie-charts or bar graphs and are summarized in percentages with respect to their contribution to the whole sample group. Data collected using Likert-type scale questions are reported in the form of bar graphs and open-ended questions are summarized individually. Observations and assumptions of this data are further discussed in the discussions section. All of the questions here are discovery questions targeted to know what is happening at the user's end and to understand their current world to formulate a solution for their problems.

4.4.2 Post-test Questionnaire (Quantitative data)

This section consists of two parts. The first part has three multiple choice questions with only one right answer and three Likert-scale questions. Answers provided for these questions in condition one and condition two by the sample group are recorded and their statistical summary measures are reported, compared and analyzed.

In the second part of this section the NASA Task Load Index is used to assess work load while performing a task (to correctly answer the multiple-choice questions in the first part) in condition one and in condition two. This Task Load Index scale helps in deriving

an overall workload score based on the average rating on six subscales pertaining to different variables. These scales increment from very low, medium and very high estimates of Mental demand, Physical demand, Temporal demand, Performance, Effort, and Frustration. Statistical summary measures of three multiple choice questions, three Likert-scale questions and paired sample t-test statistics of six NASA TXL subscales of two conditions are compared and analyzed.

Chapter 5: Results

5.1 Pre-Test Questionnaire

Previous investment experience and familiarity

80% of the participants have declared that they have successfully invested in properties or businesses in the past and 20% of the participants have been through the process but haven't invested yet (as in Illustration 21). Out of this 60% of the participants identified themselves as "Beginners" who understand basic factors involved in investment decisions and have at least invested once. 30% identified themselves as "Experts" who are well aware of the market and are confident about different variables which are to be factored

in while investing. The rest 10% identified themselves as “Neophytes” (as in Illustration 22).

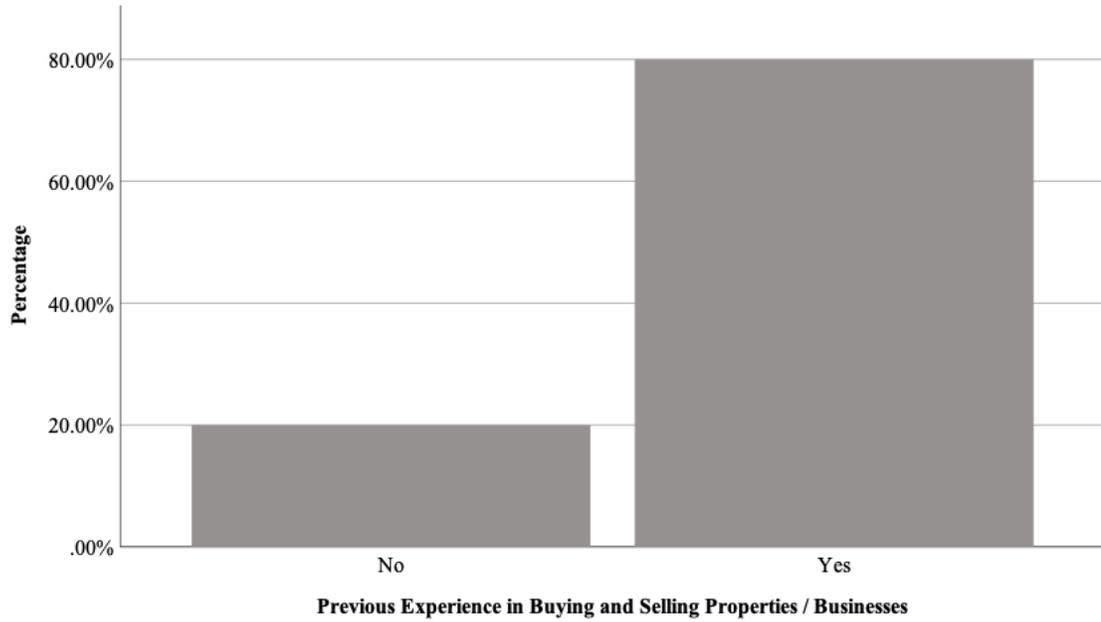


Illustration 21: Simple bar of percentage by previous investment experience

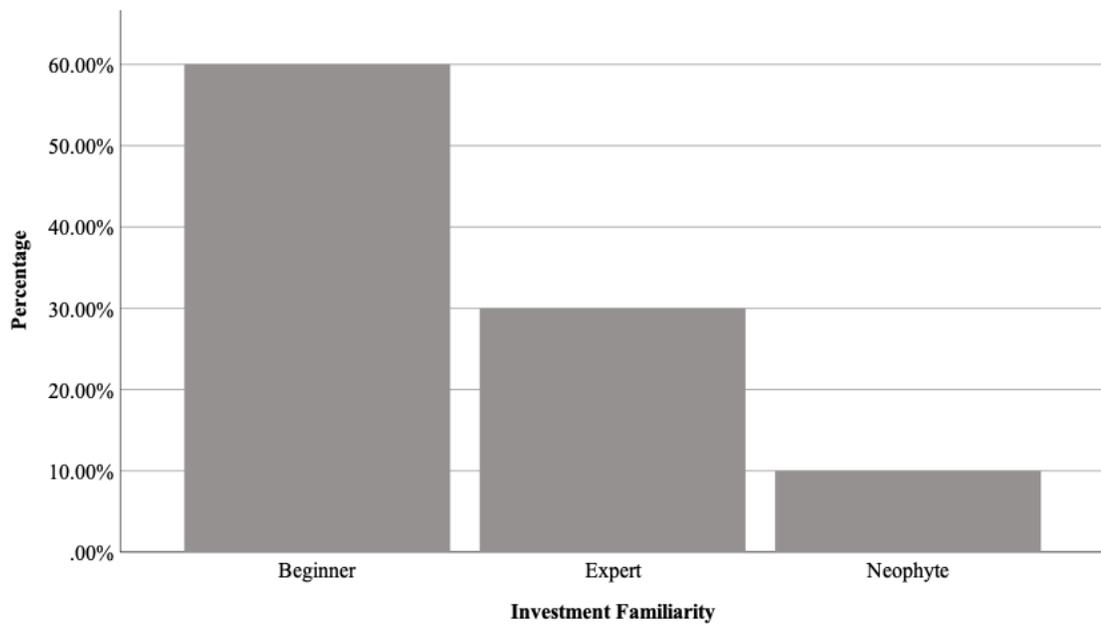


Illustration 22: Simple bar of percentage by investment familiarity

Participants investment portfolio interests

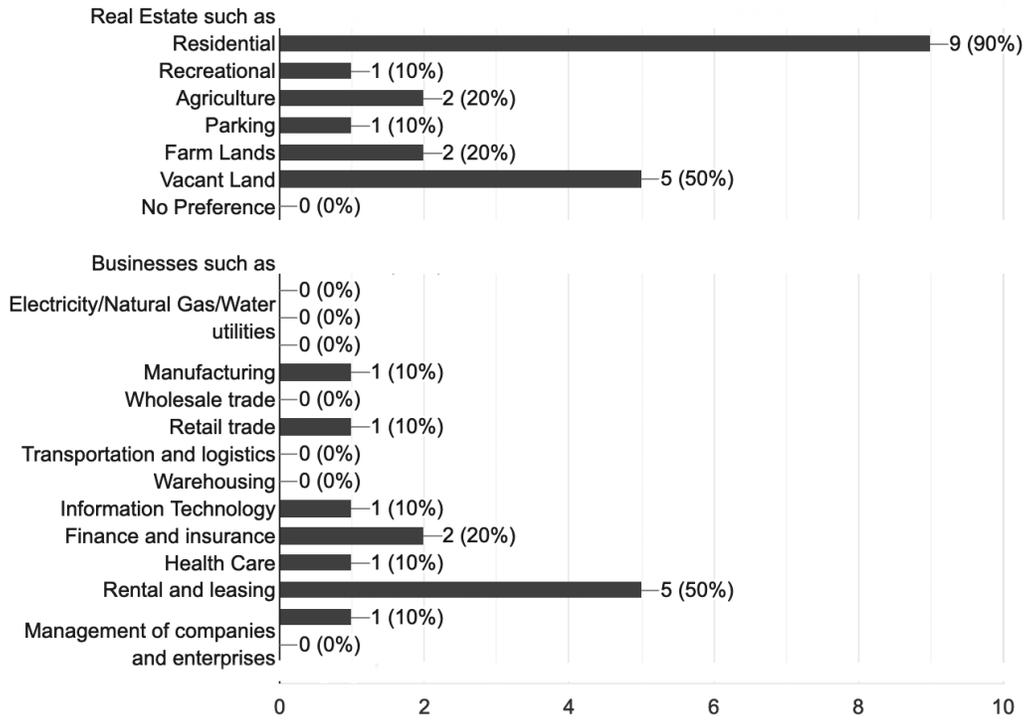


Illustration 23: Simple bar of participants interest in properties and business types by percentage

Primary goals of investment

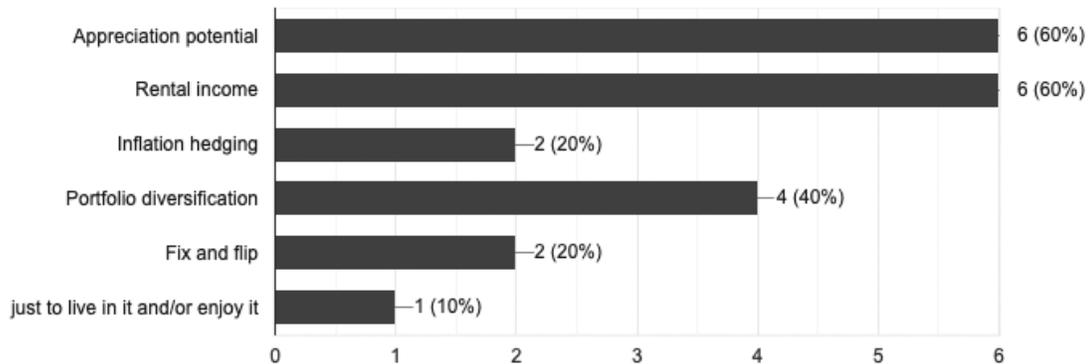


Illustration 24: Simple bar of participants investment goals by percentage

Opinion on investing in unfamiliar location

Though a majority of the participants leaned towards investing in locations that they are not currently living in or unfamiliar with, they were concerned about demographics, price, market stats, travelling distance, surroundings, and amenities. They preferred to closely inspect and manage before closing a deal.

“I don't know the demographics of other areas, prices and how competitive is the market.” (Participant 3, Jan 18, 2019). “Travel distance, unfamiliar with surround and housing and amenities, - Always like to see and inspect anything before purchase.”

(Participants 6, Jan 30, 2019). “Depends on my comfort level of the investment and how

far away it is.” (Participants 7, Jan 31, 2019). “I want to be able to manage it closely.”
(Participants 10, Feb 19, 2019)

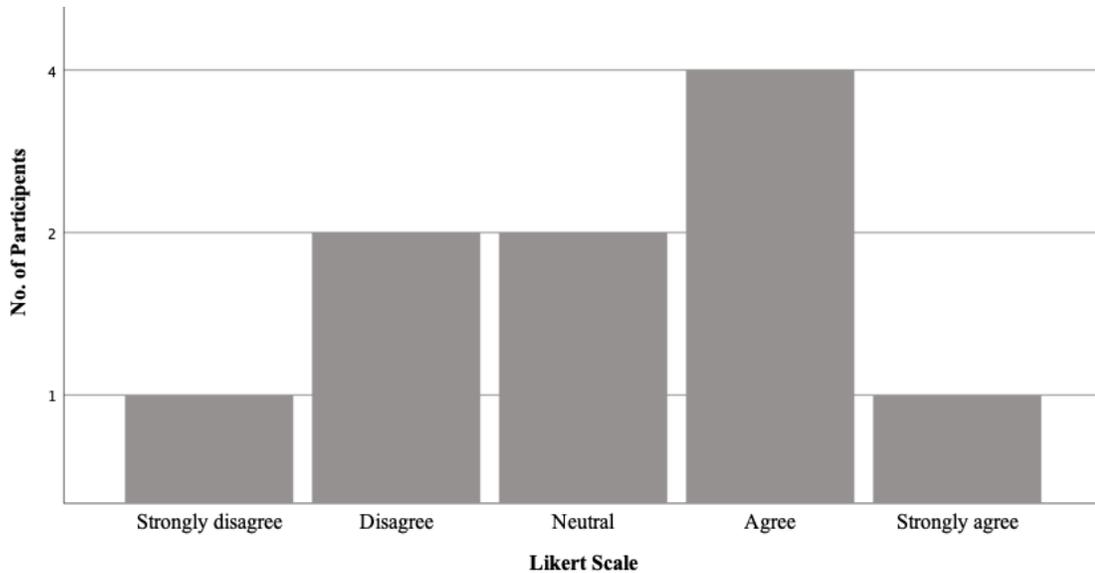


Illustration 25: Simple scale of Participants agreement on investing in an unfamiliar location

Importance of property’s location, neighborhood, and accessibility

Not only did participants strongly agree that location, neighborhood and accessibility plays a strong role in their investment decision but they have also mentioned that they would be keen on knowing other aspects such as neighborhood facilities, size of the land, construction type, property’s current condition, age, renovation history, crime rate, job

opportunities around, schools, other buildings and industries (example: waste facilities, Gas stations), rent potential, current and future developments, economy, traffic, etc.,

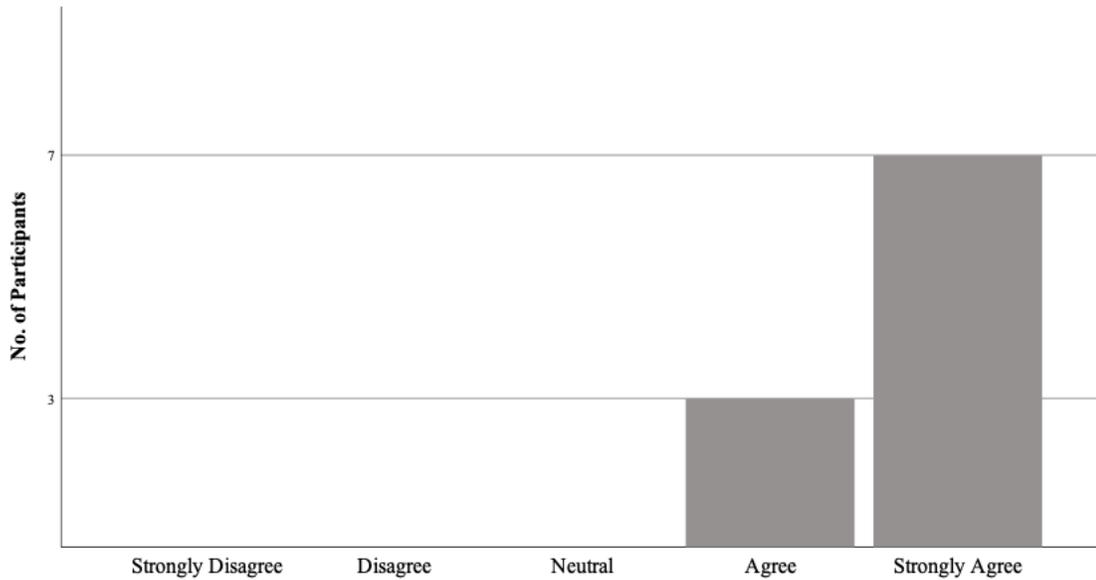


Illustration 26: Simple scale of participants agreement on considering property’s location/neighborhood and accessibility

Biggest frustrations of investing

Unknown and undisclosed details of a property or a location, market price fluctuation, investment efforts versus investment returns, legal obstacles, difficult to understand operational language.

Measures to build confidence

Market research, second opinion, online research, consulting people with experience, finding property’s growth potential in the past 10 years, consulting multiple inspectors, contractors and lawyers to track all the past owners and exchanges of the property,

Municipal assessment and taxes, maintenance costs (gas, water, hydro, etc.), news/media, consulting realtors and brokers.

“I study the market rate and how it has been performing for the last 5 to 10 years and its growth potential” (Participant 4, Jan 20, 2019). “Consulting with people that have experience” (Participant 3, Jan 18, 2019). “Speak to neighbors, word of mouth, being optimistic... Lawyer tracking of all past owners and exchanges on the property, - Municipal assessment of the building and taxes for previous years and look into charges / cost for maintenance (gas, electrical, furnace, etc.)” (Participant 6, Jan 30, 2019). “Research, local opinions, outside perspectives less emotional with risk” (Participant 9, Feb 06, 2019). “Market research, word of mouth, realtors and brokers” (Participant 10, Feb 19, 2019).

Tools and assistance

60% of the participants have previously taken help of a real-estate agents or used applications such as realtor.ca to analyze the market as they felt the need for guidance. Information gathered from real-estate agents, lawyers, financial advisers were again cross checked with credible sources such as news, banks, media and other government sources. Real-estate agents were frequently hired because they were believed to have expertise in the market, knowledge of the geographical area, soil conditions, other land uses, and future prospects.

“I would go with a real estate agent for the residential properties who can work on finding the right property in my price range and also in the neighborhood I am looking

for” (Participant 4, Jan 20, 2019). “I am using data from different government sources to analyze the market” (Participant 5, Jan 26, 2019). “News/media/bank/ accessible, reliable, credible source (Participant 7, Jan 31, 2019). “Sometimes hire an agent, review bylaws, conduct research (soil conditions, location of other land uses, designations, etc.)” (Participant 8, Feb 04, 2019). “Real estate expert, lawyer consultant, financial adviser” (Participant 9, Feb 06, 2019). “Realtor app, and Real-estate agent. Because I need guidance” (Participant 10, Feb 19, 2019).

Investment challenges

Time and reaching to a minimal risk strategy were referred as the main challenges during an investment process.

“Negotiating the deal and closing it in a timely manner” (Participant 4, Jan 20, 2019).

“The time for research and to find the appropriate professional to close the deal”

(Participant 5, Jan 26, 2019). “Negotiate the deal and wait on all the inspections to clear,

etc...” (Participant 6, Jan 30, 2019). “Time, research on the right deal, building my own

confidence in the decision” (Participant 7, Jan 31, 2019). “Opportunity cost of other

investments i.e. Stock Market vs property. Quickest way to a return with minimal risk”

(Participant 9, Feb 06, 2019). “All the obstacles with the real state rules, lawyers, not

understanding well the legal aspects of things, make sure you are doing the right thing”
 (Participant 10, Feb 19, 2019).

5.2 Task 1 – Text Search

Table 2

Post-test Questionnaire Descriptive Statistics

		Q1	Q2	Q3	Q4	Q5	Q6
N	Valid	10	10	10	10	10	10
	Missing	0	0	0	0	0	0
Mode		1	11	1	3	4	3 ^a
Range		10	7	2	3	3	3
Minimum		1	4	1	1	2	2
Maximum		11	11	3	4	5	5

a. Multiple modes exist. The smallest value is shown

As you can see from Table 2. In Task 1, most frequently participants found the right answer for Q1 i.e., they chose “**Oil seed and grain farming**” as the most contributed industry in Northumberland. But also, most frequently they were not able to find the answer for Q2 and chose “**Not sure**” for the least contributed industry of Northumberland. For Q3, most frequently participants felt there was a potential gap for greenhouse business in Northumberland. But, for Q4, they were neutral when asked about how confident they were about pursuing their business idea even after learning about the potential gap. For Q5, most frequently participants agreed on using this tool

than not having anything but for Q6 they were neutral on recommending this tool to others.

Table 3

NASA – TLX Descriptive Statistics

		Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
N	Valid	10	10	10	10	10	10
	Missing	0	0	0	0	0	0
Mean		7.40	5.30	4.60	4.30	6.00	5.90
Mode		5 ^a	2 ^a	2 ^a	3	8	7
Std. Deviation		1.838	2.869	2.989	2.312	2.828	2.644
Variance		3.378	8.233	8.933	5.344	8.000	6.989
Range		5	8	9	6	9	7
Minimum		5	2	1	2	1	2
Maximum		10	10	10	8	10	9

a. Multiple modes exist. The smallest value is shown

5.3 Task 2 – Data Visualization

Table 4

Post-test Questionnaire Descriptive Statistics

		Q1	Q2	Q3	Q4	Q5	Q6
N	Valid	10	10	10	10	10	10
	Missing	0	0	0	0	0	0
Mode		1	4	1	4	5	5
Range		3	6	2	2	1	1
Minimum		1	4	1	3	4	4
Maximum		4	10	3	5	5	5

As you can see from Table 4. In Task 2 most frequently participants found the right answer for Q1 & Q2 i.e., they chose “**Oil seed and grain farming**” as the most contributed industry in Northumberland and “**Greenhouse, nursery and floriculture production**” as the least contributed industry of Northumberland. For Q3 most frequently participants felt there was a potential gap for greenhouse business in Northumberland and for Q4 they were frequently **confident** about pursuing their business idea after learning about the potential gap. For Q5 most frequently participants **strongly**

agreed on using this tool than not having anything and for Q6 most frequently they **strongly agreed** on recommending this tool to others.

Table 5

NASA – TLX Descriptive Statistics

		Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
N	Valid	10	10	10	10	10	10
	Missing	0	0	0	0	0	0
Mean		2.70	2.20	2.00	8.20	2.00	2.20
Mode		2	2	2	9 ^a	2	2
Std. Deviation		1.767	.919	.667	2.394	.816	.919
Variance		3.122	.844	.444	5.733	.667	.844
Range		6	3	2	8	2	3
Minimum		1	1	1	2	1	1
Maximum		7	4	3	10	3	4

a. Multiple modes exist. The smallest value is shown

5.4 Task Load Index of Task 1 & Task 2

5.4.1 Mental demand

Table 6

Mental Demand Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	7.40	10	1.838	.581
	Task 2	2.70	10	1.767	.559

Table 7

Mental Demand Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	-.335	.344

Table 8

Mental Demand Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference		t	Sig.	
					Lower	Upper		df (2-tailed)	
Pair 1	Task 1 - Task 2	4.700	2.946	.932	2.593	6.807	5.045	9	.001

A dependent sample t-test showed that on average, participants experienced significantly greater mental demand while using text search ($M = 7.40$, $SD = 1.83$) than while using data visualizations ($M = 2.70$, $SD = 1.76$), $t(9) = 5.04$, $p = 0.00$.

5.4.2 Physical demand

Table 9

Physical Demand Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	5.30	10	2.869	.907
	Task 2	2.20	10	.919	.291

Table 10

Physical Demand Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	.059	.871

Table 11

Physical Demand Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Task 1 - Task 2	3.100	2.961	.936	.982	5.218	3.311	9	.009

A dependent sample t-test showed that on average, participants experienced significantly greater physical demand while using text search ($M = 5.30$, $SD = 2.86$) than while using data visualizations ($M = 2.20$, $SD = 0.91$), $t(9) = 3.31$, $p = 0.00$.

5.4.3 Temporal demand

Table 12

Temporal Demand Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	4.60	10	2.989	.945
	Task 2	2.00	10	.667	.211

Table 13

Temporal Demand Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	-.279	.435

Table 14

Temporal Demand Paired Samples Test

		Paired Differences							
		95% Confidence Interval of the Difference							
		Mean	Std. Deviation	Std. Error	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Task 1 - Task 2	2.600	3.239	1.024	.283	4.917	2.539	9	.032

A dependent sample t-test showed that on average, participants experienced significantly greater temporal demand while using text search ($M = 4.60$, $SD = 2.98$) than while using data visualizations ($M = 2.00$, $SD = 0.66$), $t(9) = 2.53$, $p = 0.03$.

5.4.4 Performance

Table 15

Performance Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	4.30	10	2.312	.731
	Task 2	8.20	10	2.394	.757

Table 16

Performance Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	.108	.766

Table 17

Performance Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference		t	Sig.	
					Lower	Upper		df (2-tailed)	
Pair 1	Task 1 - Task 2	-3.900	3.143	.994	-6.148	-1.652	-3.924	9	.003

A dependent sample t-test showed that on average, participants experienced significantly lesser performance while using text search ($M = 4.30$, $SD = 2.31$) than while using data visualizations ($M = 8.20$, $SD = 2.39$), $t(9) = 3.92$, $p = 0.00$.

5.4.5 Effort

Table 18

Effort Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	6.00	10	2.828	.894
	Task 2	2.00	10	.816	.258

Table 19

Effort Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	.048	.895

Table 20

Effort Paired Samples Test

		Paired Differences						
		Mean	Std. Deviation	Std. Error	95% Confidence Interval of the Difference		t	Sig. (2-tailed)
					Lower	Upper		
Pair 1	Task 1 - Task 2	4.000	2.906	.919	1.921	6.079	4.353	.002

A dependent sample t-test showed that on average, participants experienced significantly greater effort while using text search ($M = 6.00$, $SD = 2.82$) than while using data visualizations ($M = 2.00$, $SD = 0.81$), $t(9) = 4.35$, $p = 0.00$.

5.4.6 Frustration

Table 21

Frustration Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	5.90	10	2.644	.836
	Task 2	2.20	10	.919	.291

Table 22

Frustration Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	-.220	.542

Table 23

Frustration Paired Samples Test

		Paired Differences							
		95% Confidence Interval of the Difference							
		Mean	Std. Deviation	Std. Error	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Task 1 - Task 2	3.700	2.983	.943	1.566	5.834	3.922	9	.004

A dependent sample t-test showed that on average, participants experienced significantly greater frustration while using text search ($M = 5.90$, $SD = 2.64$) than while using data visualizations ($M = 2.20$, $SD = 0.91$), $t(9) = 3.92$, $p = 0.04$.

5.4.7 Time

Table 24

Time Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Task 1	8.1150	10	2.60510	.82381
	Task 2	4.2450	10	1.91280	.60488

Table 25

Time Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Task 1 & Task 2	10	.902	.000

Table 26

Time Paired Samples Test

		Paired Differences						
					95% Confidence Interval of the Difference			
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	Sig. (2-tailed)
Pair 1	Task 1 - Task 2	3.87000	1.20832	.38211	3.00562	4.73438	10.128	.000

A dependent sample t-test showed that on average, participants needed significantly more time to complete the task while using text search ($M = 8.11$, $SD = 2.60$) than while using data visualization ($M = 4.24$, $SD = 1.91$), $t(9) = 10.12$, $p = 0.00$.

In this study, we compared two different methods of displaying geospatial information. For each method under consideration, a scenario-based usability test was carried out with participants. Each task was rated on six subscales of perceived workload, assessed by number of errors, and total time taken to complete the task. The effectiveness and performance of text search vs data visualizations are analyzed and compared.

The study demonstrated that the subjects were able to perform with a higher accuracy rate and confidence while using data visualizations. In addition, subjects were able to complete the task in less time using the data visualizations. These findings were reinforced by a task load index scale rated immediately after each test; subjects accorded a lower degree of mental demand (MD), physical demand (PD), temporal demand (TD), effort (E), frustration (F) and higher degree of performance (P) while using data visualizations.

Nonetheless, the dependent sample t-test conducted for six dimensions in NASA TLX indicated an overall higher workload scores associated with text search. Mental demand was perceived as highest among the other five dimensions in the NASA TLX. Next were the physical demand and effort required to accomplish the task. It is also crucial to notice that despite the higher MD, PD, TD, and E, subjects' performance was significantly inaccurate and inconsistent while using text search resulting in higher frustration (F). This can be observed by subject's not being able to identify the most and least contributed industries of Northumberland and majority of them feeling neutral about

pursuing their investment decision further in task one. This clearly indicates the effects of cognitive load created by different methods used for displaying large data sets.

Chapter 6: Conclusion

Chapter 1 introduced the scope of the study and established an understanding of text search, data visualizations, Asset Information Model and Ontario East Economic Development. A summary and explanation of the current state of knowledge on urban planning, usability evaluation methods, cognitive load theory, and commercially existing data visualization tools was provided in Chapter 2. Further the steps involved in formulation of Asset Information Model were explained following “the 5 steps to design data visualizations” proposed by Riccardo Mazza (Mazza, 2009). Chapter 4 described the methodology, involved in answering the research questions. Map based text search and data visualizations were evaluated in a scenario based, within subject design, usability test to compare uncertainty, accuracy and decision making in an investment process. Results showed, a significant difference of all the dependent variables between text search and data visualizations.

Economic development officers of Ontario East are aggressively looking to attract new businesses by marketing and selling to investors across the globe more than ever before through an innovative approach – Asset Information Model. After carefully reviewing their day to day tasks - data gathering and analysis, examining opportunities and

constraints of economic development and developing targeted strategies were identified as key capabilities AIM should provide.

Fragmented pieces of information in the form of text layered on the maps required much more work to read, understand and remember, making it harder to rely upon. Whereas in Data visualizations step-wise refinement of data and grouping of similar information accompanied by numbers augmented the balance between form and function. The results displayed in the form of visualizations clearly supported subject's cognitive process in decision making. The research supported the hypothesis - data visualizations significantly reduce the work load allowing the discovery of useful information with ease and support decision-making.

All in all, the prototype was widely accepted by the users and they saw value in a tool that displays governmental data. Unlike their current research strategies, they did not feel the need for validating the information found on Asset Information Model. This tool was not seen as a replacement for humans but will help in adding intelligence to the data in today's business world. Users saw value in having both the features as they strongly felt text search helped in gaining spatial knowledge where in visualizations helped in finding statistics. They said, "investment is all about numbers."

6.1 Limitations

Though a high-fidelity prototype was used during the study a complete replication of Asset Information Model couldn't be reached. Most of the results for interactions on the

prototype were predefined and were the only results the prototype could generate when a participant wished to explore. The results in the prototype were worked up on by using the geospatial data and NAICS codes. In reality though this information exists it needs data cleaning and only then the actual tool can be fully functional. Current prototype has a two-level text search feature. Users can select county as a location, and an industry from an industry hierarchical tree or by NAICS code. As users are un aware of these hierarchical levels or NAICS codes, search at sub sector, industry group and industry level became difficult during the study. Using the current feedback received from the usability study future research studies can be conducted by building fully functional search quires for more accurate results and feedback. Though the 3D globes and maps were integrated into the prototype by hosting them on an internal server they were not functional while remotely testing the prototype. This resulted in not being able to demonstrate a fully functional text search and Asset Information Model to the users.

Since AIM prototype is a highly functional and interactive model resembling a final product, skewness in the results can be anticipated. There is a possibility that participants felt positively about the design and demonstrated tolerance for usability issues. This can be avoided by conducting the study using simple wireframes. Furthermore, the order in which participants were introduced to text search and data visualizations may have influenced their rating on NASA Task load index scale and time required to complete the tasks. Participants may have gradually improved due to familiarity with the prototype and

learnability of the questionnaire. To investigate this further a between subject usability study should be conducted to support the results and argument in this study.

6.2 Discussion

Visualizations are an effort to help people understand significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easily with visualizations. Data visualizations will definitely add help in analyzing computationally, to reveal patterns, trends, and associations. Any relevant data is difficult to oversee because no meaningful decisions can be made without the analysis of relevant data. The data analysis not only drives the decision making but also takes an active part in developing strategies and methods to ensure the existence and success of organizations. Presently the AIM only displays location and industry specific information. Adding dashboards with additional information such as demographics, evolution of asset value, market stats, job markets, crime rates, taxes, etc., can add more facts to decision making process. More data variables result in more engaging and accurate conclusions. Connecting AIM to other business analytics services can help in tracking user interactions and further measure user engagement. Reports generated through these can help in further context development, making AIM a rich data source for OEED. AIM can be EDO's one stop shop for instantly generating specific map-based information and reports rather than reaching out to GIS experts or spending time on preparing client presentations. Business Investors can search for data on nearby potential sites, logistics, demographics, storage facilities, raw material producers, proximity, etc., to plan future and current business establishments. All in all,

data visualizations can accelerate economic development and investment decision making.

In this study, we limited our evaluations to domain experts with pre-defined goals. We recruited participants with investment experience not specific to an industry. For more effective user research it is important to also test AIM with EDO. As mentioned in Chapter 2. EDO have similar skills as town planners and city engineers. They may bring in specific domain knowledge, which is not clearly defined. Therefore, involving them in usability testing can reveal new and interesting findings which can be valuable for further development of AIM.

The findings from this research can provide practical knowledge for user experience designers and academic knowledge to reflect on managing and analyzing large data sets. Based on the literature review performed there is very little reference to effectiveness of data visualizations in the field of architecture and economic development. Further research with in these fields can facilitate in building new cognitive artifacts that can maintain and operate upon information structures these fields use effecting both users and their tasks.

Appendices

Appendix A Ethics Clearance



Office of Research Ethics
503 Robertson Hall | 1125 Colonel By Drive
Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The Carleton University Research Ethics Board-B (CUREB-B) has granted ethics clearance for the research project described below and research may now proceed. CUREB-B is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Ethics Protocol Clearance ID: Project # 109940

Research Team: Ms. Naga Mahathi Meruga (Primary Investigator)
Dr. Stephen Fai (Research Supervisor)

Project Title: Data Visualization of Geospatial Data to plan future business investments

Funding Source (If applicable):

Effective: **December 14, 2018**

Expires: **December 31, 2019.**

Please ensure the study clearance number is prominently placed in all recruitment and consent materials: CUREB-B Clearance # 109940.

Restrictions:

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to CUREB-B via a Change to Protocol Form. All changes must be cleared prior to the continuance of the research.
3. An Annual Status Report for the renewal of ethics clearance must be submitted and cleared by the renewal date listed above. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.
4. A closure request must be sent to CUREB-B when the research is complete or terminated.
5. During the course of the study, if you encounter an adverse event, material incidental finding, protocol deviation or other unanticipated problem, you must complete and submit a Report of Adverse Events and Unanticipated Problems Form, found here:
<https://carleton.ca/researchethics/forms-and-templates/>

Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

Please contact the Research Compliance Coordinators, at ethics@carleton.ca, if you have any questions.

CLEARED BY:

Date: December 14, 2018

Bernadette Campbell, PhD, Chair, CUREB-B

Natasha Artemeva, PhD, Vice-Chair, CUREB-B

Appendix B Recruitment Poster



Participate in a study on Data Visualization of Geospatial Data to Plan Future Business Investments

I am looking for volunteers to participate in a 90-minute session where you will be asked to engage in a series of activities focusing on decision-making practices while investing in a property or a business. During the study you will be asked to provide feedback on your experience, findings as well as the evaluation of the components of an Asset Information Modelling prototype.

Please be advised that you will not be paid or compensated for your participation.

To participate in this study, you must be:

A working professional between the age of 25-64 years
Have an experience or interest in investing in properties or businesses
Fluent in speaking, writing, and reading English
An avid user of internet
Experienced in using mobile phones, tablets, laptops, and like

Interested in participating? Please email
Mahathi Meruga at mahathimeruga@cmail.carleton.ca

The ethics protocol for this research (Clearance# 109940) has been reviewed and cleared by the Carleton University Research Ethics Board. If you have any ethical concerns with the study, please email ethics@carleton.ca.

Research Supervisor contact information:
Dr. Stephen Fai, School of Architecture /
Engineering & Design, Carleton University
email stevefai@cunet.carleton.ca

B.1 Consent Form



Consent Form

Title: Data Visualization of Geospatial Data to Plan Future Business Investments

CUREB-B Clearance Number: 109940

Date of clearance: December 14, 2018

Ethics Clearance for the Collection of Data Expires: December 31, 2019

I _____, choose to participate in a study on Planning future business investments. I acknowledge that this study aims to assess the strengths and weakness of data visualization for planning future business investments. **The researcher for this study is a Masters' student Naga Mahathi Meruga.** She is working under the supervision of Dr. Stephen Fai in **The School of Architecture / Engineering & Design.**

The goal of this study is to study how data visualizations of geospatial data facilitate and equip investors in their decision-making process by minimizing cognitive load and investment risks. This study will take 90 minutes in total. It involves a pre-test questionnaire, usability testing of a prototype and a post-test questionnaire. With your consent, notes will be taken while you are interacting with the prototype. Your identity will not be revealed and only the summary of data collected during the study will be used in reports and presentations.

Please note that you will not be paid or compensated for your participation in this study.

The information in this form is intended to help you understand what we are studying so that you can decide whether or not to participate. Your participation is voluntary, and a decision not to participate will not be used against you in any way. As you read this form, and decide whether to participate, please ask all the questions you might have, take whatever time you need, and consult with others as you wish. You have the right to end your participation in the study at any time during the session or after the session. Simply tell the researcher that you want to end the session or email the researcher any time before March 15th, 2019. If you withdraw from the study, all information you have provided will be immediately destroyed.

All responses will be kept anonymous. All research data and notes will be kept on a password protected computer of the researchers. Research data will only be accessible by the lead researcher and the research supervisor.

Page 1 of 2

**This document has been printed on both sides of a single sheet of paper.
Please retain a copy of this document for your records.**

Note: If you chose to do the study online the data will be collected using Google forms and Google hangouts/Skype. Information shared using these platforms are end-to-end encrypted and are stored in google and skype data centres located in the United States. The data collected using these platforms may be disclosed via a court order or data breach.

Once the project is completed, all the research data will be securely destroyed. Electronic data and hard copies will be erased and shredded one year after the collection or after the completion of the study.

The ethics protocol for this project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 2517 or via email at ethics@carleton.ca).

Researchers contact information:

Naga Mahathi Meruga
School of Information Technology
Carleton University
mahathimeruga@cmail.carleton.ca

Supervisor contact information:

Dr. Stephen Fai
School of Architecture / Engineering & Design
Carleton University
613-520-2600
stevefai@cunet.carleton.ca

I voluntarily agree to participate in this study. Yes No

Signature of participant

Date

Signature of a researcher

Date

Page 2 of 2

**This document has been printed on both sides of a single sheet of paper.
Please retain a copy of this document for your records.**

B.2 Task Scenario

Usability Testing

Task Scenario

You are one of those people who live in cities and share the wish to be active in food production. Kitchen gardens, backyards, barren grounds and rooftops look tempting to you.

You live in Canada and since the weather is unpredictable here, you have decided to start a greenhouse business using hydroponic techniques. You have done your research on costs, profits and labor and now you are in a search of parcel of land to invest and start the business. You found out from your real-estate agent about few parcels of lands for sale in Northumberland that are within your budget.

Competitor research

You are unaware of the existing competition in Northumberland for greenhouse farming / traditional farming and therefore, need clarity on the below mentioned aspects to start the business with as low investment risk as possible.

1. Information on **existing Greenhouses** in Northumberland
2. **Market for specific crop types** in Northumberland
3. **Weather conditions** in Northumberland

B.3 Pre-Test Questionnaire

My name is Mahathi Meruga and I am currently studying for a Master's in Human & Computer Interaction at Carleton University. Through this questionnaire I am conducting research into factors affecting the investment decision-making process and current investment practices. The questionnaire consists of 20 questions and will take no longer than 30 minutes to complete. There are two types of questions, multiple-choice questions and open questions. All responses will be kept anonymous and no one will be identified in this research. Please answer them honestly, thank you.

1. What is your age?

- 25-34 years old
- 35-44 years old
- 45-54 years old
- 55-64 years old
- 65-74 years old

2. Are you a Canadian citizen?

- Yes
- No

3. If No, you are a citizen of?

4. Do you currently reside in Canada?

- Yes
- No

5. If yes, how long you are been staying in Canada?

6. Do you have any previous experience in buying and selling or investing in properties/businesses?

- Yes
- No

7. How familiar are you with investment process?

- Expert - I have purchased investment properties before and know what I'm doing.
- Beginner - I've read a lot and know the basics, but this is my first investment.
- Neophyte: - Eager to learn, but not sure what I'm doing.

8. What type of properties or businesses did you or would you be interested to invest in?

Real Estates

- Residential

- Recreational
- Agriculture
- Parking
- Farm Lands
- Vacant Land
- No Preference
- Others (Please specify)

Businesses

- Mining, quarrying and oil and gas extraction
- Electricity/Natural Gas/Water utilities
- Construction
- Manufacturing
- Wholesale trade
- Retail trade
- Transportation and logistics
- Warehousing
- Information Technology
- Finance and insurance
- Health Care
- Rental and leasing
- Accommodation & food services (Hotels, resorts, casinos, cottages, etc.)
- Management of companies and enterprises

- Educational services (Schools, colleges, universities, training centers, etc.)
- No Preference
- Others (Please specify)

9. What are your primary goals for an investment?

- Appreciation potential
- Rental income
- Inflation hedging
- Portfolio diversification
- Fix and flip
- Others (Please specify)

10. What type of an investor are you?

- Emotional
- Logical
- Others (Please specify)

11. Would you be interested in investing on properties or businesses that are away from your current location?

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

12. If No, can you briefly describe the reasons for not being interested in investing on properties and businesses that are away from your current location?

Example: I believe in what I see

(Please specify)

13. Would you consider a property's location/neighborhood/accessibility before making an investment decision?

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

14. Things you would consider while investing in a property or a business

Example: Location

(Please specify)

15. What frustrates you the most in the investment process and why?

Example: Market's price fluctuation because you either make money or lose money (Briefly describe) _____

16. What strategies and help do you use and take to build confidence while proceeding forward with your investment decisions?

Example: Word of mouth, second opinion, market research, being optimistic, etc.
(Please specify) _____

17. Do you use any specific tools or hire an agent to build your strategy?

- Yes
- No

18. If Yes, what are they and why do you choose them?

(Please specify and briefly describe) _____

19. If you have ever hired an agent how would you describe an ideal one?

Must be experienced and knowledgeable about

20. What are your biggest challenges in closing on an investment property?

Example: The time it takes to properly research, evaluate and negotiate the deal.

B.4 Post-Test Questionnaire

My name is Mahathi Meruga and I am currently studying for a Masters in Human & Computer Interaction at Carleton University. Through this questionnaire I am conducting research into end user's point of view, learnability and usability of the prototype. The questionnaire consists of 6 questions and will take no longer than 15 minutes to complete. The questions below are to test the prototype and not you. You can't do anything wrong here and all responses will be kept anonymous. Please answer them honestly, thank you.

1. Which industry contributes the most to Northumberland Agriculture?

- Oilseed and grain farming
- Vegetable and melon farming
- Fruit and tree nut farming
- Greenhouse, nursery and floriculture production
- Other crop farming
- Cattle ranching and farming

- Hog and pig farming
- Poultry and egg production
- Sheep and goat farming
- Another animal production
- Not sure

2. Which industry contributes the least to Northumberland Agriculture?

- Oilseed and grain farming
- Vegetable and melon farming
- Fruit and tree nut farming
- Greenhouse, nursery and floriculture production
- Other crop farming
- Cattle ranching and farming
- Hog and pig farming
- Poultry and egg production
- Sheep and goat farming
- Another animal production
- Not sure

3. Do you think there is potential gap for greenhouse businesses in Northumberland? Briefly describe why?

- Yes
- No

- Not sure

4. Based on your research so far by using this prototype, how confident are you about pursuing your business idea?

Extremely not confident	Not confident	Neutral	Confident	Extremely confident

5. If I were to invest in a property or business, I would prefer using this tool than not have anything.

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

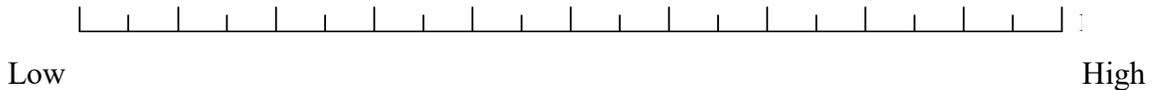
6. Would you recommend this tool to others?

Strongly disagree	Disagree	Neutral	Agree	Strongly agree

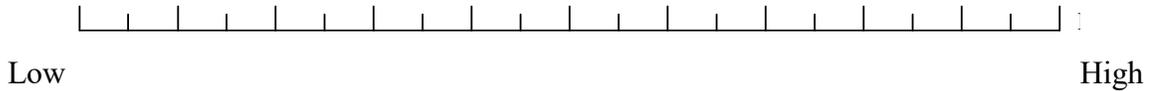
B.5 NASA-TLX Mental Workload Rating Scale

Please place an “X” along each scale at the point that best indicates your experience with the display configuration.

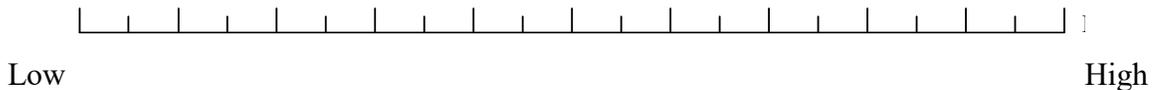
Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)?



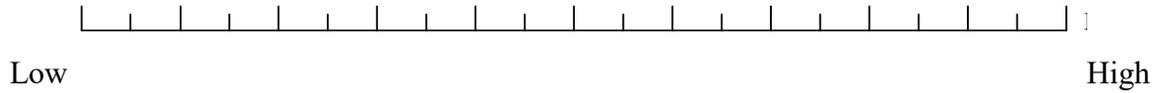
Physical Demand: How much physical activity was required (e.g., clicking, closing, zooming, scrolling, shifting tabs, etc.)?



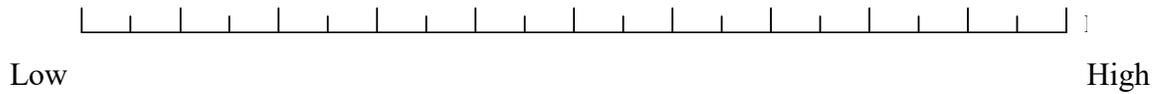
Temporal Demand: How much time pressure did you feel due to the rate or pace at which the analysis occurred? Was the pace slow and leisurely or rapid and frantic?



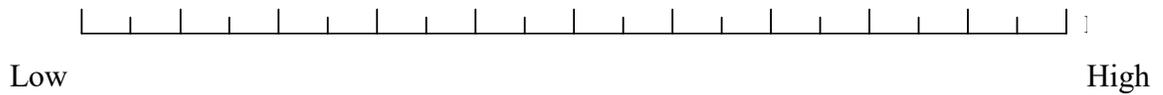
Performance: How successful do you think you were in accomplishing the goals of the study? How satisfied were you with your performance in accomplishing these goals?



Effort: How hard did you have to work (mentally) to accomplish your level of performance?



Frustration: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during the study?



Bibliography or References

- Carvajal, A. (n.d.). Quantitative Comparison between the Use of 3D vs 2D Visualization Tools to Present Building Design Proposals to Non-Spatial Skilled End Users. Ninth International Conference on Information Visualization (IV05).
doi:10.1109/iv.2005.103
- Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. (2012). Empirical Studies in Information Visualization: Seven Scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9), 1520-1536.
doi:10.1109/tvcg.2011.279
- Sedlák, P., Komárková, J., Hub, M., Struška, S., & Pásler, M. (2015). Usability Evaluation Methods for Spatial Information Visualisation - Case Study: Evaluation of Tourist Maps. Proceedings of the 10th International Conference on Software Engineering and Applications. doi:10.5220/0005557904190425
- Trujillo, A. C. (2011, August). Evaluation of Electronic Formats of the NASA Task Load Index. Retrieved from <https://ntrs.nasa.gov/search.jsp?R=20110014454>
- Kiesling, E., Gettinger, J., Stummer, C., & Vetschera, R. (2011). An Experimental Comparison of Two Interactive Visualization Methods for Multicriteria Portfolio

- Selection. *Portfolio Decision Analysis International Series in Operations Research & Management Science*, 187-209. doi:10.1007/978-1-4419-9943-6_9
- (2017, May 08). Retrieved from <https://www.statcan.gc.ca/eng/subjects/standard/naics/2012/introduction>
- Carter, O., Mitchell, M., Porfirio, L., Hugh, S., Lockwood, M., Gilfedder, L., & Lefroy, E. (2017). Mapping Scenario Narratives: A Technique to Enhance Landscape-scale Biodiversity Planning. *Conservation and Society*, 0(0), 0. doi:10.4103/0972-4923.211071
- Tracy, J. P., & Albers, M. J., Dr. (n.d.). Measuring Cognitive Load to Test the Usability of Web Sites. Retrieved September 08, 2015.
- Shamim, A., Balakrishnan, V., & Tahir, M. (2014). Evaluation of opinion visualization techniques. *Information Visualization*, 14(4), 339-358. doi:10.1177/1473871614550537
- Dudycz, H. (2011). Research on usability of visualization in searching economic information in topic maps based application for return on investment indicator. Retrieved from <https://www.cceol.com/search/article-detail?id=136194>.
- Cairo, A. (2013). *The functional art an introduction to information graphics and visualization*. Berkeley: New Riders.
- Wang, C., Pouliot, J., & Hubert, F. (2016). How users perceive transparency in the 3D visualization of cadastre: Testing its usability in an online questionnaire. *GeoInformatica*, 21(3), 599-618. doi:10.1007/s10707-016-0281-y
- Da Silva, C. G., Meidanis, J., Moura, A. V., Douza, M. A., Viadanna, P., Jr., De Oliveira, M. R., . . . De Barros, R. S. (2016). An improved visualization-based approach for

project portfolio selection. *Computers in Human Behavior*.

doi:10.1016/j.chb.2016.12.083

Dockerty, T., Lovett, A., Appleton, K., Bone, A., & Sünnerberg, G. (2006). Developing scenarios and visualisations to illustrate potential policy and climatic influences on future agricultural landscapes. *Agriculture, Ecosystems & Environment*, 114(1), 103-120. doi:10.1016/j.agee.2005.11.008

Amicis, R. D. (2009). *GeoSpatial visual analytics: Geographical information processing and visual analytics for environmental security* ; Dordrecht: Springer.

Dong, X., & Hayes, C. C. (2012). Uncertainty Visualizations. *Journal of Cognitive Engineering and Decision Making*, 6(1), 30-56. doi:10.1177/1555343411432338

Anderson, E. W., Potter, K. C., Matzen, L. E., Shepherd, J. F., Preston, G. A., & Silva, C. T. (2011). A User Study of Visualization Effectiveness Using EEG and Cognitive Load. *Computer Graphics Forum*, 30(3), 791-800. doi:10.1111/j.1467-8659.2011.01928.x

Mcardle, G., & Kitchin, R. (2016). The Dublin Dashboard: Design And Development Of A Real-Time Analytical Urban Dashboard. *ISPRS Annals of Photogrammetry*,

Remote Sensing and Spatial Information Sciences, IV-4/W1, 19-25.

doi:10.5194/isprs-annals-iv-4-w1-19-2016

Kucher, K., Paradis, C., & Kerren, A. (2017). The State of the Art in Sentiment

Visualization. *Computer Graphics Forum*, 37(1), 71-96. doi:10.1111/cgf.13217

Penteado, H. M. (2013). Assessing the effects of applying landscape ecological spatial concepts on future habitat quantity and quality in an urbanizing landscape.

Landscape Ecology, 28(10), 1909-1921. doi:10.1007/s10980-013-9940-7

Banting, P. M., & Ross, R. E. (1973). The marketing mix: A Canadian perspective.

Journal of the Academy of Marketing Science, 1(1), 1-11.

doi:10.1007/bf02729310

Google Marketing Platform - Unified Advertising and Analytics. (n.d.). Retrieved from

<https://marketingplatform.google.com/about/>

Google Analytics Daily Overview [Digital image]. (n.d.). Retrieved December 14, 2018,

from <https://www.klipfolio.com/gallery/dashboards/google-analytics-daily-overview>

Tableau. (n.d.). Retrieved December 14, 2018, from <https://www.tableau.com>

Anna, M. (2018, November 15). Sales Summary [Viz created by Tableau]. Retrieved

December 14, 2018, from

https://en.wikipedia.org/wiki/Tableau_Software#/media/File:Executive_Salesforce_Sales_Summary.png

Gupta, S. (2015). A Comparative study of Usability Evaluation Methods. *International*

Journal of Computer Trends and Technology, 22(3), 103-106.

doi:10.14445/22312803/ijctt-v22p121

- Pieters, R., & Zeelenberg, M. (2007). A Theory of Regret Regulation 1.1. *Journal of Consumer Psychology, 17*(1), 29-35. doi:10.1207/s15327663jcp1701_6
- Mazza, R. (2009). *Introduction to information visualization*. London: Springer.
doi:10.1007/978-1-84800-219-7_1
- Waterhouse, M. D. (1997). Professionalizing the Economic Developer. *New Directions for Higher Education, 1997*(97), 84-95. doi:10.1002/he.9708
- Choosing Ontario East is Easy... (n.d.). Retrieved from <http://ontarioeast.ca/>
- Gupta, S., Dumas, M., McGuffin, M. J., & Kapler, T. (2016). MovementSlicer: Better Gantt charts for visualizing behaviors and meetings in movement data. *2016 IEEE Pacific Visualization Symposium (PacificVis)*.
doi:10.1109/pacificvis.2016.7465265
- Zhu, B., & Watts, S. A. (2010). Visualization of Network Concepts: The Impact of Working Memory Capacity Differences. *Information Systems Research, 21*(2), 327-344. doi:10.1287/isre.1080.0215
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist, 38*(1), 1-4.
doi:10.1207/s15326985ep3801_1
- Faiola, A., Srinivas, P., & Hillier, S. (2015). Improving Patient Safety: Integrating Data Visualization and Communication Into Icu Workflow to Reduce Cognitive Load. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care, 4*(1), 55-61. doi:10.1177/2327857915041013

- Leigh, N. G. (2017). *Planning Local Economic Development: Theory and Practice* (6th ed.). Thousand Oaks, CA: SAGE
- McPartland, R. (2017, July 18). What is the Asset Information Model (AIM)? Retrieved from <https://www.thenbs.com/knowledge/what-is-the-asset-information-model-aim>
- Building Information Modelling Scottish Futures Trust. (n.d.). Retrieved from <https://bimportal.scottishfuturestrust.org.uk/page/what-is-bim>
- Perer, A., & Shneiderman, B. (2009). Integrating Statistics and Visualization for Exploratory Power: From Long-Term case Studies to Design Guidelines. *IEEE Computer Graphics and Applications*, 29(3), 39-51. doi: 10.1109/mcg.2009.44
- Barcellos, R., Viterbo, J., Bernardini, F., & Trevisan, D. (2018). An Instrument for Evaluating the Quality of Data Visualizations. *2018 22nd International Conference Information Visualizations (IV)*. doi: 10.1109/iv.2018.00038
- Knaflic, C.N. (2015). *Storytelling with data: A data visualization guide for business professionals*
- Andrews, K. (2008). Evaluation comes in many guises. In *AVI Workshop on BEyond time and errors (BELIV) Position Paper* (pp. 7-8).
- Nielsen, J. (1994, April). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.
- Zuk, T., Schlesier, L., Neumann, P., Hancock, M. S., & Carpendale, S. (2006). Heuristics for information visualization evaluation. *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors Novel Evaluation Methods for Information Visualization - BELIV 06*.doi:10.1145/1168149.1168162

Choosing Ontario East Is Easy... (n.d.). Retrieved from <http://ontarioeast.ca/node/16>

Task Scenarios for Usability Testing. (n.d.). Retrieved from

<https://www.nngroup.com/articles/task-scenarios-usability-testing/>

Yi, J. S., Kang, Y., Stasko, J. T., & Jacko, J. A. (2008). Understanding and characterizing insights. *Proceedings of the 2008 Conference on BEyond Time and Errors Novel EvaLuation Methods for Information Visualization - BELIV 08*.

doi:10.1145/1377966.1377971

Mazza, R. (2009). *Introduction to information visualization*. London: Springer.

How to Design an Information Visualization. (n.d.). Retrieved from

<https://www.interaction-design.org/literature/article/how-to-design-an-information-visualization>