RESEARCH ETHICS BOARDS
APPLICATION FORM

Prospective Research

This form should only be used if new data will be collected. For research involving only secondary use of existing information (such as health records, student records, survey data or biological materials), use the *REB Application Form – Secondary Use of Information for Research.*

This form should be completed using the *[Guidance for Submitting an Application for Research Ethics Review](#)*.

SECTION 1. ADMINISTRATIVE INFORMATION   [File No: office only]

Indicate the preferred Research Ethics Board to review this research:

[ ] Health Sciences OR [ ] Social Sciences and Humanities

Project Title: Exploring Knowledge Generation during Visual Data Analysis to Promote Data Analytics Democratization though Knowledge Graphs

1.1 Research team information				
Lead researcher (at Dalhousie)	Name	Leonardo Milhomem Franco Christino		
	Email (@dal)	christinoleo@dal.ca	Phone	6046559491
	Banner #	B00839709	Academic Unit	Faculty of Computer Science
Co-investigator names, affiliations, and email addresses	Taylor Hill, PhD student, Department of Psychology and Neuroscience, Dalhousie University, taylor.hill@dal.ca Fernando Vieira Paulovich, Associate Professor in Visual Analytics for Data Science, Department of Mathematics and Computer Science, TU Eindhoven, f.paulovich@tue.nl			
	Name			

Contact person for this submission (if not lead researcher)	Email		Phone	
Study start date	November 1, 2022	Study end date	January 31, 2023	

1.2 For student submissions (including medical residents and postdoctoral fellows)	
Degree program	PhD
Supervisor name and department	Evangelos Milios Faculty of Computer Science

Supervisor Email (@dal)	eem@cs.dal.ca	Phone	902-494-7111

Department/unit ethics review (if applicable). **Undergraduate minimal risk research only**.
Attestation: [ ] I am responsible for the unit-level research ethics review of this project and it has been approved. Authorizing name: Date:

1.3 Other reviews		
Other ethics review (if any) for this research	Where?	
	Status?	
Scholarly/scientific peer review (if any)		
Is this a variation on, or extension of, a previously approved Dal REB submission?	[x] No [ ] Yes Dal REB file #_____	
If yes, describe which components of the current submission are the same as the previously approved submission (list section numbers), and which components are different from the previously approved submission (list section numbers). You may also use highlighting to clearly indicate revised text.		

1.4 Funding		[x] Not Applicable
Funding (list on consent form)	Agency	
	Award Number	
	Institution where funds are/will be held	[ ] Dalhousie University [ ] Other: _____
	Was a Release of Funds issued for this award?	[ ] No [ ] Yes Date of RoF Agreement: _____

1.5 Attestation(s). The appropriate boxes *must* be checked for the submission to be accepted by the REB

[X] I am the **lead researcher** (at Dalhousie) named in section 1.1. I agree to conduct this research following the principles of the Tri-Council Policy Statement *Ethical Conduct for Research Involving Humans* (TCPS) and consistent with the University *Policy on the Ethical Conduct of Research Involving Humans*.

I have completed the TCPS Course on Research Ethics (CORE) online tutorial.

[X] Yes   [ ] No

For Supervisors (of student / learner research projects):

[x] I am the **supervisor** named in section 1.2. I have reviewed this submission, including the scholarly merit of the research, and believe it is sound and appropriate. I take responsibility for ensuring this research is conducted following the principles of the TCPS and University Policy.

I have completed the TCPS Course on Research Ethics (CORE) online tutorial.

[X] Yes   [ ] No

SECTION 2. PROJECT DESCRIPTION

2.1 Lay summary

2.1.1 In **plain language**, describe the rationale, purpose, study population and methods to be used. Include a summary of background information or literature to contextualize the study. What new knowledge, or public or scientific benefit is anticipated? [maximum 500 words]

The primary goal of *Visual Analytics (VA)* is knowledge generation. In this process, VA knowledge models and ontologies have shown to be beneficial to better understanding how users obtain new insights when executing a VA workflow. Yet, the gap between theoretical models and the practice of knowledge generation analysis is wide, and theory has mainly been used as a baseline for practical works. Also, two concepts are typically ambiguous and intermixed when analyzing VA workflows: the *temporal* aspect, which indicates sequences of events, and the *atemporal* aspect, which suggests the workflow's *state-space*, which is the set of all states of the VA tool and its user occupied during a VA workflow. Due to this and the lack of guidelines on how to *analyze* the recorded user's knowledge-gathering process led to the submission of **Visual Analytics Knowledge Graph (VAKG)**, a conceptual framework that bridges the gap between VA workflow modeling theory and application. VAKG structures a VA workflow by temporal sequences of human and machine changes over time and how they relate to the workflow's *state-space*. To apply this conceptual framework in practice, two Visual Analytic tools were enhanced to capture the user's knowledge generation process. One is an exploratory tool of quality of life (QoL) data, which provides a glimpse of the QoL status within the province (https://datatool.nsqualityoflife.ca/). The other is an undisclosed tool to visually analyze marine behavior in the Bay of Fundy. In both tools, users utilize a map with visual indicators and, through interactivity with side panels, may explore or search for insights concerning each dataset. In this study, we propose certain tasks to users and collect their process of interacting with the tools and harnessing insights from them. This data will then be structured as a knowledge graph following VAKG and be used for analysis and recommendation. With this, we aim to publish a paper that delineates the process of applying VAKG in practice, including collecting and analyzing the user behavior stored within its knowledge graph. The outcome of such analysis are many, including but not limited to "insight recommendation" to users with similar behavior, "learning to analyze" providing a learning experience to users who did not reach an insight by presenting them an aggregate of those who did, "recollection" to retrace a user's path to reach an insight and "tool performance" to analyze whether the tool was able to allow users to reach their desired insights.

[ ] This is a pilot study.

[x] This is a fully developed study.

2.1.2 Phased review. If a phased review is being requested, describe why this is appropriate for this study, and which phase(s) are included for approval in this application. Refer to the guidance document before requesting a phased review.

[x] Not applicable

2.2 Research question

State the research question(s) or research objective(s).

We have four hypotheses and one research objective:
1. **Hypothesis I** : The data collected from user interactivity, intentions and insights properly convey the user's experience and knowledge gathered while using the tool.
2. **Hypothesis II:** The structure of the collected data as a knowledge graph used in VAKG allows for intuitive recollection and analysis of a user's knowledge gathering sessions.
3. **Hypothesis III:** Through VAKG we can compare user's experiences of knowledge gathering, identifying issues with the tool used or its data.
4. **Hypothesis IV:** Through VAKG we can compare the two tools in order to improve each with lessons learned from the other.
5. **Hypothesis V:** Through VAKG we can forecast or predict what insight a new user's wishes to reach based on their behavior and/or intentions.
6. **Hypothesis VI:** Through VAKG we can forecast or predict what behaviour a new user will do based on their intentions and what insight they wish to reach.

- **Objective**: We aim for confirmation or denial of the usefulness of VAKG claimed in the theoretical literature when used in a real-world application.

2.3 Recruitment

2.3.1 Identify the study population. Describe and justify any inclusion / exclusion criteria. Also describe how many participants are needed and how this was determined.

We need between 15 participants to have a credible statistical analysis but will attempt to reach 30 participants if possible. This number of participants is selected based on the similar user studies performed in this area, such as *Explainable Matrix – Visualization for Global and Local Interpretability of Random Forest Classification Ensembles* (https://arxiv.org/abs/2005.04289) and *Q4EDA: A Novel Strategy for Textual Information Retrieval Based on User Interactions with Visual Representations of Time Series* (https://www.mdpi.com/2078-2489/13/8/368).

2.3.2 Describe recruitment plans and append recruitment instruments. Describe who will be doing the recruitment and what actions they will take, including any screening procedures.

The main population for our study is partners who aided or participated in the development of the two tools to be used in the study. The participants are employees of Engage Nova Scotia, a company which studies the Quality of Life of the province, and employees of Dalhousie, FORCE and OTN, all of which participated in prior discussions related to the undisclosed tool for marine life analysis. Participants are already aware of the study and have verbally accepted to participate. Participants will be reached out by email with details once this REB is accepted following Appendix A. Participants are also optionally allowed to disseminate the invitation throughout their organizations on their own discretion. There is no social, cultural or safety consideration, and no special inclusion or exclusion of participants. We expect the participants to be honest while doing the study. Since participation is on a volunteer basis, and the target group has shown interest in such a study, we do not have any reasons to think that the participants who may have been invited through a disseminated email will feel pressured to participate in the study.

2.3.3 If you require permission, cooperation, or participation from a community, organization, or company to recruit your participants, describe the agreement obtained from the relevant group(s). Attach correspondence indicating their cooperation and/or support (required). Describe any other community consent or support needed to conduct this research. (If the research involves Indigenous communities complete section 2.11).

[x] Not applicable

2.4 Informed consent process

2.4.1 Describe the informed consent process:

A) How, when and by whom will the study information be conveyed to prospective participants? How will the researcher ensure prospective participants are fully informed?

See appendices. The participants are already aware of the tools and of the objectives/hypothesis of this study. Once the REB is accepted, they will be presented an informed consent form before they start the study. This form contains a brief introduction of the study, information about confidentiality and anonymity of the participant's data, the participant's right to withdraw and the compensation. The form will be provided and administered by PI at the very beginning of the study.

B) Describe how consent will be documented (e.g. written signature, audio-recorded, etc).

The consent will be digitally signed before the study and will be saved with the study results.

[x] Append copies of all consent information that will be used (e.g. written consent document, oral consent script, assent document/script, etc).

Note: If the research will involve third party consent (with or without participant assent), and/or ongoing consent, ensure these are described above.

2.4.2 Discuss how participants will be given the opportunity to withdraw their participation (and/or their data) and any time (or content) limitations on this. If participants will not have opportunity to withdraw their participation and/or their data explain why.

The participants will be told that they can withdraw from the study at any time they want without penalty. They will have one week after the study is finished to ask for having their data to be discarded and not used in our analysis. This is mentioned in the consent form.

2.4.3 If an alteration/exception to the requirement to seek prior informed consent is sought, address the criteria in TCPS article 3.7A. If the alteration involves deception or nondisclosure, also complete section 2.4.4.

[x] Not applicable

2.4.4 Describe and justify any use of deception or nondisclosure and explain how participants will be debriefed.

[x] Not applicable

2.5 Methods, data collection and analysis

2.5.1

A) Where will the research be conducted?

The research will be conducted online though online forms using Microsoft Forms for demography questions and a CSV file to retain the collected data. All data will be safely stored in the SharePoint secured vault of PI's university account Microsoft Office.

B) What will participants be asked to do?

Five problems or tasks whose solutions exist in the system are given to each participant (see Appendix E). They are required to investigate the data through the system to find their answers. Each task is done expecting three types of inputs from the participant: a description of what is the intention of the participant at any given time, a description of any new insight or knowledge found by the participant at any given time, and any interactions done within the tool, such as clicking or interacting with visualizations. The three types of information are stored following VAKG's design as a CSV file. The amalgamation of all participants' data will allow us to measure and confirm/deny the hypothesis of how well VAKG can be applied in real-world scenarios. No external data is collected, such as user interaction outside the tool or any data not originating from interactions within the tool.

The users are asked to perform a specific series of actions on the tools and search for the task's answer while writing any new intention or insight during the process. When a relevant answer is found, the user should input the final insight and answer the questionnaire.

The user interfaces of both tools are implemented as a web page, and participants will have a specific amount of time to perform all parts of the study. Participants will be asked to perform the following tasks in 70 minutes:

- Participants will sign the consent form and fill the demographic questionnaire. (10 minutes)
- Participants will be given an optional short training on the system and tasks to be performed. (5 minutes)
- The post-condition (evaluation) questionnaire will be given to the users as well.
- The participants will receive 5 sets of pre-defined instructions, and they have up to 12 minutes to complete each problem.
- After finishing the user study, the participant will answer the post-condition (evaluation) questionnaire (10 minutes).
- Total time of the study is 70 minutes.

At any point, participants may take breaks of any amount of time as long as they submit the form by the deadline specified in the consent form.

C) What data will be collected using what research instruments? *(Note that privacy and confidentiality of data will be covered in section 2.6)*

The data collected will be the answers of the forms, which are found in the appendix, intentions and insights texts collected during the execution of each task and all interactions the user has done within the tool (clicks and selections). This questionnaire is of type multiple-choice, and will require the participants to perform several actions onto the tools and then to either rate an affirmation (from strongly agree to strongly disagree) or to specify what output they have found while executing the steps provided. The questionnaire result will be a table of answers for each participant. The intention texts, insight texts and interactions within the tool are recorded as a CSV file.

D) How much of the participant's time will participation in the study require?

70 minutes

[x] Append copies of all research instruments (questionnaires, focus group questions, standardized measures, etc)

[ ] This is a clinical trial (physical or mental health intervention) – ensure section 2.12 is completed

2.5.2 Briefly describe the data analysis plan. Indicate how the proposed data analyses address the study's primary objectives or research questions.

Data analysis will focus on verifying the hypothesis and research objective: verify that VAKG can be used in real-world applications. The tasks given to participants will allow us to gather the information and structure it by following the specifications of VAKG. If the result successfully allows us to perform each hypothesis listed in 2.2, we will be able to affirm our objective. For that, graph data mining techniques are planned to be used. For instance, the path-length technique allows us to check how many steps each user took from beginning to end on each task, which allows us to compare the user's opinion of how simple a task was to the number of steps they required to perform the task (Hyp1). Theoretically, we also can reproduce step-by-step what each user did during each task, a concept we aim to confirm with this study (Hyp2). By using the interaction of multiple users, we aim to compare and find similarities or dissimilarities among the population to discover any issues with the tools (Hyp3 and Hyp4) and to verify if a new hypothetical user would be able to use the aggregated insights from past users to complete tasks faster (Hyp5 and Hyp6). The techniques used will be varied, and many will be tested to verify if all hypotheses can be supported, but a small list of some of them is: page-rank, shortest-path, breadth-first search, monte-carlo prediction, graph to matrix embedding, and graph neural networks. We will also qualitatively analyze the results from tasks and questionnaires, comparing the answers and plotting the results using different charts; for instance, box plots of answers to observe distributions and scatter plots for correlation between answer distributions. Once the information is visualized, we will be able to gain insights into how data experts make use of the tool's functionalities, and what patterns they find.

2.5.3 Describe any compensation that will be given to participants and how this will be handled for participants who do not complete the study. Discuss any expenses participants are likely to incur and whether/how these will be reimbursed.

No compensation will be given.

2.6 Privacy and confidentiality

2.6.1

A) Describe who will have knowledge of participants' identities.

B) Describe the level of identifiability of the study data (anonymous, anonymized, de-identified/coded, identifying) (see TCPS Chapter 5A – types of information for definitions).

C) Specify which members of the research team (or others) will have access to participants' data and for what purpose.

D) Describe measures to ensure privacy and confidentiality of study documents and participant data during the data collection and analysis phase. *[Note that plans for long term storage will be covered in 2.6.2]*
- Address: handling of documents/data during data collection; transportation or transfer of documents/data; storage of documents/data (during the study).
- If a key-code will be maintained, describe how it will be kept secure.
- For electronic data, describe electronic data security measures, including file encryption and/or password protection as applicable.
- For hard copy documents, describe physical security measures (specify location).

There will not be any personal or sensitive data collected in this study. Our software (online user interface) will only collect relevant interactivity data within the tool and the questionnaire answers. The demographic questionnaire will also not collect sensitive data, and participants are to their discretion to answer them or not. Data will be stored in either a CSV file (text-based format of comma separated lists) or in a spreadsheet format (such as xlsx). All data will be anonymous. Only the four researchers listed at the researcher team will have access to the participant's data. The records will be stored in the university's SharePoint and allocated only to the PI's account, allowing only him and his supervisor to have access to it. The data will be deleted from the university's secure vault five years after the experiment or final publication.

[ ] This research involves personal health records (ensure section 2.13 is completed)

2.6.2 Describe plans for data retention and long-term storage (i.e. how long data will be retained, in what form and where). Will the data eventually be destroyed or irreversibly anonymized? If so, what procedures will be used for this? Discuss any plans for future use of the data or materials beyond the study currently being reviewed.

The records will be deleted from PI's university Microsoft SharePoint vault five years after the experiment or publication. The results may be used for further revisions of the same publication within the above timeframe, if needed.

[ ] This research will be deposited in a data repository (ensure section 2.14 is completed)

2.6.3

Describe if/how participant confidentiality will be protected when research results are reported:

A) For quantitative results - In what form will study data be disseminated?

  [x] Only aggregate data will be presented

[ ] Individual de-identified, anonymized or anonymous data will be presented

[ ] Other. If "other", briefly describe dissemination plans with regard to identifiability of data.

[ ] Not applicable, only qualitative data will be presented

B) For qualitative results - Will identifiable data be used in research presentations/publications? If participants will be quoted, address consent for this and indicate whether quotes will be identifiable or attributed.

[] Not applicable, only quantitative data will be presented

Anonymous excerpts of the two open-ended questions and excerpts of the intentions and insight texts provided by participants during the tasks will be used within publications. No form of specific participant identification will be used.

2.6.4 Address any limits on confidentiality, such as a legal duty to report abuse or neglect of a child or adult in need of protection, and how these will be handled. Ensure these are clear in the consent documents. (See the guidance document for more information on legal duties and professional codes of ethics).

[x] Not applicable

2.6.5 Will any information that may reasonably be expected to identify an individual (alone or in combination with other available information) be accessible outside Canada? And/or, will you be using any electronic tool (e.g. survey company, software, data repository) to help you collect, manage, store, share, or analyze personally identifiable data that makes the data accessible from outside Canada?

[] No

[X] Yes. If yes, refer to the University *Policy for the Protection of Personal Information from Access Outside Canada*, and describe how you comply with the policy (such as securing participant consent and/or securing approval from the Vice President Research and Innovation).

The PI will be for some months present in Europe and will have access to the data during this time though Dalhousie's VPN. The data will only be accessible though a secure connection through Dalhousie's own VPN encryption solution.

2.7 Risk and benefit analysis

2.7.1 Discuss what risks or discomforts are anticipated for participants, how likely risks are and how risks will be mitigated. Address any particular ethical vulnerability of your study population. Risks to privacy from use of identifying information should be addressed. If applicable, address third party or community risk. (If the research involves Indigenous communities also complete section 2.11)

No risks or discomforts are anticipated for participants caused due to the research. No identifiable data will be collected or stored. Participants are at their discretion to answer demographic-related questions, and all results will be reported as an aggregate statistic. Even considering the 70 minutes of expected time to complete the research, participants are allowed as many breaks as they wish, and each break can be as long as they wish after considering the deadline of the from submission of 1 week.

2.7.2 Identify any direct benefits of participation to participants (other than compensation), and any indirect benefits of the study (e.g. contribution to new knowledge).

Participants are actively invested in the tools they will be using during the study. The results will be given to them upon request to allow them to better plan the future of said tools and to, potentially, use VAKG again in the future.

2.8 Provision of results to participants and dissemination plans.

2.8.1 The TCPS encourages researchers to share study results with participants in appropriate formats. Describe your plans to share study results with participants and discuss the process and format.

The users will be notified by email when the publication resulting from the study is available. An official copy of the publication and a friendly summary of it will be send to the participants up on request.

2.8.2 If applicable, describe how participants will be informed of any material incidental findings – a discovery about a participant made in the course of research (screening or data collection) that is outside the objectives of the study, that has implications for participant welfare (health, psychological or social). See TCPS Article 3.4 for more information.

[x] Not applicable

2.8.3 Describe plans for dissemination of the research findings (e.g. conference presentations, journal articles, public lectures etc.).

The results will be used as a foundation to justify research in the VAKG technique and its use in real world applications for collecting and analysing user behaviour within visual analytics tools.

2.9   Research Team

2.9.1 Describe the role and duties of all research team members (including students, RA's and supervisors) in relation to the overall study.

The lead researcher will reach out to possible participants, deliver the online research process webpage, and collect the results. He will be the main author of articles with the results of this study,

therefore all statistical analysis, data retrieval and future communications with participants who requests results of the study will also be handled by him.

The rest of the research team will supervise and double-check the collected data and results. The co-investigator is focused on how to best use the results of the study to aid the participants and tools in the future. The supervisor will manage the operation to assert good progress on article publications and that the results of the study is used in a useful and relevant manner to make the system and the area of research better.

2.9.2 Briefly identify any previous experience or special qualifications represented on the team relevant to the proposed study (e.g. professional or clinical expertise, research methods, experience with the study population, statistics expertise, etc.).

The research supervisor has previous experiences with many other similar user-test research, and this one will be equivalent to the others, only differentiating on what is the tool being tested. Several articles of systems, applications and techniques were published by the Supervisor (e.g. Explainable Matrix – Visualization for Global and Local Interpretability of Random Forest Classification Ensembles), and he will follow on the same procedures, some of which also involved similar study population. The research lead is a specialist in the technical aspect (research-wise and programming-wise) and the co-researcher is a specialist in one of the tools and on the overall concept of using data to provide a better QoL to the local population. Both aim to aid on the specific work required to build and use the tools and will use the results of the study to improve certain aspects of the work in order to publish findings about the usage of VAKG in real-world applications.

2.10 Conflict of interest

Describe whether any dual role or conflict of interest exists for any member of the research team in relation to potential study participants (e.g. TA, fellow student, teaching or clinical relationship), and/or study sponsors, and how this will be handled.

[x] Not applicable

Although the research team have invested interest in improving the tools being used, the focus of the study is the usage of VAKG, which has no conflict of interest with any participant.

2.11 Research involving Indigenous peoples

Consult TCPS Articles 9.1 and 9.2 in determining whether this section is applicable to your research.

[x] Not applicable – go to 2.12

2.11.1 If the proposed research is expected to involve people who are Indigenous, describe the plan for community engagement (per TCPS Articles 9.1 and 9.2). Attach supporting letters, research agreements and other relevant documents, if available. If community engagement is not sought, explain why the research does not require it, referencing TCPS article 9.2.
2.11.2 State whether ethical approval has been or will be sought from Mi'kmaw Ethics Watch and if not, why the research does not fall under their purview. If the research falls under the purview of other Indigenous ethics groups, state whether ethical approval has been or will be sought.
2.11.3 Describe any plans for returning results to the community and any intellectual property rights agreements negotiated with the community with regard to data ownership (see also 2.11.4 if applicable). If there are specific risks to the community involved, ensure these have been addressed in section 2.8.1.
2.11.4 Does this research incorporate OCAP (Ownership, Control, Access, and Possession) principles as described in TCPS Article 9.8? [ ] Yes. Explain how. [ ] No. Explain why not.

2.12 Clinical trials [x] Not applicable – go to 2.13
2.12.1 Will the proposed clinical trial be registered? [ ] No. Explain why not. [ ] Yes. Indicate where it was/will be registered and provide the registration number.
2.12.2 If a novel intervention or treatment is being examined, describe standard treatment or intervention, to indicate a situation of clinical equipoise exists (TCPS Chapter 11). If placebo is used with a control group rather than standard treatment, please justify.

2.12.3 Clearly identify the known effects of any product or device under investigation, approved uses, safety information and possible contraindications. Indicate how the proposed study use differs from approved uses.

[ ] Not applicable

2.12.4 Discuss any plans for blinding/randomization.

2.12.5 What plans are in place for safety monitoring and reporting of new information to participants, the REB, other team members, sponsors, and the clinical trial registry (refer to TCPS Articles 11.6, 11.7, 11.8)? These should address plans for removing participants for safety reasons, and early stopping/unblinding/amendment of the trial. What risks may arise for participants through early trial closure, and how will these be addressed? Are there any options for continued access to interventions shown to be beneficial?

2.13 Use of personal health information

[x] Not applicable

2.13.1 Research using health information may be subject to Nova Scotia's *Personal Health Information Act*. Describe the personal health information (definition explained in the guidance document) required and the information sources, and explain why the research cannot reasonably be accomplished without the use of that information. Describe how the personal health information will be used, and in the most de-identified form possible.

2.13.2 Will there be any linking of separate health data sets as part of this research?

[ ] No

[ ] Yes

If yes:

A) Why is the linkage necessary?

B) Describe how the linkage will be conducted (it is helpful to append a flow diagram)

C) Does that linkage increase the identifiability of the participants?

2.13.3 Describe reasonably foreseeable risks to privacy due to the use of personal health information and how these will be mitigated.

2.14 Data Repositories
[x] Not applicable
2.14.1 Identify and describe the data repository in which the research data will be deposited. What is its focus, who are its target users, who can access deposited data and under what circumstances? For how long will the data be kept in the repository?
2.14.2 Describe the data set to be released to the repository. If there is personal and/or sensitive information in the data, describe how you will prepare the data for submission to the repository and mitigate risks to privacy. Identify all fields that will be included in the final data set (include as an appendix).
2.14.3 Is agreeing to have one's data deposited a requirement for participation in the study? If yes, provide a justification. If no, indicate how participants can opt in or out.

SECTION 3. APPENDICES

Appendices Checklist. Append all relevant material to this application in the order they will be used. This may include:

[ ] List of References

[ ] Permission letters (e.g. Indigenous Band Council, School Board, Director of a long-term care facility)

[ ] Support/cooperation correspondence

[x] Recruitment documents (posters, oral scripts, online postings, invitations to participate, etc.)

[ ] Screening documents

[x] Consent/assent documents or scripts

[x] Research instruments (questionnaires, interview or focus group questions, etc.)

[ ] Debriefing forms

[ ] List of data fields included in data repository

Consent Form Templates

Sample consent forms are provided on the Research Ethics website and may be used in conjunction with the information in the *Guidance* document to help you develop your consent form.
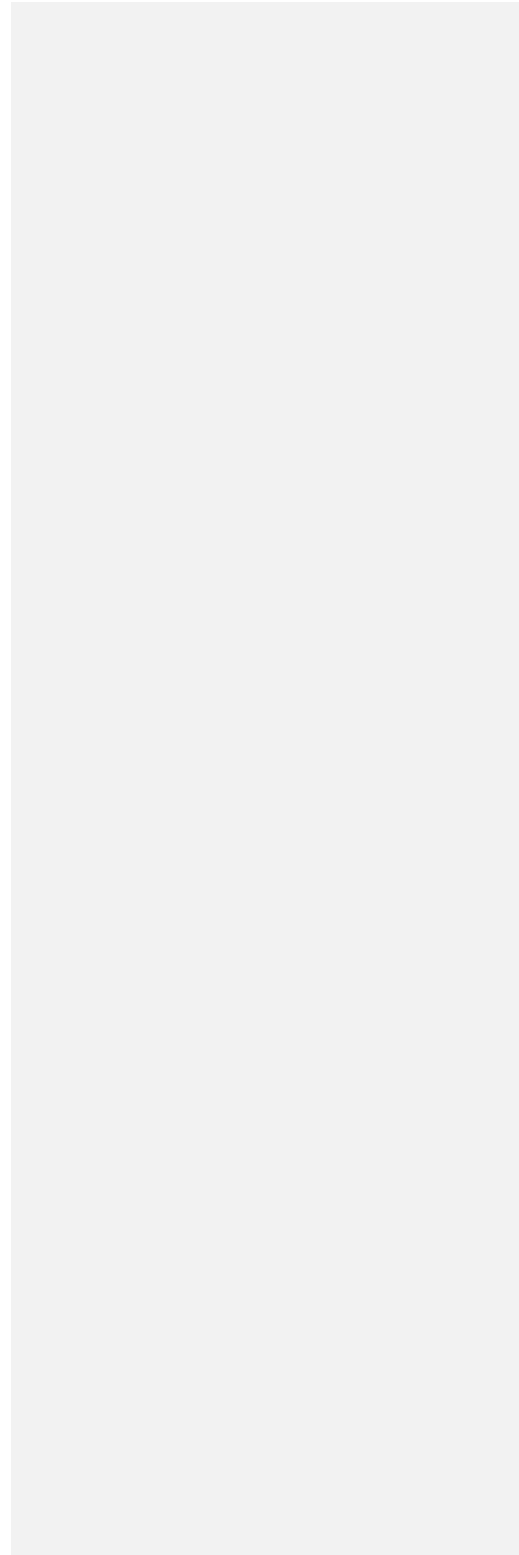
Appendix A – Email Recruitment Notice

We are recruiting 30 participants to take part in a research study to collect and examine your intentions and insights when using a web application for data analysis of the quality of life within our province and the fish and shark behavior in the Bay of Fundy. We are inviting those interested in participating in this study. You receive an overview of the study details and a consent form. If you accept the consent form, you will watch a short video on how to use each of the tools for about 5 minutes, and then you will perform pre-defined operations on them while describing your intentions and insights during the process. The entire study is online. After the study is completed, you will also fill out and submit a small evaluation questionnaire on your experience. You should be able to complete the study in 70 minutes and are allowed breaks of your own discretion, being only required to submit your form within 1 week of accepting the consent form. If you are interested in participating, please contact Leonardo Christino (christinoleo@dal.ca).

Appendix B – Consent Form

CONSENT FORM

User evaluation of VAKG in real-world applications: Analysing user's intentions and insights during visual data analysis of Quality of Life and Ocean data visualization tools

You are invited to take part in a research study being conducted by Leonardo Christino, a Ph.D. graduate student in the Faculty of Computer Science at Dalhousie University. The purpose of this research is to analyze and verify the ease-of-use and usefulness of our Visual Analytic Knowledge Graph (VAKG).

If you choose to participate in this research, you will be asked to perform pre-set operations and analysis through our Webapps while describing your intentions and insights and anonymously answering questions regarding its usability. The survey should take approximately 70 minutes, and you are allowed breaks of your own discretion.
- You will complete a demographic questionnaire regarding only your level of education and knowledge, and experience with data/visualization analysis.
- You will be given a tutorial on how to use the software.
- You will be given a practice session to use the software.
- You will be given an evaluation questionnaire.
- You will perform four tasks of performing data analysis using the software.
- You will submit the post-study questionnaire and comment.

Your participation in this research is entirely your choice. You do not have to answer questions you do not want to answer (by selecting prefer not to answer), and you are welcome to stop the survey at any time if you no longer want to participate. All you need to do is close your browser or browser window. I will not include any incomplete surveys in my analyses. If you do complete your survey and you change your mind later, I will not be able to remove the information you provided as I will not know which response is yours. After accepting this consent, you will have 1 week to perform the study, after which time you will be considered as no longer wanting to participate in it.

Your responses to the survey will be anonymous. This means that there are no questions in the survey that ask for identifying details such as your name or email address. All responses will be saved on a secure Dalhousie computer. Only Leonardo Christino, Taylor Hill, Prof. Fernando Paulovich and Prof. Evangelos Milios will have access to the survey results.

I will describe and share general findings of this research in a journal and/or conference publication, including aggregate/statistical data from this study. All data with answers collected from participants will be deleted 5 years after reporting the results, where only fully anonymous aggregated data and text excerpts will be available through the published material.

The risks associated with this study are no greater than those you encounter in your everyday life.

Your contact information will not be linked in any way to your survey responses.

You should discuss any questions you have about this study with Leonardo Christino or Prof. Evangelos Milios. Please ask as many questions as you like before or after participating. My contact information is christinoleo@dal.ca.

If you have any ethical concerns about your participation in this research, you may contact Research Ethics, Dalhousie University at (902) 494-3423, or email ethics@dal.ca (and reference REB file # 20XX-XXXX)."

If you agree to complete the survey, please answer this email with "I accept the consent agreement", and the link to the survey will be sent to you.

Appendix C–Demographic Questionnaire

1   How long have you used a computer?
   - Less than one week
   - 1 week to less than 1 month
   - 1 month to less than 1 year
   - 1 year to less than 2 years
   - 2 years to less than 4 years
   - 4 or more years
   - I prefer not to answer
2   On average, how much time do you spend per week on a computer?
   - Less than one hour
   - One to less than 4 hours
   - to less than 10 hours
   - 10 to less than 20 hours
   - 20 to less than 40 hours
   - Over 40 hours
   - I prefer not to answer
3   How comfortable are you at using interactive user interface?
   - Extremely comfortable
   - Very comfortable
   - Comfortable
   - Uncomfortable
   - Very uncomfortable
   - Extremely uncomfortable
   - I prefer not to answer
4   How familiar are you with Data Analytics Tools, such as Microsoft Excel or Tableau?
   - Very well
   - Well
   - Neutral
   - Not well
   - Not well at all
   - I prefer not to answer
5   At what level do you think your understanding of written English is?
   - Excellent
   - Very good
   - Good
   - Acceptable
   - Bad
   - Very bad
   - None
   - I prefer not to answer
6   What is the highest level of education you have completed?
   - Little or no formal education
   - High school or equivalent
   - College or university
   - Master
   - Doctoral
   - Post-Doctoral
   - I prefer not to answer

7   In case you have or are pursuing a degree, what is your primary area of study?
- Computer Science
- Information technology
- Internetworking
- Social Science
- Health Science
- Other_____
- I have no primary area of study
- I prefer not to answer

Appendix D –Questionnaires

Prelude - Video tutorial

Please view this video tutorial (link) only once and respond to the following statements about the visualization-based interface, using the given scale:

Questionnaire 1.1 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Which FSA shows the most concern with access to educational opportunities?

Question	Answers				
I was able to follow the steps without any problems	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I was able to quickly understand what I needed to do to perform the required steps in the webapp	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe I was able to answer the requested question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

Questionnaire 1.2 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Is there a "happiest" community that shows the highest life satisfaction score?

Question	Answers				
I was able to follow the steps without any problems	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

I was able to quickly understand what I needed to do to perform the required steps in the webapp	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe I was able to answer the requested question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

Questionnaire 1.3 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Where do parents experience barriers to access recreation in terms of facilities not offering childcare?

Question	Answers				
I was able to follow the steps without any problems	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I was able to quickly understand what I needed to do to perform the required steps in the webapp	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe I was able to answer the requested question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

Questionnaire 1.4 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Where do people tend to buy local food? (or, where is the lowest amount of those who "never" buy local food?)

Question	Answers				
I was able to follow the steps without any problems	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I was able to quickly understand what I needed to do to perform the required steps in the webapp	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
While performing the steps, I totally ignored other information not relevant to the question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe I was able to answer the requested question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 1.5 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Where do people report the highest overall work life balance?

Question	Answers				
I was able to follow the steps without any problems	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I was able to quickly understand what I needed to do to perform the required steps in the webapp	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
While performing the steps, I totally ignored other information not relevant to the question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe I was able to answer the requested question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 2 – Quality of Life Tool's Interface Features Questionnaire

Please respond to the following statements about the visualization-based interface, using the given scale:

Question	Answers				
The tool was easy to use	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

The questions were simple to answer	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I answered the questions confidently	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
It was intuitive to write down my **intentions** during the process	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
It was intuitive to write down my **insights** during the process	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe if I were to see other's **intentions**, I would be able to better understand the question.	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe if I were to see others **insights**, I would be able to answer the question faster	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe if I were to see others **insights**, I would be able to answer the question with more precision	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 3.1 - Pre-defined execution Questionnaire of the Ocean Data Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: Considering the cardinal directions (north, northeast, east, etc), where are most of the sensors (also called stations) located within the Bay of Fundy?

Question	Answers				
I was able to follow the steps without any problems	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I was able to quickly understand what I needed to do to perform the required steps in the webapp	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
While performing the steps, I totally ignored other information not relevant to the question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe I was able to answer the requested question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 3.2 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: In which year and month were fish most detected on the passage?

Question	Answers				
I was able to follow the steps without any problems	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I was able to quickly understand what I needed to do to perform the required steps in the webapp	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe I was able to answer the requested question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

Questionnaire 3.3 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: How much did the tide affect the presence and number of sharks in the avon river in 2019?

Question	Answers				
I was able to follow the steps without any problems	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I was able to quickly understand what I needed to do to perform the required steps in the webapp	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

I believe I was able to answer the requested question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 3.4 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: On what day of the week are sharks most and least active in 2018?

Question	Answers				
I was able to follow the steps without any problems	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I was able to quickly understand what I needed to do to perform the required steps in the webapp	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
While performing the steps, I totally ignored other information not relevant to the question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I believe I was able to answer the requested question	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

Questionnaire 3.5 - Pre-defined execution Questionnaire of the Quality of Life Tool

Now open the website at this link and attempt to answer the following question. As the video has shown, write the individual intentions and/or insights you discover in the process. If you find the answer, type as a final insight and then respond to the following statements about your experience using the given scale. Note that the process of finding the answer is what is important to us, so don't worry about how well you answer the question, just attempt to within the allotted time of 5min.

Question: How does the number of detected striped bass compare between the sensors running across the north shore of the main basin, the southern shore sensors and the avon river sensors?

Question	Answers				
I was able to follow the steps without any problems	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*
I was able to quickly understand what I needed to do to perform the required steps in the webapp	*Strongly Disagree*	*Somewhat Disagree*	*Neutral*	*Somewhat Agree*	*Strongly Agree*

While performing the steps, I totally ignored other information not relevant to the question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe I was able to answer the requested question	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

Questionnaire 4 – Ocean Tool's Interface Features Questionnaire

Please respond to the following statements about the visualization-based interface, using the given scale:

Question	Answers				
The tool was easy to use	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
The questions were simple to answer	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I answered the questions confidently	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
It was intuitive to write down my **intentions** during the process	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
It was intuitive to write down my **insights** during the process	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe if I were to see other's **intentions**, I would be able to better understand the question.	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe if I were to see others **insights**, I would be able to answer the question faster	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
I believe if I were to see others **insights**, I would be able to answer the question with more precision	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree

1. Please give us more comments about your experience:

_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____

2. Is there any way you expect that the intentions and insights could be used to empower you while you were answering the questions?
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____
_____