Transcript:

Slide:
Project Title:

Hello and thank you for joining me for this research proposal. My name is Simon Miller, I am a student of the University of Essex, Computing Department currently taking part in a two years masters program. Today I will be discussing what I feel offers a unique opportunity to study the regional impacts of cyberbullying on communities across Canada. Without further adieu, let's dive into my presentation titled:

Regional Demographic Influences Motivating Cyberbullying in Urban Centers Across Canada

Before we look at what we'd like to do, let me briefly give an overview of the current trends in research. The current body of literature regarding cyberbullying often looks at the phenomenon as if it occurs in a bubble. This is often required to limit the scope of any research paper, or to better define their own set of research questions. The result, however, is that a large amount of current research fails to define motivations or influences that could alter the nature of cyberbullying itself outside of a single place, be it a country as a whole, or by reviewing datasets that don't geographically distinguish where the data came from.

Currently, much of the research coming out of academic institutions focuses on applying an array of machine learning models to the detection of cyberbullying. The drawbacks to these studies are that they are often limited in scope; limited by dataset availability as well as limitations in the ability to effectively codify context awareness into a machine learning algorithm. An example which might be defining the incapability of distinguishing malintent or levity.

These models can also be limited in that they pick one small facet of cyberbullying as an indicator, such as bad language, something that can be a poor "true" reflection of the existence of cyberbullying.

I feel the emphasis on algorithmic detection, while potentially effective at detecting the presence of cyberbullying, does nothing to outline the motivations behind cyberbullying. Machine learning doesn't tell us if events of cyberbullying are racially, sexually, politcally or professionally motivated. Now I know these factors are out of the scope for these types of research projects as it is not the intent of AI cyberbullying detection to identify the motivations or the "why" of cyberbullying.

Furthermore, non-machine learning oriented research does little to address regional influences on cyberbullying, instead addressing the questions of awareness, impacts and exposure.
I believe there is a gap in the current body of literature that fails to address whether regionality plays a factor in the motivations that impact cyberbullying victims as a whole. What factors

motivate cyberbullying and how do these factors change based on region? Do cyberbullying victims in Vancouver, British Columbia experience the same type of online discrimination as those in St. Johns, Newfoundland?

Key Literature

In preparation for this research proposal, a literature review was carried out with the objective of assessing a range of relevant sources on the field's current body of research into cyberbullying and more specifically what is being done for detection and prevention as well as what populations this research is being applied to.

One theme that came out of this review was that cyberbullying as a phenomenon is not a singular problem across the global community, but instead a phenomenon that is unique to distinct cultures, political climates and anthropogenic currents as countries have different web-use habits that change the threat landscape from region to region.

Slide:

Throughout this review, 21 journal articles and conference papers were analyzed, spanning a range of countries from Ecuador, to China, Korea or Japan. Topics emphasized the detection and prevention of cyberbullying through the use of machine learning algorithms, as well as less technically oriented papers that analyzed the challenges of addressing cyberbullying in countries that don't use alphabetic languages or had differing cultural understandings of digital abuse.

Slide:

RQ

While my literature review held two research questions to be answered throughout analysis of the literature, I feel a new question needs to be answered regarding the aforementioned motivations that affect the nature of cyberbullying.

The research question we aim to answer is:

What effect do culturally distinct Canadian locales have on how individuals perceived motivations of cyberbullying?

Slide:

Methodology:

So how do we go about identifying regional motivations to cyberbullying? While I believe there could be further work under an expanded scope; comparing the perceptions of cyberbullying across national borders or even continents far exceeds the capabilities or timeframe of this project and would likely require collaboration with institutions in candidate countries.

Let's begin by defining which locales throughout Canada are of interest to us:

These locations represent geographically unique areas with dense population centers such as:

- Vancouver in British Columbia
- Calgary and Edmonton in Alberta
- Regina in Saskatchewan
- Winnipeg in Manitoba
- Kitchener/Waterloo and Ottawa in Ontario
- Montreal in Quebec
- St. Johns in Newfoundland
- Halifax in Nova Scotia

Aside from population density, these cities are culturally distinct

Slide

Our chosen method of data collection is through the distribution of a survey. The survey's design an important step that can make or break the level of engagement. Too long and the response rate will be poor. If the survey is too short the response rate may be higher, but may not garner enough useful data to answer our research question.
Our survey will have no more than 10 questions and will not collect any personally identifiable information including names, addresses or phone numbers or email addresses. The survey is prefaced with a definition of cyberbullying along with a description of the study and its objectives.
Survey data collected is to be predominantly if not entirely qualitative with numeric description from multiple choice selections. This means that “yes and no” responses will be represented as integers as well as any locational or frequency descriptors. Qualtrics will be used to design the survey and distribution will be through paid social media targeting as services such as facebook and instagram offer enormous exposure potential as well as effective demographic targeting that allows us to define appropriate selection and exclusion criteria such as age and region.

Example questions would include:

- Where is your permanent residence?
    - With a list of target cities
    - Or none of the above
        - This selection marks exclusion criteria that may terminate followup questions
- Have you been a victim of cyberbullying within the last year?
    - If “Yes” is selected the survey provides followup questions

Multiple selections can be made on some questions such as:

- Would you say your experiences were motivated by:
    - Sexuality
    - Personal
    - Race

    - Spiritual belief
    - Politics
- Were experiences isolated or repeated?
    - This is a yes no question

Analysis of Data:
From the data collected from the survey, demographic analysis will provide the most insight, given the type of boolean data collected from the survey.

Analysis of data will start by exporting survey results to csv or xlsx format for analysis in Excel. This research project won't require specialized statistics software, such as Stata, SPSS or R. As mentioned before, because the data from respondents is largely formatted as boolean or "choose all that apply" format questions, analysis is largely frequency based. "What percentage of respondents in Calgary faced politically motivated online discrimination? What was the average age of respondents who experienced religious, or harassment based on sexuality?"

While it would be interesting and valuable to test for correlation across datasets, this may not garner accurate results as the correlation tests are best performed on data easily represented on scatter plots. Since survey results are largely represented as boolean figures we can't anticipate data that lends itself to correlation tests via Pearson Correlation Coefficients or Spearman's rank coefficients (Kent State University, 2018).

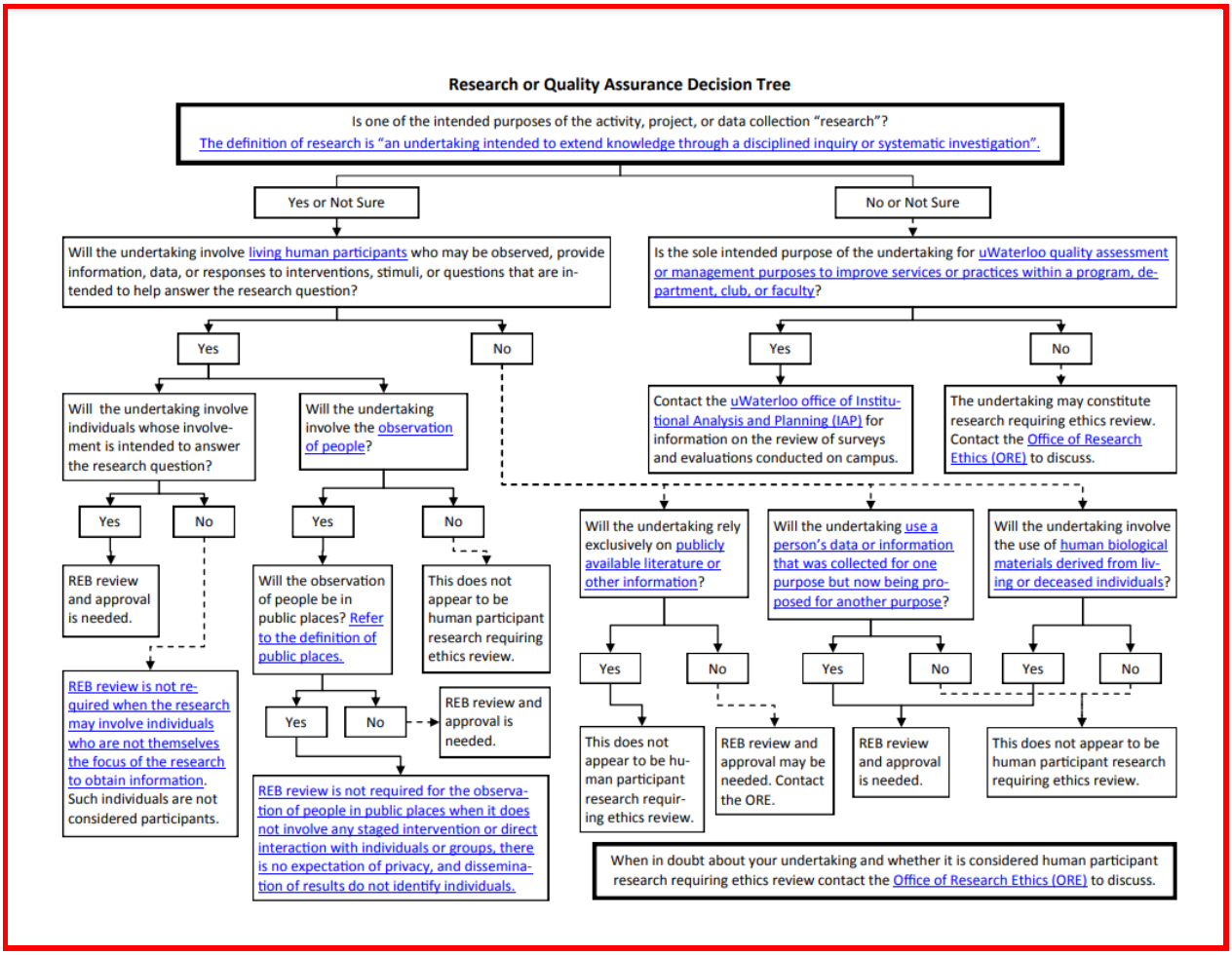
Slide
This data can also be clearly represented graphically and the use of bar and pie charts will be the mainstay of illustrating findings throughout the report.

Key points this research project aims to collect and analyze is the relationship between, age, sex and race relative to the types of cyberbullying experienced. This analysis will be repeated across each of the cities where data was collected. From these findings we can report on the rates of different types of cyberbullying and compare them to the results of different locations.

Risks:
The risks of this project are minimal. Because no personally identifiable information is collected, the responses are kept entirely confidential and anonymous; no party, including researchers, have access to participant phone numbers, email addresses, residential locations or names because that information is not collected. Participation is completely voluntary.

Slide
That being said, in Canada, under TCPS2, Chapter 2, article 2.1, those individuals submitting information as the subject of a study are considered participants of research, thus an ethics committee review may be required (anon, 2014).

To confirm if review by an ethics board would be required, we used the University of Waterloo decision tree as an assessment tool. It was determined that a review by the research and ethics board would be required, as human participants are used in the collection of data to answer our research question.



Artifacts and Outcomes:

The outcome of this research project won't produce any created or developed software artifacts as the objective isn't to create a piece of software or an algorithm. There will however be artifacts in the form of a survey as well as the data collected via the survey. Furthermore, there are the spreadsheets where statistical analysis and any graphed resources sit

What I hope is the outcome is a report that shows regional trends in cyberbullying. Places such as Calgary and Edmonton may experience higher rates of politically influenced cyberbullying. Vancouver is known for its racial diversity, but this may also contribute to higher incidents of racially motivated cyberbullying. Newfoundland may make a great location "control" group as they're generally universally beloved by all.

Timeline:

The total timeframe this project is expected to take, start to finish is approximately eight months to one year to complete.
The rationale behind this largely sits with the time it will take to collect enough relevant data to be actionable. I anticipate the following timeline:
Month 1: This first month encompasses the time up till now to complete the research proposal before next steps are taken
Month 2: Perform an expanded literature review that may provide further insight into the demographics surrounding cyberbullying
Month 3: Planning and methodology are established. Here we will develop our survey questionnaire in Qualtrics as well as develop a comprehensive plan to distribute the survey via sponsored social media with targeted demographics.
Month 4 and 5: Is where we'll begin running our surveys and monitoring data collection. Depending on the response rate, preliminary analysis can be done on results pulled before the conclusion of the data collection phase. We can begin to get an understanding on the shape and quality of the data as it comes in.
Month 6: Will be spent analyzing the data in earnest to be used to write the reports section of the report
Month 7: Will focus on completing the analysis and methodology sections of the final submission
Month 8 and 9: Is all editing and revision and finishing any supportive writing not under the results and methodologies sections
Month 10: Is the final proofing prior to submission

Conclusion:
Thank you for sticking with me till the end of this proposal. Hopefully, you see the same value in performing such a study as I believe one such as this hasn't been performed yet, at least not for populations within Canada and I believe we could find real insight into whether there is a correlation between where someone lives and what might influence the type of cyberbullying experienced.
If results are promising, research could be expanded to include a broader scope of demographic data to compare the differences of not just cities in one country, but perhaps even differences in cyberbullying motivations across borders and cultures.

Kent State University (2018). *LibGuides: SPSS Tutorials: Pearson Correlation*. Kent.edu. Available at: https://libguides.library.kent.edu/SPSS/PearsonCorr.

University of Waterloo (2014). *Does my data collection activity require ethics review?* uwaterloo.ca. Available at: https://uwaterloo.ca/research/office-research-ethics/research-human-participants/pre-submission-and-training/human-research-guidelines-and-policies-alphabetical-list/does-my-data-collection-activity-require-ethics-review [Accessed 25 Mar. 2022].