Check for updates

Semi-Automated Approach for Evaluating Severe Weather Risk Communication

Sophia Jit
University of Toronto
sophia.jit@mail.utoronto.ca

Jennifer Spinney
York University
jspinney@yorku.ca

Priyank Chandra
University of Toronto
priyank.chandra@utoronto.ca

Robert Soden
University of Toronto
soden@cs.utoronto.edu

ABSTRACT

Communicating risk to the public in the lead-up to and during severe weather events has the potential to reduce losses from such events. Globally, severe weather events are anticipated to increase due to climate change rendering effective risk communication an integral component of climate adaptation policies. The identification of best practices in risk communication has long been studied by risk communication professionals. However, few articles have attempted to quantify the compliance of severe weather risk messages with these best practices at scale or developed tools that can be used by this population. The current work makes two contributions toward this goal. First, we use string-matching to evaluate the degree to which risk communication issued in Canada complies with best practices and suggest ways to improve this messaging. Second, we suggest a writing support tool to be used by risk communication professionals to evaluate risk messages.

CCS CONCEPTS

• **Human-Computer Interaction**; • **Empirical Studies in HCI**; • **Information Systems**;

KEYWORDS

Severe Weather Risk Communication, String-Matching, Writing Support Tool

ACM Reference Format:
Sophia Jit, Jennifer Spinney, Priyank Chandra, and Robert Soden. 2023. Semi-Automated Approach for Evaluating Severe Weather Risk Communication. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3585753

1 INTRODUCTION

Weather risk communication is an interdisciplinary field of research and practice that concerns the communication of accurate and actionable information about health, safety, and risk management strategies during the course of weather-related hazardous events [58, 67]. Weather risk messages may be issued before, during, and after such events occur and may thus span both the short- and long-term. Short-term messaging concerns the communication of immediate threats while long-term messaging is intended to encourage preparation and planning in advance. Regardless of the timeframe, both are integral to public safety and may potentially save lives [37]. The creation of effective risk messaging remains a challenging task as risk communication professionals must often contend with varying degrees of certainty and competing priorities in a high-stress situation [36, 52]. This study developed a semi-automated approach for evaluating the adoption of best practices in risk communication. Our approach enables risk communication researchers and practitioners to identify opportunities for improving the effectiveness of risk communication and informs the design of writing support tools for risk communication professionals.

Climate change is predicted to increase the frequency and duration of severe weather events across the globe [29, 54]. A warmer planet will likely result in more extreme hot days and heat waves, which are predicted to increase in length, frequency, and intensity [8]. Many areas in Asia, Australia, and Europe have already begun to experience these phenomena [8], with the latter continent suffering over 20,000 excess deaths due to a summer heat wave in 2022 [60]. Regions such as southern Europe and the Mediterranean, central Europe, and central North America will experience rising temperatures that will increase evapotranspiration and consequently drought events [8]. Increased temperatures are also predicted to increase atmospheric water vapor, resulting in more heavy precipitation in both North America and Europe. The 2021 European floods are evidence of this claim, with countries such as Belgium, Germany, Luxembourg, and the Netherlands, experiencing severe flooding. Germany experienced the highest casualty rate among the European countries, with at least 184 Germans reported dead and an estimated € 30 billion in damages [5]. In response to this flood, the United Nations University Climate Resilience Initiative noted hazard-related gaps in the flow of information and risk perception in the population, which may benefit from technological and digital innovation [2].

The evaluation of risk communication at scale and the development of writing support tools for risk communication professionals offers opportunities for HCI and crisis informatics researchers to contribute to risk communication and practice. To date, research in these disciplines has largely focused on the exchange of risk information between the public and official agencies [44, 70, 72], or online risk communication during public health emergencies [34, 35, 58]. The development of quantitative approaches to risk

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585753

communication remains a nascent area of research. Meanwhile, much of the literature that has looked to develop writing support tools in non-fiction settings has focused on sentence completion [17] and prediction in online communication [17, 39]. Recent work in idea generation [32, 55] offers particular insights into the development of a writing support tool for risk communication professionals as the inclusion of particular ideas or themes is integral to preventing losses.

This paper uses string-matching [16, 65], a methodology that has been used in fields such as computational biology [4, 18, 69] and text retrieval [3, 38], to semi-automatically evaluate structured risk communication at scale issued by Environment and Climate Change Canada (ECCC). Canada was chosen as the subject of the analysis as it is currently warming at twice the global rate and is expected to experience significant losses, damages, and disruptions due to climate change over a 20-year period [1]. Thematic analysis – a methodology employed in much of the HCI literature to analyze qualitative data [19] – acts as a starting prompt for much of this work. We first use this methodology to inductively and deductively identify key themes in a subsample of risk messages issued by ECCC. To evaluate the entire corpus, we calculate the readability score of the messages and use our string-matching algorithm to semi-automatically label the entire sample. In doing so, we are able to overcome the small sample sizes that often characterize data in this literature and evaluate the corpus' compliance with best practices in risk communication. In addition to informing risk communication policy, the methodology supports the creation of a writing support tool that can notify risk communication professionals of themes that are omitted from a given message in real time.

2 RELATED WORK

2.1 HCI and Crisis-Informatics

Risk communication has been a growing subject of interest to both HCI and crisis informatics researchers. Much of the work in the latter discipline has examined the exchange of risk information between authoritative bodies and the public to understand the role of ICTs in aiding contextual awareness between these groups [44, 70, 72]. For example, Bica et al. [11] examine the diffusion, prevalence, and public reception of hurricane risk images from the 2017 Atlantic hurricane season shared by authorities on social media. Recent work in this field has similarly turned its attention to informal risk communication, including how the public uses social media to communicate and understand localized risk [53] and how volunteer technology communities organize to support such communication [64, 68]. Work in the former field has largely focused on online risk messaging during public health emergencies, such as the COVID-19 pandemic [58] and the outbreak of the Zika virus [34, 35]. Severe weather risk communication (SWRC) specifically has been the focus of recent design-related HCI scholarship. Researchers have outlined a series of design opportunities for HCI practitioners and designers to support digital risk communication and used the public's behavior during storms to help inform the design of tools to better meet their needs [67].

2.2 Risk Communication

Creating effective risk communication has been the focus of much of the emergency management literature. This field of research largely employs qualitative and quantitative methods to understand how individuals make risk-related decisions during hazardous events, and how to communicate complex information to the public. This literature has found that trust in organizations, governmental agencies, or spokespeople is a major determinant of risk reduction and protective actions during weather-related hazards [49, 71]. These trusted authorities should work with a diverse group of institutions, scientists, and public officials to communicate accurate information with a unified message or voice [25]. Although most risk communication during hazards is often unidirectional – going from authorities to the public – a growing body of work has found that two-way dialogue between creators and recipients of risk communication is more effective during crisis situations and disaster mitigation efforts [7, 24]. Furthermore, educational materials and risk messages should be targeted to a specific audience and use multiple communication channels [10, 63, 66]. Successful weather risk communication should thus be designed to meet the needs, vulnerabilities, and cultural beliefs of the intended audience [41, 48, 51, 66].

Identifying the elements of effective risk communication has been of particular interest to emergency management researchers. Research in this domain has found that effective risk messaging is often consistent, clear, and concise and does not rely on fear-based language [23, 46, 50]. Specificity is a key determinant of protective actions as the audience is more likely to both believe a threat is credible and personalize the hazard risk [26]. As a general guideline, risk communication should thus include discussion of hazards, location, timing, vulnerability, and recommended actions [21, 22, 30, 45–47, 57]. The inclusion of hazards and its associated impacts may help inform the logic of protective actions, while hazard-related risk is determined by its location and the time that the event will strike – information which allows individuals to respond accordingly [46, 47]. Similarly, an understanding of one's vulnerability and recommended actions are integral to the adoption of protective behaviors [21, 22, 57]. Stylistically, discussion of the hazard and safety-related actions should be delivered with a high degree of certainty in plain language that is understandable to the public [56].

The current work expands upon previous literature by using string-matching to quantify the corpus' compliance with best practices in risk communication. The proposed string-matching algorithm is predicated on inductive and deductive thematic analysis, where the previously identified themes related to specificity – hazards, location, timing, vulnerability, and recommended actions – informed the latter. These themes were chosen by virtue of their role in informing risk perception and motivating protective actions during severe weather events [12, 13, 22, 46]. To evaluate the plainness and comprehensibility of the language, reading level metrics for the corpus were also computed.

2.3 Writing Assistant Tools

With the development of improved and advanced language models, the academic literature on writing support tools has experienced

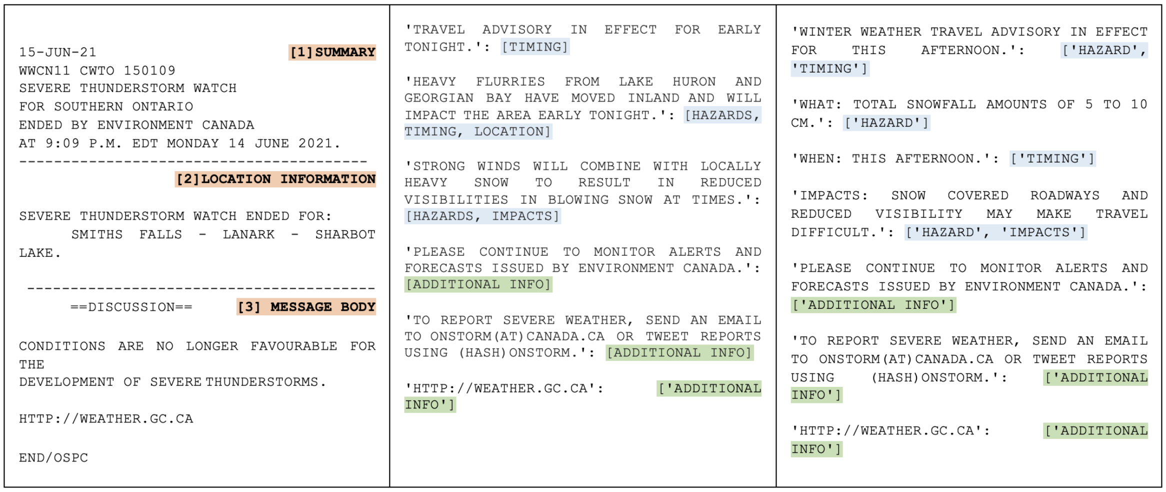

Figure 1: ECCC Severe Weather Risk Messages. From left to right, an example of: (a) an ECCC risk message, (b) a message labelled using thematic analysis, and (c) a message labelled using semi-automatic message classification. Text in green denotes a theme – "Additional Information" – that was omitted from the primary analysis.

a resurgence in recent years, with tools to support both creative and scientific writing. Much of the work in the former focused on narrow support for writing tasks, such as storytelling [62] and metaphor writing [15]. More varied and diverse assistance to writers, such as assistance with descriptions, plot points, or providing questions [6, 20] has been the subject of recent work in the field. Scholarship in the domain of scientific writing has thus far been more limited, with much of this work focusing on sentence completion [17] or smart replies in emails [39]. As opposed to providing complete phrases or sentences, Peng et al. [55] develop a tool to assess and provide suggestions to posts created for online mental health communities. Similarly, Gero et al. [32] follow up on their nascent research on online scientific communication on Twitter [33] and use pre-trained language models to generate suggestions for science-related text. The present study builds upon this literature on idea generation to suggest a writing support tool for risk communication professionals. In lieu of the machine learning (ML) methods employed in previous studies, this research employs string-matching to identify the absence of themes related to best practices in risk communication, which can be used to both analyze risk communication at scale and suggest the inclusion of omitted themes.

3 METHODOLOGY

3.1 Data

The current work employs 9,181 public weather alerts issued by ECCC – the branch of the Canadian federal government responsible for coordinating environmental policies and programs – for Ontario between June 2020 and December 2021. In the data, weather alerts span four categories – special weather statements, advisories, watches, and warnings – and 23 hazards of varying intensity (see Figure 2). The structure of weather alerts is comprised of three sections as indicated in Figure 1 :(1) summary, (2) location information, and (3) message body. The summary provides an overview of the message, including the type of weather event, the affected locations, and the date and time the message was issued. The location section summarizes all the affected locations, and the message body contains the main risk message. Multiple alerts may have been issued for a single weather event, where the maximum number of messages associated with an event in the sample is five. The current work focuses on an analysis of the message bodies. Although the data set is composed of weather alerts issued by the federal government to the public, the compiled corpus was purchased by the research team and thus cannot be shared publicly.

3.2 Reading Level Metrics

To understand the reading level of the messages, the Flesch Reading Ease (FRE) score – one of the most common methods to assess text readability [28, 31] – was computed for the message bodies of the sample. Developed in 1948 by Rudolph Flesch, the FRE score is a metric to indicate the difficulty level of a text based on a weighing of the ratio of total words to total sentences and the total syllables to total words. The resulting score ranges from 0 to 100, where higher scores indicate that the text is relatively easy to read for the average adult, resulting from either a lower number of words or syllables (see Table 3). For his work, Flesch was lauded as the "Apostle of 'Plain Talk"' [43] where the author himself set a score of 60 as the minimum for plain English [61].

To calculate the FRE score using the TextStat library, basic text pre-processing was undertaken such as replacing all website names and social media accounts with generic word identifiers such as "website URL" and "email account". Like most readability formulas, the FRE score does not evaluate features of text such as content, organization, word order, format, or imagery [61]. Such metrics primarily pertain to style, although they are unable to assess elements of writing such as mood, tone, and persuasiveness [61]. An assessment of linguistic features such as consistency, clarity, and conciseness may thus require ML-based methods for automatic text evaluation [9, 27, 42, 59] which are beyond the methodological scope of the present work.

3.3 Manual Thematic Analysis

To assess the messages' compliance with best practices in risk communication, the research team first engaged in deductive and inductive thematic analysis [19]. Given the size of the dataset (N=9,181), each member of the research team reviewed ten message bodies for each of the five most prevalent weather events to inductively find repeated patterns and themes. This process helped identify a preliminary list of qualitative codes, which the team then updated and deductively refined through discussion of similarities and differences between the patterns in the message bodies and best practices in risk communication. This deductive and inductive process was repeated iteratively until a final list of themes, was agreed upon: hazards, location, timing, affected populations[1], impacts, and recommended actions. The inductive method generated two additional themes – "impacts" and "additional information" – relative to the deductive approach although the latter is omitted from the remainder of the analysis as it is not a theme associated with best practices. Table 4 provides a summary of the identified themes and their corresponding interpretation.

3.4 Semi-Automatic Message Classification

To systematically analyze the samples' compliance with best practices in risk communication, the research team used the labels identified in 3.3 and keywords to semi-automatically label the data set. For each of the five most prevalent weather events, a set of key words and phrases associated with the identified themes were selected to label the data [36, 52]. Given the general uniformity of the messages within weather events and the frequent recycling of text, keywords and phrases were chosen by virtue of their repeated appearance in weather alerts and their ability to uniquely identify a theme. The presence of a given keyword or phrase within a sentence was used to label a sentence with that theme, where sentences could have multiple labels. Keywords or phrases that uniquely identified a given theme in one context, but whose meaning varied given the context of another theme, was used to create a list of exclusion criteria. For example, the words "north" and "northerly" were used as inclusion and exclusion criteria respectively to identify discussion of locations. To develop a full list of keywords and phrases, a similar iterative process as in 3.3 was followed where the preliminary list of keywords and phrases was used to randomly label a subset of the data. The keywords and phrases were then iteratively updated based on mislabeled data, until no further errors were found. Figure 1 includes an example of results from both the manual thematic analysis and the semi-automatic message classification.

4 RESULTS

4.1 Summary Statistics

Table 1 provides a summary of the data set, where the hazards are organized by the event category they belong to. The percent column notes the total percent of the entire sample that comprises each category and the total percent of the category that comprises a given hazard. As indicated in the table, of the total sample, statements, advisories, watches, and warnings amount to 17.04%, 19.26%, 9.51%, and 54.2% of the sample respectively. Across all weather events, five events – severe thunderstorm warnings (23.3%), special weather statements (17.04%), weather advisories (11.53%), heat warnings (7.87%), and snow squall warnings (6.42%) – make up 66.18% of the total sample (see Figure 2).

Table 1 also displays summary statistics for the word count of the message bodies. Across the message types, the word count of the messages ranged from one word to a maximum of 372 words (Tornado Watch). Messages that only included a single word reflected a message body that simply included a URL for additional information. Of the total sample, 2.97% contain only URLs of which 2.86% are the first message to be issued in the sample, and the remaining 0.11% are the second message to be issued. For the analysis, this data is dropped from the sample. Although the content and structure of the message bodies varied across weather events, message bodies within a given weather event demonstrated significant similarity such that many messages repeated the same text with event-specific details. The uniformity of messages within weather events is thus conducive to the use of string-matching [14, 16, 65] to analyze the data.

4.2 Reading Level Results

The final columns of Table 1 provide summary statistics for the FRE score for each event category and hazard. Across the four categories, the average FRE score ranges between 51.89 (Watch) and 67.1 (Statement) suggesting that the difficulty level across the categories ranges from fairly difficult to standard difficulty. The minimum FRE score across the four categories was difficult (30 < FRE < 49), while the maximum was either easy (80 < FRE < 89) or very easy (90 < FRE < 100) where the higher scores reflected a lower ratio of total words to total sentences, a lower ratio of total syllables to total words or both. For the five most prevalent weather events, the range of average FRE scores was narrower than that of the four event categories and ranged from 58.23 (Severe Thunderstorm Warning) to 67.1 (Special Weather Statement); the difficulty level ranged from fairly difficult to standard difficulty. The range of FRE scores was larger compared to the four categories, with minimum scores that were difficult (30 < FRE < 49) and maximum scores that were either fairly easy (70 < FRE < 79), easy (80 < FRE < 89) or very easy (90 < FRE < 100). The results suggest that, for both event categories and hazards, the average reading score of the message body was difficult to varying degrees, although there was significant variation within the given unit of analysis.

4.3 Semi-Automatic Message Classification Results

For the five most prevalent weather events, Table 2 shows the percent of messages that include discussion of the themes identified in 3.3. For this analysis, weather alerts that simply noted an event had come to an end were dropped from the sample. This decision was motivated by the assumption that these messages no longer required the audience to engage in protective actions and thus did

[1]The theme "affected populations" corresponds to the deductively identified theme of "vulnerability". The research team chose to rename the theme to more accurately reflect the specificity in the corpus.

Table 1: Summary Statistics for Risk Messages Issued by ECCC

	Word Count				Flesch Reading Ease Score [a,b]		
Weather Event	Percent	Mean (Std)	Min	Max	Mean (Std)	Min	Max
Statement	17.04	93.7 (52.2)	1	273	67.1 (8.44)	31.55	103.63
Special Weather	100	93.7 (52.2)	1	273	67.1 (8.44)	31.55	103.63
Advisory	19.26	88.65 (49.15)	1	231	59.72 (8.95)	37.47	95.17
Blowing Snow	1.98	90.11 (42.89)	12	166	56.96 (9.00)	41.36	77.33
Freezing drizzle	3.9	88.45 (42.61)	1	167	65.11 (5.45)	54.83	77.91
Frost	16.18	71.08 (31.38)	7	161	62.19 (6.75)	47.79	81.29
Fog	18.04	69.33 (34.23)	1	154	54.7 (9.52)	37.47	89.75
Weather	59.9	99.18 (54.17)	1	231	60.31 (8.82)	43.9	95.17
Watch	9.51	131.06 (63.23)	1	253	51.89 (9.05)	36.79	81.63
Winter storm	0.46	171 (8.12)	166	183	67.13 (7.58)	57.06	72.97
Tornado	4.12	183.69 (72.14)	11	253	50.58 (5.97)	41.26	59.3
Snow Squall	29.55	130.45 (50.86)	1	247	62.02 (6.15)	44.64	81.63
Severe Thunderstorm	65.86	127.76 (66.35)	1	242	47.32 (6.02)	36.79	77.91
Warning	54.2	143.69 (68.39)	1	372	60.86 (10.5)	36.96	100.24
Blizzard	0.24	152.58 (22.77)	129	191	68.31 (6.22)	57.67	76.62
Extreme Cold	0.68	81.76 (41.93)	12	132	67.32 (10.7)	49.82	77.53
Winter Storm	1.91	117.91 (52.76)	8	204	60.63 (5.02)	49.31	72.97
Wind	3.64	86.77 (39.9)	11	150	73.9 (4.75)	58.79	80.28
Tornado	4.18	213.49 (76.71)	5	372	56.69 (7.66)	43.19	99.23
Freezing Rain	5.47	101.91 (48.83)	1	193	64.04 (6.61)	45.93	77.91
Snowfall	5.83	115.87 (49.23)	1	239	61.39 (9.62)	37.98	81.33
Rainfall	7.6	125.33 (63.05)	7	270	60.71 (10.56)	37.98	88.74
Snow Squall	11.84	124.33 (54.71)	8	239	62.29 (5.16)	41.87	83.96
Heat	14.53	166.63 (57.02)	11	347	63.89 (6.08)	47.79	77.13
Severe Thunderstorm	44.09	153.46 (71)	1	287	58.23 (12.47)	36.96	100.24

[a] The interpretation of the Flesch Reading Ease score is as follows: 90-100: Very Easy; 80-89: Easy; 70-79: Fairly Easy; 60-69: Standard; 50-59: Fairly Difficult; 30-49: Difficult; 0-29: Very Confusing.

[b] A feature of the TextStat library is that the computed Flesch Reading Ease scores have an upper bound of 121.22 and no lower bound where values above 100 and below zero are interpreted as "Very Easy" and "Very Confusing" respectively.

Table 2: Prevalence of Themes Related to Best Practices in Risk Communication in ECCC Risk Messages

Event	Timing Information	Affected Population	Recommended Actions	Hazards	Locations	Impacts
Snow Squall Warning	96.46%	23.62%	79.72%	99.41%	78.15%	98.82%
Special Weather Statement	94.88%	14.93%	27.82%	96.49%	52.05%	55.86%
Weather Advisory	85.47%	39.78%	59.42%	86.67%	41.48%	71.74%
Heat Warnings	100.0%	93.19%	100.0%	99.11%	51.7%	88.74%
Severe Thunderstorm Warning	83.33%	62.19%	83.15%	99.77%	76.12%	82.42%

not necessitate the corresponding elements of effective risk communication. In the remaining sample, there is significant variation in the presence of themes related to best practices in risk communication. Across all five weather events, the majority of the sample included discussion of timing and hazards with a range of 83.33% - 100% and 86.67% - 99.77% respectively. Discussion of the remaining themes demonstrated significant variability across the events. The inclusion of affected populations showed the greatest variation, ranging from 14.93% (Special Weather Statement) to 93.19% (Heat Warnings). Locations and impacts were mentioned within a narrower interval of messages relative to affected populations, with the discussion of the latter included in almost 100% of the messages in one event (Snow Squall Warning). Ideally, for weather alerts to comply with best practices, the relevant themes would be present in all the messages across all weather events. These results are hardly born out in the data as no weather event includes all the themes and only a handful of select themes top 99%.

5 DISCUSSION

5.1 Feedback to Improve Risk Communication Policy and Practice Summary Statistics

The results of this analysis offer insights into practical changes that can be made to SWRC to improve their effectiveness. Specifically, current risk messages issued by ECCC may benefit from both simpler language to improve their readability and the inclusion of omitted themes such as recommended actions, impacts, and affected populations. Traditional risk communication literature has found that messages that include the level of risks, impacts, and recommended actions generate the greatest change in people's intentions and behavior [30]. This finding is supported by recent work in hurricane risk communication which found that including impact statements may improve a message's efficacy [67]. Similarly, a sense of vulnerability has been shown to be an important determinant of risk perception [12, 13] with the combination of high vulnerability and high efficacy being integral to the adoption of protective actions [21, 22, 57]. The inclusion of the relevant themes may thus improve the audience's propensity to engage in protective decision making which may reduce the losses caused by severe weather events. Reducing the reading level would also make them more accessible, inclusive, and easier to understand [40]. This research can potentially inform individual risk communicators and risk communication policy by providing detailed guidance on the current state of practice in the field and practical, hazard-specific guidance on avenues for improving these messages.

5.2 Design Recommendations for a Risk Communication Writing Support Tool

The algorithm proposed in this work can be used to design a writing support tool that can identify themes that are lacking in risk communication issued in real time. Prior work has largely relied on ML techniques to generate questions or ideas to prompt writing [32, 55]. For example, Peng et al. [55] use traditional ML to develop a tool that predicts a comment's level of support and provides two types of feedback: (1) an assessment of the text and (2) recommendations of high-support comments as a reference for revision. Similarly, Gero et al. [32] train a language model (GPT-2) to generate new ideas from a prompt to be used as inspiration for new research projects. The proposed writing support tool would similarly take messages and the type of weather event as inputs and identify themes that are absent from the message and suggest their inclusion. Given that sentences are often reused in the text, the tool could also suggest sentences to include based on the most frequently recycled messages in historical risk communication. The benefit of this approach, relative to artificial intelligence-based models, is that it is computationally efficient and easily interpretable. This ease of interpretation renders it more accessible to risk communication professionals and lends itself to open collaboration. Furthermore, natural language processing models can only be employed in contexts where the language of the training and test data is also the language in which risk communication is disseminated. In settings where risk communication is relatively structured and text is recycled, the proposed method can easily be adapted to small and large data sets by updating the relevant keywords in the language the risk is communicated.

6 CONCLUSION

Severe weather events are anticipated to increase in duration and frequency due to climate change [29, 54], rendering effective risk communication an integral component of climate adaptation policies. The present study employed string-matching to evaluate the adoption of best practices in risk communication in current practice. The analysis provides guidance on ways that SWRC can be improved and acts as a starting prompt for the creation of an authoring support tool for risk communication professionals. Future work will consider taking steps, such as expanding the patterns specified in the current analysis to improve the algorithm's performance and extending the analysis to the remaining 18 weather events, to operationalize the suggested writing support tool. Furthermore, an evaluation of linguistic features, such as consistency, clarity, and conciseness, of the corpus would greatly supplement the current analysis and provide additional guidance on improving SWRC.

REFERENCES

[1] Climate change adaptation in Canada. Government of Canada City.
[2] 2021 floods: UN researchers aim to better prepare for climate risks. United Nations Regional Information Centre for Western Europe, City, 2022.
[3] Aho, A. V. and Corasick, M. J. Efficient string matching: an aid to bibliographic search. Communications of the ACM, 18, 6 (1975), 333-340.
[4] Alsmadi, I. and Nuser, M. String matching evaluation methods for DNA comparison. International Journal of Advanced Science and Technology, 47, 1 (2012), 13-32.
[5] Apel, H., Vorogushyn, S. and Merz, B. Brief communication: Impact forecasting could substantially improve the emergency management of deadly floods: case study July 2021 floods in Germany. Natural Hazards and Earth System Sciences, 22, 9 (2022), 3005-3014.
[6] Arnold, K. C., Volzer, A. M. and Madrid, N. G. Generative Models can Help Writers without Writing for Them. City, 2021.
[7] Árvai, J. The end of risk communication as we know it. Journal of Risk Research, 17, 10 (2014), 1245-1249.
[8] Banholzer, S., Kossin, J. and Donner, S. The impact of climate change on natural disasters. Springer, City, 2014.
[9] Barzilay, R. and Lapata, M. Modeling local coherence: An entity-based approach. Computational Linguistics, 34, 1 (2008), 1-34.
[10] Bica, M., Demuth, J. L., Dykes, J. E. and Palen, L. Communicating hurricane risks: Multi-method examination of risk imagery diffusion. City, 2019.
[11] Bica, M., Weinberg, J. and Palen, L. Achieving accuracy through ambiguity: The interactivity of risk communication in severe weather events. Computer Supported Cooperative Work (CSCW), 29, 5 (2020), 587-623.
[12] Brewer, N. T., Chapman, G. B., Gibbons, F. X., Gerrard, M., McCaul, K. D. and Weinstein, N. D. Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. Health psychology, 26, 2 (2007), 136.
[13] Bubeck, P., Botzen, W. J., Kreibich, H. and Aerts, J. C. Detailed insights into the influence of flood-coping appraisals on mitigation behaviour. Global environmental change, 23, 5 (2013), 1327-1338.
[14] Calzolari, N. Detecting patterns in a lexical data base. City, 1984.
[15] Chakrabarty, T., Zhang, X., Muresan, S. and Peng, N. MERMAID: Metaphor generation with symbolism and discriminative decoding. arXiv preprint arXiv: 2103.06779 (2021).
[16] Charras, C. and Lecroq, T. Handbook of exact string matching algorithms. King's College, 2004.
[17] Chen, M. X., Lee, B. N., Bansal, G., Cao, Y., Zhang, S., Lu, J., Tsay, J., Wang, Y., Dai, A. M. and Chen, Z. Gmail smart compose: Real-time assisted writing. City, 2019.
[18] Cheng, L.-L., Cheung, D. W. and Yiu, S.-M. Approximate string matching in DNA sequences. IEEE, City, 2003.
[19] Clarke, V., Braun, V. and Hayfield, N. Thematic analysis. Qualitative psychology: A practical guide to research methods, 222, 2015 (2015), 248.
[20] Coenen, A., Davis, L., Ippolito, D., Reif, E. and Yuan, A. Wordcraft: A human-AI collaborative editor for story writing. arXiv preprint arXiv:2107.07430 (2021).
[21] De Boer, J., Wouter Botzen, W. and Terpstra, T. More than fear induction: Toward an understanding of people's motivation to be well-prepared for emergencies in flood-prone areas. Risk analysis, 35, 3 (2015), 518-535.

[22] De Hoog, N., Stroebe, W. and De Wit, J. B. The impact of vulnerability to and severity of a health risk on processing and acceptance of fear-arousing communications: A meta-analysis. Review of General Psychology, 11, 3 (2007), 258-285.
[23] Demuth, J. L., Morss, R. E., Morrow, B. H. and Lazo, J. K. Creation and communication of hurricane risk information. Bulletin of the American Meteorological Society, 93, 8 (2012), 1133-1145.
[24] Donovan, A. R., Borie, M. A. H. and Blackburn, S. E. Changing the paradigm for risk communication: integrating sciences to understand cultures. United Nations Office for Disaster Risk Reduction (UNDRR), City, 2019.
[25] Dow, K. and Cutter, S. L. Crying wolf: Repeat responses to hurricane evacuation orders (1998).
[26] Drabek, T. E. and Stephenson Iii, J. S. When Disaster Strikes. Journal of applied social psychology, 1, 2 (1971), 187-203.
[27] Ellouze, S., Jaoua, M. and Belguith, L. H. Automatic evaluation of a summary's linguistic quality. Springer, City, 2016.
[28] Farr, J. N., Jenkins, J. J. and Paterson, D. G. Simplification of Flesch reading ease formula. Journal of applied psychology, 35, 5 (1951), 333.
[29] Field, C. B., Barros, V., Stocker, T. F. and Dahe, Q. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. Cambridge University Press, 2012.
[30] Fischhoff, B. Helping the public make health risk decisions. Springer, City, 1989.
[31] Flesch, R. A new readability yardstick. Journal of applied psychology, 32, 3 (1948), 221.
[32] Gero, K. I., Liu, V. and Chilton, L. Sparks: Inspiration for science writing using language models. City, 2022.
[33] Gero, K. I., Liu, V., Huang, S., Lee, J. and Chilton, L. B. What Makes Tweetorials Tick: How Experts Communicate Complex Topics on Twitter. Proceedings of the ACM on Human-Computer Interaction, 5, CSCW2 (2021), 1-26.
[34] Gui, X., Kou, Y., Pine, K., Ladaw, E., Kim, H., Suzuki-Gill, E. and Chen, Y. Multidimensional risk communication: public discourse on risks during an emerging epidemic. City, 2018.
[35] Gui, X., Kou, Y., Pine, K. H. and Chen, Y. Managing uncertainty: using social media for risk assessment during a public health crisis. City, 2017.
[36] Henderson, J., Nielsen, E. R., Herman, G. R. and Schumacher, R. S. A hazard multiple: Overlapping Tornado and flash flood warnings in a National Weather Service Forecast Office in the Southeastern United States. Weather and Forecasting, 35, 4 (2020), 1459-1481.
[37] Henstra, D. Toward the climate-resilient city: extreme weather and urban climate adaptation policies in two Canadian provinces. Journal of Comparative Policy Analysis: Research and Practice, 14, 2 (2012), 175-194.
[38] Kalyanathaya, K. P., Akila, D. and Suseendren, G. A fuzzy approach to approximate string matching for text retrieval in NLP. J. Comput. Inf. Syst. USA, 15, 3 (2019), 26-32.
[39] Kannan, A., Kurach, K., Ravi, S., Kaufmann, T., Tomkins, A., Miklos, B., Corrado, G., Lukacs, L., Ganea, M. and Young, P. Smart reply: Automated response suggestion for email. City, 2016.
[40] Kher, A., Johnson, S. and Griffith, R. Readability assessment of online patient education material on congestive heart failure. Advances in preventive medicine, 2017 (2017).
[41] Kox, T. Criteria affecting people's decision to take protective measures during winter storm XAVER on 5 December 2013. City, 2015.
[42] Lin, Z., Liu, C., Ng, H. T. and Kan, M.-Y. Combining coherence models and machine translation evaluation metrics for summarization evaluation. City, 2012.
[43] Longo, B. Toward an informed citizenry: Readability formulas as cultural artifacts. Journal of technical writing and communication, 34, 3 (2004), 165-172.
[44] Ludwig, T., Reuter, C., Siebigteroth, T. and Pipek, V. CrowdMonitor: Mobile crowd sensing for assessing physical and digital activities of citizens during emergencies. City, 2015.
[45] Mayhorn, C. B. and McLaughlin, A. C. Warning the world of extreme events: A global perspective on risk communication for natural and technological disaster. Safety science, 61 (2014), 43-50.
[46] Mileti, D. S. and Peek, L. The social psychology of public response to warnings of a nuclear power plant accident. Journal of hazardous materials, 75, 2-3 (2000), 181-194.
[47] Mileti, D. S. and Sorensen, J. H. Warning systems: A social-science perspective. Westview Press, Boulder, Co, 1995.
[48] Morrow, B. and Lazo, J. Effective tropical cyclone forecast and warning communication: Recent social science contributions. Tropical Cyclone Research and Review, 4, 1 (2015), 38-48.
[49] Morrow, B. H., Lazo, J. K., Rhome, J. and Feyen, J. Improving storm surge risk communication: Stakeholder perspectives. Bulletin of the American Meteorological Society, 96, 1 (2015), 35-48.
[50] Morss, R. E., Cuite, C. L., Demuth, J. L., Hallman, W. K. and Shwom, R. L. Is storm surge scary? The influence of hazard, impact, and fear-based messages and individual differences on responses to hurricane risks in the USA. International journal of disaster risk reduction, 30 (2018), 44-58.
[51] Myers, L. NOAA Weather Information and Dissemination All Hazards Stakeholder Project: Phase 1 (2016).
[52] Nielsen, E. R., Herman, G. R., Tournay, R. C., Peters, J. M. and Schumacher, R. S. Double impact: When both tornadoes and flash floods threaten the same place at the same time. Weather and Forecasting, 30, 6 (2015), 1673-1693.
[53] Palen, L. and Anderson, K. M. Crisis informatics—New data for extraordinary times. Science, 353, 6296 (2016), 224-225.
[54] Panteli, M. and Mancarella, P. Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies. Electric Power Systems Research, 127 (2015), 259-270.
[55] Peng, Z., Guo, Q., Tsang, K. W. and Ma, X. Exploring the effects of technological writing assistance for support providers in online mental health community. City, 2020.
[56] Perry, R. W., Lindell, M. K. and Greene, M. R. Threat Perception and Public Response to Volcano Hazard. The Journal of social psychology, 116, 2 (1982), 199-204.
[57] Peters, G.-J. Y., Ruiter, R. A. and Kok, G. Threatening communication: a critical reanalysis and a revised meta-analytic test of fear appeal theory. Health psychology review, 7, sup1 (2013), S8-S31.
[58] Pine, K. H., Lee, M., Whitman, S. A., Chen, Y. and Henne, K. Making sense of risk information amidst uncertainty: individuals' perceived risks associated with the COVID-19 pandemic. City, 2021.
[59] Pitler, E., Louis, A. and Nenkova, A. Automatic evaluation of linguistic quality in multi-document summarization (2010).
[60] Portala, J. Europe's heatwave may have caused more than 20,000 'excess' deaths. Reuters, City, 2022.
[61] Redish, J. Readability formulas have even more limitations than Klare discusses. ACM Journal of Computer Documentation (JCD), 24, 3 (2000), 132-137.
[62] Roemmele, M. Writing stories with help from recurrent neural networks. City, 2016.
[63] Schulze, K., Lorenz, D. F., Wenzel, B. and Voss, M. Disaster Myths and their Relevance for Warning Systems. City, 2015.
[64] Shklovski, I., Palen, L. and Sutton, J. Finding community through information and communication technology in disaster response. City, 2008.
[65] Singla, N. and Garg, D. String matching algorithms and their applicability in various applications. International journal of soft computing and engineering, 1, 6 (2012), 218-222.
[66] So, M., Franks, J. L., Cree, R. A. and Leeb, R. T. An evaluation of the literacy demands of online natural disaster preparedness materials for families. Disaster medicine and public health preparedness, 14, 4 (2020), 449-458.
[67] Soden, R., Chilton, L., Miles, S., Bicksler, R., Villanueva, K. R. and Bica, M. Insights and Opportunities for HCI Research into Hurricane Risk Communication. City, 2022.
[68] Starbird, K. and Palen, L. " Voluntweeters" self-organizing by digital volunteers in times of crisis. City, 2011.
[69] Tarhio, J. and Peltola, H. String matching in the DNA alphabet. Software: Practice and Experience, 27, 7 (1997), 851-861.
[70] Vieweg, S., Hughes, A. L., Starbird, K. and Palen, L. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. City, 2010.
[71] Wachinger, G., Renn, O., Begg, C. and Kuhlicke, C. The risk perception paradox—implications for governance and communication of natural hazards. Risk analysis, 33, 6 (2013), 1049-1065.
[72] Zade, H., Shah, K., Rangarajan, V., Kshirsagar, P., Imran, M. and Starbird, K. From situational awareness to actionability: Towards improving the utility of social media data for crisis response. Proceedings of the ACM on human-computer interaction, 2, CSCW (2018), 1-18.

A APPENDICES

A SUPPLEMENTARY TABLES AND FIGURES

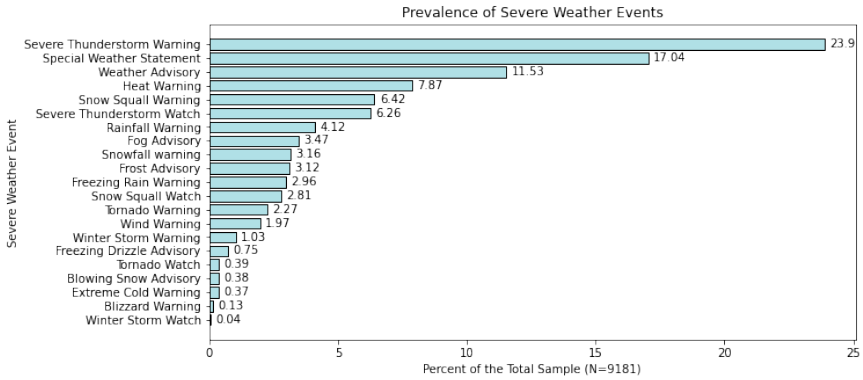

Figure 2: Prevalence of Severe Weather Events in Ontario between June 2020 and December 2021

Table 3: Interpreting the Flesch Reading Ease Score [a]

Score	Difficulty
90 - 100	Very Easy
80 - 89	Easy
70 - 79	Fairly Easy
60 - 69	Standard
50 - 59	Fairly Difficult
30 - 49	Difficult
0 - 29	Very Confusing

[a] Source: TexstStat 0.7.3 (https://pypi.org/project/textstat/)

Table 4: Best Practices in Risk Communication Themes

Theme	Definition
Hazards	Health and/or safety hazards associated with the event
Location	Areas affected by the hazardous event
Timing	Information related to time such as onset, duration, and end time
Affected Populations	Groups vulnerable to hazard impacts
Impacts	Impacts of the hazardous event
Recommended Actions	Any actions or guidance to the public to mitigate the effects of the hazardous event