An Integrated Approach for Automated Acquisition of Bridge Data and Deficiency Evaluation

Abdelhady Omar[1,2] and Osama Moselhi[3]

[1]Ph.D. Candidate, Department of Building, Civil and Environmental Engineering, Concordia University, Canada. [2]Assistant Lecturer, Department of Structural Engineering, Faculty of Engineering, Alexandria University, Egypt. [3]Professor and Director of Center for Innovation in Construction and Infrastructure Engineering and Management (CICIEM), Department of Building, Civil and Environmental Engineering, Concordia University, Canada. Email: abdelhady.omar@mail.concordia.ca, moselhi@encs.concordia.ca

Abstract –

Bridges are essential components of civil infrastructures that enable joining highways and cities, as well as facilitating movement between different geographical locations. Deterioration of bridges can pose a threat to public safety and hinder economic activities. The development of innovative and reliable asset management tools to support the optimization of maintenance and timely intervention plans is of paramount concern to transportation agencies. However, such development is usually impeded by a lack of data and/or difficulties in collecting and organizing data from different sources. This paper introduces an integrated scheme for automated acquisition of bridge-related data. The proposed methodology enables the creation of a unified structured and comprehensive data repository that can be used for further processing and modeling purposes. Data from multiple sources, including web pages and bridge inspection reports, were extracted, collected, prepared, and compiled. As such, web scraping and rule-based information extraction were employed to achieve this objective. Finally, text mining was applied to the inspection comments extracted from bridge inspection reports to get insights into such inspection information and evaluate the severity of bridge deck deficiencies.

Keywords –

Data Acquisition; Data Scraping; Bridge Inspection Report; Document Parsing; Information Extraction; Text Mining; Bridge Deck Deficiencies.

1 Introduction

Bridges are strategic points in the road network because of the severe consequences of their failure or closure to traffic. Measures must therefore be taken to prevent these structures from deteriorating to the point where they compromise the safety of users or require a diversion. Transportation agencies should offer service levels that meet the users' needs. As such, transportation agencies implement several types of inspection programs to identify defects in the structural elements as soon as possible so that the necessary measures can be taken to ensure the durability of the structures and the safety and comfort of the users. In addition, these programs are carried out at regular intervals or, if necessary, depending on the problem identified to gather all the data needed to plan preventive and corrective actions; hence, ensure the safety of structures and protect the invested capital [1].

A crucial first step in maintaining the integrity of the bridge, preventing deterioration proliferation, and avoiding unnecessary Maintenance, Rehabilitation, and Replacement (MR&R) work is determining the best intervention plan [2]. However, current bridge deterioration models and management systems have limited ability to enable justifiable and well-informed decisions regarding bridge intervention strategies since they have concentrated on using just a small collection of abstract data (e.g., year of construction, material type, and bridge geometry). Due to the absence of thorough descriptions of bridge deficiencies of different bridge components (e.g., deck, superstructure, and substructure), such abstract data are insufficient since they restrict the ability to learn from these data to guide intervention planning decision-making and make deficiency-based forecasts [3]. On the other hand, a lot of information regarding bridge deficiencies and deterioration conditions is buried in textual inspection reports without being used. Federal Highway Administration claims that bridge owners are unable to effectively use the "mountain" of inspection data contained in bridge inspection reports [3, 4]. Even worse, there are instances where the abstract data are either unavailable or difficult to collect, which can impede the development of effective decision-making support tools. This issue is prevalent in Quebec, Canada. In these circumstances, planning and executing preventive or corrective measures becomes challenging, compromising the bridge's and its users' safety.

In this context, this paper serves as a key first step in developing innovative and reliable asset management tools to support the optimization of maintenance and timely intervention plans by introducing an integrated scheme for automated acquisition of bridge-related data and defects' severity evaluation for bridges in Quebec. As such, it enables the creation of a structured comprehensive data repository that can be used for further processing and modeling purposes. The repository of data includes not only information regarding the bridge and crossing road characteristics but also data obtained from bridge inspection reports. This data encompasses a wide range of information, such as inspection frequency, percentages of defective material, structural behavior ratings, and inspection comments made by inspectors, which provide information about the various types of defects observed. Web scraping techniques were used to collect and extract bridge and crossing road characteristics data. This process involved automatically retrieving and parsing relevant information from government agency websites. A rule-based information extraction method was employed to extract the important information buried in bridge inspection reports. These reports were automatically downloaded first by navigating through the website and retrieving the reports without any manual intervention. Moreover, text mining techniques were employed to get insights into inspection comments related to the bridge deck in order to explore the most common types of defects, their association, and their severity.

2 Background

Data acquisition methods are required to automatically identify and extract data/information from different sources and formats, e.g., transportation agencies' websites and unstructured textual bridge inspection reports, and then represent the extracted data in a structured format ready for additional data analytics. In this light, web scraping and rule-based information extraction methods were employed to extract and collect bridge-related data. Web scraping is an automated technique for extracting copious unstructured volumes of scattered data from websites. There are several ways to scrape web pages, including using internet services, Application Programming Interfaces (APIs), or writing particular code [5-7]. The application of web scraping and machine learning algorithms was considered to assess the real estate price in the secondary housing market in Moscow [8]. A data-based price indices research was implemented using web scraping [9]. In addition, web scraping and convolutional neural networks were integrated to verify fraudulent content from social media, news articles, and other web pages [10]. On the other hand, information extraction is the process of searching through natural language text for information pertinent to an interest, such as entities, relationships, or events. Information extraction methods are either rule-based or machine-learning-based [3, 11]. Among both methods, the most closely related to this work is rule-based information extraction. Rule-based techniques use manually created pattern-matching-based rules to direct the identification and extraction of target information from unstructured textual input [3, 12]. Syntactic and/or semantic text properties are used to construct the pattern-matching-based rules [3]. Zhang and El-Gohary [13] and Zhou and El-Gohary [14] developed pattern-matching-based rules that integrate syntactic and semantic elements to extract building regulation information for compliance assessment.

Once the data have been extracted from bridge inspection reports, performing text mining to analyze and gain valuable insights from the text-based information becomes essential. Text mining is a data analytics technique that involves searching for patterns in text data through various approaches such as data, analysis, visualization, and application of statistical and machine learning algorithms. Williams and Betak [15] used Latent Dirichlet Analysis, a text mining algorithm, to highlight significant issues in railroad equipment accidents from equipment accident reports published by Federal Railroad Administration. Lv and El-Gohary [16] developed a method using text analytics to extract words and phrases describing stakeholders' concerns from comments on early-stage large-scale highway projects. In order to establish a set of decision-making chains to boost workplace safety, Zhao et al. [17] analyzed occupational safety reports and investigations relevant to electrocution accidents. They identified activities and decision-making errors that raise the worker's safety risk. Mostafa et al. [18] used rule-based text mining to extract the type and extent of damage from textual information from school inspection reports. They used Principal Component Analysis and clustering to categorize events and identify assets with a high priority for renewal.

3 Motivation

Quebec bridges experience significant deterioration levels as well as have the highest average age among all provinces in Canada, followed by Nova Scotia [19]. Several agents, such as exposure to de-icing salts, cycles of freeze and thaw, deferred maintenance, aging, and increasing traffic volumes, contribute to the rapid degradation of these critical civil infrastructures posing diverse burdens and challenges to transportation agencies, specifically under budget constraints and limited resources [20]. The failure, or even service interruption, of these infrastructures can be ruinous in terms of human life, social, environmental, and economic impacts.

In September 2006, the Concorde Boulevard overpass in Laval, Quebec, collapsed, resulting in five deaths and six injuries. It was the first time a young structure had collapsed in the province without apparent cause [21]. Less than a year later, the Interstate 35W bridge over the Mississippi River in Minneapolis, Minnesota, USA, collapsed, causing 13 deaths and 145 injuries [22]. These tragedies raised concerns about deficient bridges in North America and prompted transportation agencies and academia to develop innovative asset management tools and techniques to ensure safety and preserve the value of these assets [20]. However, the development of such smart technology-based methods is greatly hindered due to the lack of effective processing/transforming of raw data into valuable information; and hence, into informed decisions regarding the MR&R of bridge structures. A large amount of raw data is available but not used efficiently for various reasons, summarized as follows.

Data are located in different sources and formats. For example, the main structures' data can be fetched from Quebec datasets in different formats, including comma-separated values, and contains many attributes of different data types (e.g., numerical, categorical, and text) [23]. Such a variety of data types can impede further analysis and processing of data. In addition, it does not include any information about, e.g., crossing road class, bridge length or width, Annual Average Daily Traffic (AADT), and deterioration conditions or defects of different bridge components. On the other hand, some important information, including the aforementioned attributes, is provided on the Quebec Ministry of Transportation and Sustainable Mobility (MTQ) websites [24]. However, four challenges were encountered while trying to collect these pieces of information. First, the website is designed to show only a limited number of records at a time (i.e., 15 records out of 9,779 records), which cannot be downloaded or processed. Second, another page needs to be opened separately to reach specific information about a particular structure. These pages contain different types of unstructured non-downloadable data, meaning that collection of these data would be tedious and error-prone. Third, each structure has a different page to reach and download its inspection report, meaning more than nine thousand websites and download links should be accessed to download and store these inspection reports. Fourth, even more challenging, is how all these data will be filtered, merged, and compiled in one data repository that can be used later for further data analytics.

Furthermore, in addition to being time-consuming, manually extracting information from bridge inspection reports can be error-prone and challenging due to the large volume of unstructured text data. These reports often contain detailed information about the condition of various bridge components and the extent of any defects or deterioration. However, due to the sheer volume of text data and the lack of standardized reporting formats, it is difficult to efficiently and accurately extract the relevant information from these reports. As a result, there is a need for automated tools and techniques that can efficiently and accurately extract information from bridge inspection reports and help identify any potential issues or areas for improvement.

4 Aim and Objectives

In the above context, the main goal of this paper is to create a structured yet comprehensive data repository that can be used for further processing, hence, developing innovative asset management tools for bridge infrastructures. This goal is broken down into the following objectives:

1. Automate the process of bridge-related data acquisition from different sources and formats, including the automatic retrieval of bridge inspection reports;
2. Analyse and pre-process the collected data to be ready for further processing steps;
3. Develop a rule-based method for information extraction from bridge inspection reports;
4. Use text mining to derive insights from text data buried, and not used, in inspection reports; and
5. Introduce how these collected data and insights can be used for future work.

5 Methodology

In order to achieve the objectives of this paper, the high-level methodology is depicted in Figure 1. As shown in Figure 1, the devised methodology comprises two phases. The first phase is a three-tier method designed for data acquisition and compiling. The output of this phase is a structured comprehensive data repository ready for further processing and modeling. The second phase intends for data processing and text mining of inspection comments buried in bridge inspection reports. This phase enables the analysis and visualization of the important information related to different types of defects in bridge decks, in addition to deriving insights from text data. It is worth pointing out that the second phase will continue in future studies and, for the scope of this paper, only text mining of inspection comments related to the bridge deck is considered. The three tiers of the first phase and the two phases are interconnected as indicated by arrows with letters a, b, and c, i.e., the output of one tier or phase can be used for the next tier or phase. The inputs, steps, and outputs of these phases are depicted in Figure 2 and Figure 3.

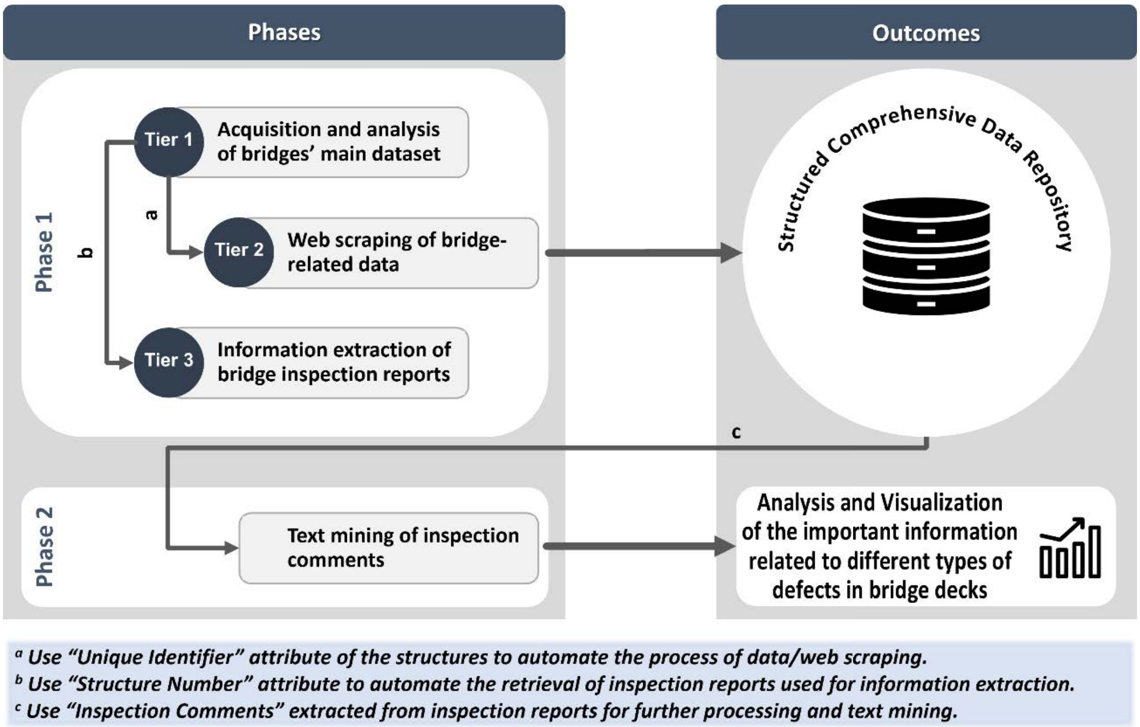


Figure 1. Automated two-phase methodology for bridge data acquisition and text mining of inspection reports

5.1 Phase 1: Data Acquisition and Compiling

As shown in Figure 2, the first stage is a three-tier phase. The first step involves obtaining the primary dataset of structures from the Quebec datasets website [23], which is then stored in a local data repository. Subsequently, an initial analysis is conducted on the data to gain a better understanding of its characteristics and identify any additional features that may need to be collected. The data is then subjected to preliminary preparation techniques, such as integer encoding, to make it suitable for subsequent processing and modeling. Finally, the dataset is filtered to include only bridges with reinforced concrete decks that are relevant to the scope of this study. The second tier of this phase involved web scraping of both the main page and separate pages of MTQ structures in order to collect additional data [24]. This task proved to be difficult due to several reasons. Firstly, the main page of MTQ structures only allows a limited number of records to be opened per page, which is also not accessible directly. Secondly, some pieces of information were stored separately for each structure on different pages with varying layouts. Extracting and manipulating data from over nine thousand structures required a systematic and well-defined approach. Accordingly, the code was designed to scrape each web page separately, merge, and manipulate the extracted data before adding them to the main data repository.

Tier three posed the greatest challenge in the project as it involved handling text data from bridge inspection reports, which lacked any standard or structured format due to being written by different inspectors using different styles and language features. These text data contain information related to deterioration conditions and deficiencies of different bridge components. The inspection reports were initially automatically downloaded from web pages and subsequently analyzed and annotated to identify the necessary information for extraction. Then, a rule-based information extraction approach was adopted, where eighteen different rules were formulated to extract various information pieces from the text. These included the structure number, structure level, last and next general inspection dates, and inspectors' comments, among others. The integration between the Docparser platform [25] and the python programming language through REST API enabled the effective and correct extraction of information into a structured format. Finally, the data were stored in a comprehensive structured data repository. It is worth pointing out that, for the scope of this paper, only inspection comments related to different elements forming the bridge deck were considered for extraction. In the Quebec inspection program [1], the bridge deck system is divided into four elements based on load transmission: two principal elements (deck and exterior sides), and two secondary elements (road/driving surface and drainage system).

Additionally, although the filtering process resulted in 4,119 inspection reports eligible for information extraction, this paper's focus is limited to only 200 reports. To ensure that the information extraction process was performed accurately, the 200 reports were divided into 10-report samples. First, each sample was annotated for the required information, then uploaded to the Docparser to create rules. The extracted information was then verified, cleaned, and prepared. The following sample underwent the same procedure if the information was accurately extracted. Otherwise, the process is repeated on the previous sample until the extracted information is correctly obtained. While these detailed steps were tedious, they were necessary to ensure precise results. Importantly, these 200 reports will serve as a training set for the remaining reports.

5.2 Phase 2: Data Processing & Text Mining

After unifying and structuring all the data, further data analytics can be performed. This paper is a part of a larger research project that aims to use this data to develop bridge management tools. This paper focuses on text mining of inspection comments extracted from bridge inspection reports. Accordingly, various text mining tools and techniques were employed to get insights into these inspection comments, as depicted in Figure 3. The text mining output includes frequent defects found on bridge decks, common comments made by inspectors, correlations between different defects, and their severity levels. This information could further inform maintenance and repair decisions and ultimately improve the safety and longevity of the bridge deck.

In Natural Language Processing (NLP), there are several steps to prepare text data for analysis. The first step is to remove stop words, which are commonly used words that do not add value to the analysis, such as "a," "an," and "the" in English. It is important to carefully select the list of stop words based on the project rather than relying on pre-built lists from libraries. Punctuation marks are also removed since they do not contribute to the analysis and may lead to incorrect results. Converting the text to the same case, preferably lowercase, is crucial to prevent different-case words from misleading the analysis. Tokenization is the process of dividing the text into token-sized pieces, which can be individual words, phrases, or complete sentences. Certain characters, such as punctuation marks, may be removed during tokenization. Typically, tokens are used as input in parsing and text-mining procedures. Lemmatization is applied to convert any word to its basic root mode and combine several word inflections with the same meaning into their root forms. Finally, the results are cleaned and visualized after each step to ensure the correct form of results is obtained since text analysis is very sensitive to changes and sometimes requires manual revision.

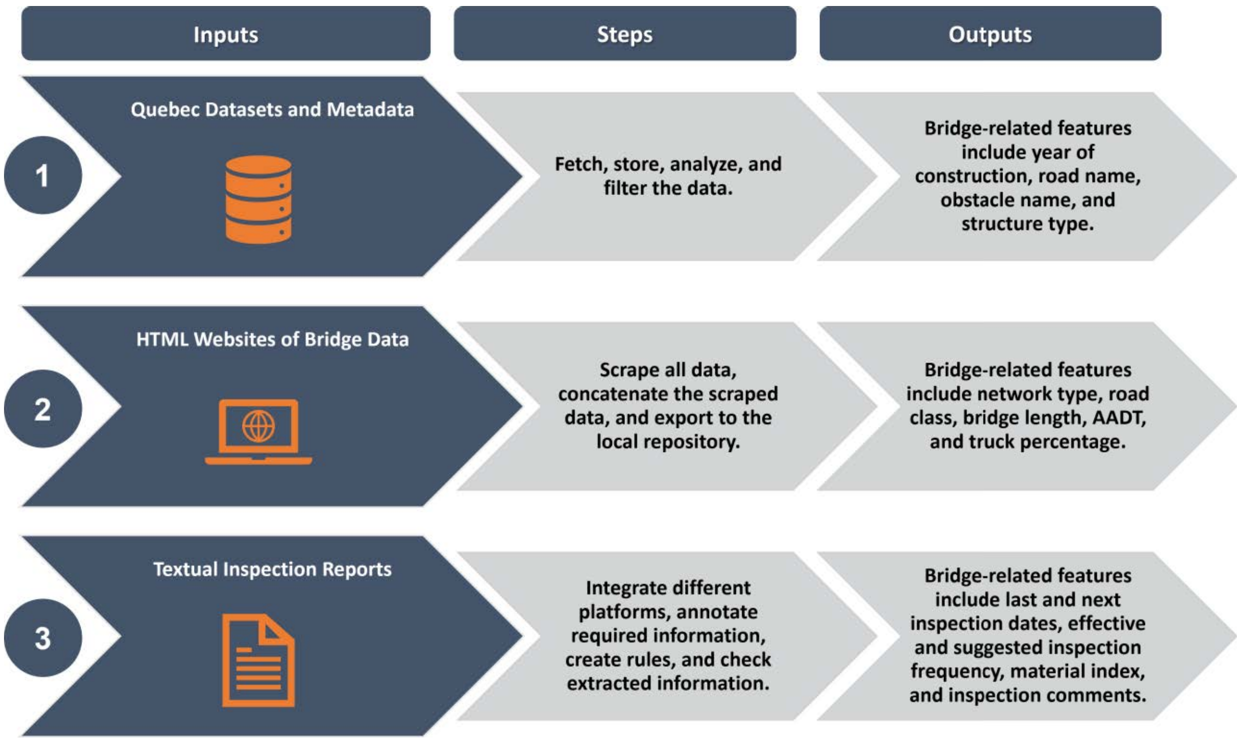


Figure 2. Inputs, steps, and outputs of the data acquisition and compiling phase and its three tiers

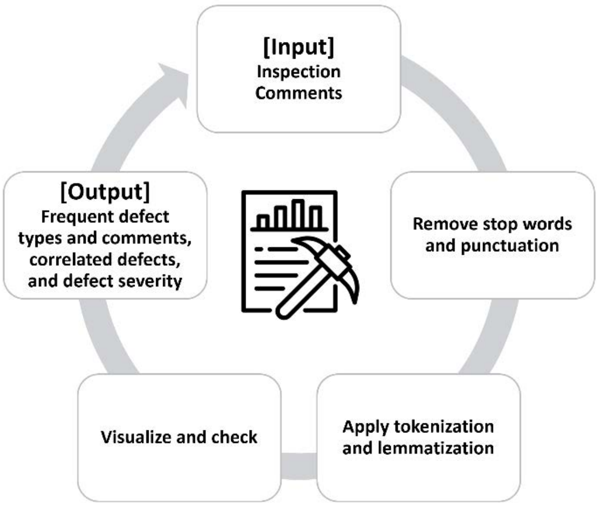


Figure 3. Inputs, steps, and outputs of the data processing & text mining phase

6 Results and Discussions

The results of the proposed approach can be summarized into two major outcomes. Firstly, a structured and comprehensive data repository is generated that is ready for further processing and modeling steps. Secondly, valuable insights into the inspection comments related to the bridge deck are obtained, including frequent defect types, common comments, correlations between different defects, and their corresponding severity levels. The first outcome involves consolidating various attributes related to bridge structures and crossing roads into a single repository. These attributes include, but are not limited to, structure number, year of construction, structure type, bridge geometry, road class, AADT, inspection frequency, and inspection comments. The repository is built upon the Statistical Package for the Social Sciences (SPSS) software, enabling effective organization, management, and data analysis. One of the most crucial features of the repository is the ability to view and modify variable properties within a dataset, including the attribute name and length; data type and measure (e.g., numeric, string/nominal, or date/time); descriptive labels for easier identification and interpretation; and the ability to specify values and assign corresponding labels similar to integer encoding. In addition to its ability to effectively import, manipulate, and analyze large datasets, this tool also offers features for identifying missing values, duplicates, and outliers. It also provides various statistical procedures, including descriptive statistics and inferential statistics. Various visualization tools, including charts, graphs, and tables, can explore and present the data meaningfully, aiding preliminary analysis before further advanced modeling.

The second outcome includes the results of the text mining of inspection comments. As shown in Figure 4, cracking, spalling, and delamination are the most common defects, followed by rust, efflorescence, corrosion, and scaling. Cracks have a significantly more absolute and relative frequency than other words. For more information about these defect types and their progression into concrete, the reader may refer to this reference [20]. Crack-related comments were further analyzed for a more thorough analysis of inspection comments, as depicted in Figure 5. Figure 5a presents the most common phrases inspectors use to describe defects of elements with cracks associated with their relative frequency. The comments also show the severity of the cracks, i.e., crack width, which ranges from “less than 0.80 mm” to “more than 15 mm,” with low severity cracks being the most frequent (relative frequency of 66%). In addition, these phrases can be transformed into information about defect severity and material condition, as shown in Figure 5b. The guidelines for classifying cracks based on their severity and associated material condition are provided in Table 1. The coexistence of cracks with another defect was examined by calculating the correlation of being in the same sentence. As shown in Figure 5c, cracks probably coexist with spalling (correlation of 67.32%) and reinforcement corrosion (correlation of more than 65%).

The created data repository can serve as a kick-off point for building innovative tools to support the optimization of maintenance and budget allocation of bridge structures. In addition, the information extracted from inspection comments can help develop probabilistic defect-based deterioration models and intervention strategies. For example, ongoing research is currently exploring the relationship between the severity levels of different defects, such as cracking, corrosion, delamination, scaling, and spalling, and the probability of a bridge deck being in a certain condition. This investigation is being accomplished through the use of a Bayesian-belief-network model, with the output of text mining being used to determine the marginal and conditional probability for the model. Moreover, other important attributes, such as the year of construction and AADT, will be included in the analysis to identify the most critical factors affecting bridge decks' deterioration.

Table 1. Material condition and degree of crack severity based on crack opening/width (adapted from [1])

Material Condition	Crack Opening (mm)	Degree of Severity
A	None	Slight
B	< 0.8	Moderate
C	≥ 0.8 & ≤ 3.0	Significant
D	> 3.0	Severe

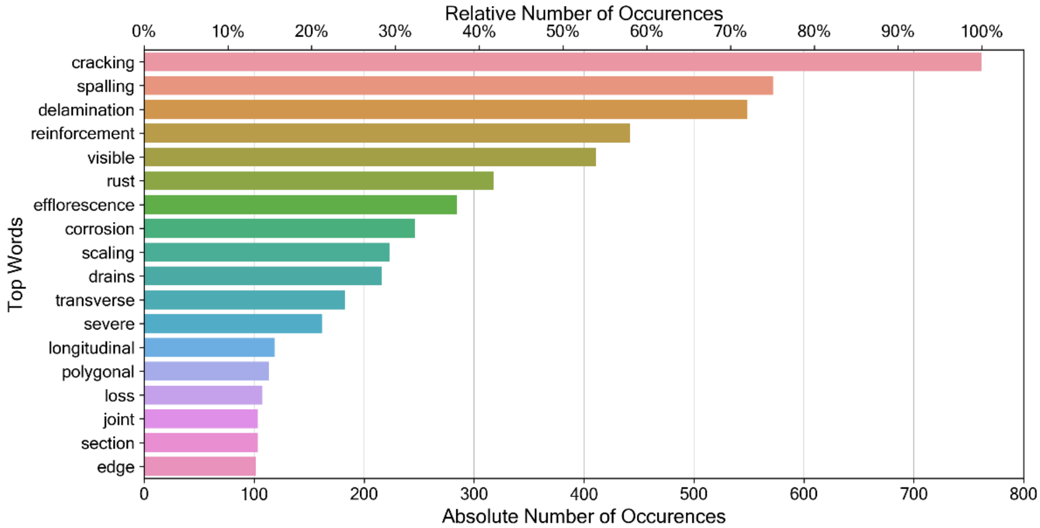
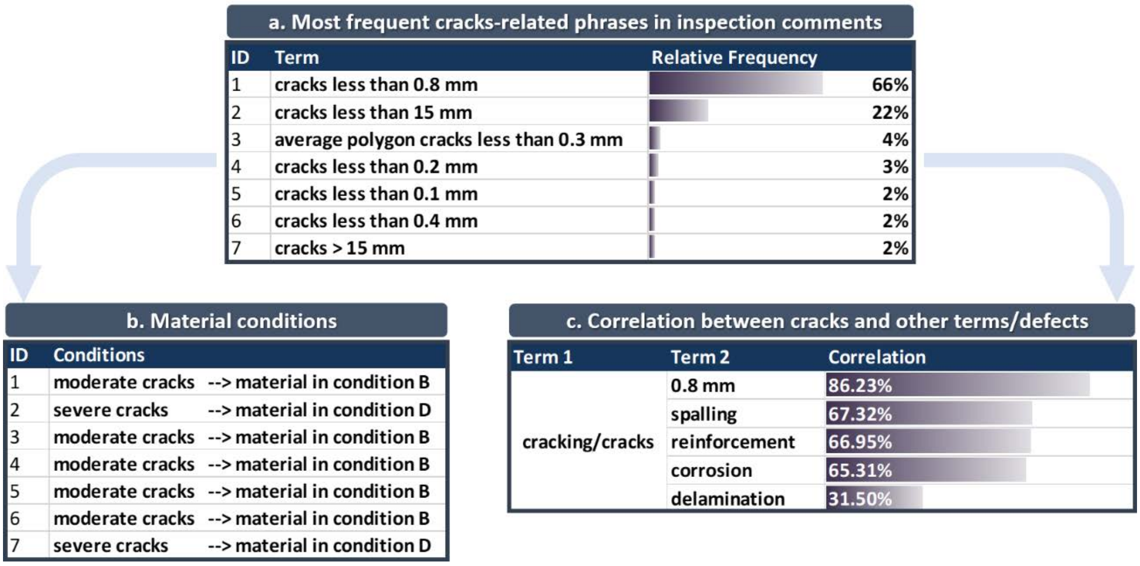


Figure 4. Absolute and relative frequencies of top words/defects appeared in inspection comments



a. Most frequent cracks-related phrases in inspection comments

ID	Term	Relative Frequency
1	cracks less than 0.8 mm	66%
2	cracks less than 15 mm	22%
3	average polygon cracks less than 0.3 mm	4%
4	cracks less than 0.2 mm	3%
5	cracks less than 0.1 mm	2%
6	cracks less than 0.4 mm	2%
7	cracks > 15 mm	2%

b. Material conditions

ID	Conditions
1	moderate cracks --> material in condition B
2	severe cracks --> material in condition D
3	moderate cracks --> material in condition B
4	moderate cracks --> material in condition B
5	moderate cracks --> material in condition B
6	moderate cracks --> material in condition B
7	severe cracks --> material in condition D

c. Correlation between cracks and other terms/defects

Term 1	Term 2	Correlation
cracking/cracks	0.8 mm	86.23%
	spalling	67.32%
	reinforcement	66.95%
	corrosion	65.31%
	delamination	31.50%

Figure 5. Analysis of cracks-related inspection comments

7 Conclusions

This study introduced an integrated approach for automatically collecting bridge-related data and deficiency evaluation. Transportation agencies' websites and unstructured textual bridge inspection reports were considered to gather information utilizing web scraping and information extraction techniques. The study enabled the creation of a structured data repository ready for further data analytics and the development of innovative asset management tools. In addition, text mining was used to get insights into the inspection comments buried in inspection reports and always overlooked. Text mining results showed that cracking, spalling, and delamination are the most common defects in bridge decks. These inspection comments were also transformed into severity levels that can be used for developing probabilistic deterioration models. The contribution of this study lies in integrating different techniques and tools to facilitate the process of data acquisition besides creating a structured data repository that encompasses the most important bridge-related attributes. It also introduced the application of text mining of inspection comments for developing probabilistic deterioration models that, to the best of the authors' knowledge, never used before in that domain, besides laying the foundation for future work.

References

[1] Ministry of Transport and Sustainable Mobility. *Structure Inspection Manual*. Quebec Government, National Library and Archives of Quebec, 2017.

[2] Abdelkhalek S. and Zayed T. Comprehensive Inspection System for Concrete Bridge Deck Application: Current Situation and Future Needs. *Journal of Performance of Constructed Facilities,* 34(5): 03120001, 2020.

[3] Liu K. and El-Gohary N. Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports. *Automation in construction,* 81: 313-327, 2017.

[4] Hooks J. M. and Frangopol D. M. LTBP bridge performance primer. On-line: https://rosap.ntl.bts.gov/view/dot/26946, Accessed: Jan 30, 2023.

[5] Lubis V. How to Scrape Data from a Website using Python for Beginner. On-line: https://medium.com/analytics-vidhya/how-to-scrape-data-from-a-website-using-python-for-beginner-5c770a1fbe2d, Accessed: Jan 30, 2023.

[6] Foo E. What is Web Scraping and How Does It Work. On-line: https://www.datasciencecentral.com/what-is-web-scraping-and-how-does-it-work/, Accessed: Jan 30, 20232.

[7] Hiremath O. S. A Beginner's Guide to learn web scraping with python! On-line: https://www.edureka.co/blog/web-scraping-with-python/, Accessed: Jan 03, 2023.

[8] Mamedli M. O. and Umnov A. V. Real estate valuation based on big data. *Voprosy Ekonomiki,* 2022(12): 118-136, 2022.

[9] Uriarte J. I., Ramírez Muñoz De Toro G. R., and Larrosa J. M. C. Web scraping based online consumer price index: The "iPC Online" case. *Journal of Economic and Social Measurement,* 44(2-3): 141-159, 2020.

[10] Vishwakarma D. K., Meel P., Yadav A., and Singh K. A framework of fake news detection on web platform using ConvNet. *Social Network Analysis and Mining,* 13(1), 2023.

[11] Hobbs J. R. and Riloff E. *Information Extraction*. HANDBOOK OF NATURAL LANGUAGE PROCESSING, SECOND EDITION. Taylor & Francis Group, 2010.

[12] Sarawagi S. Information extraction. *Foundations and Trends® in Databases,* 1(3): 261-377, 2008.

[13] Zhang J. and El-Gohary N. M. Semantic NLP-based information extraction from construction regulatory documents for automated compliance checking. *Journal of Computing in Civil Engineering,* 30(2): 04015014, 2016.

[14] Zhou P. and El-Gohary N. Ontology-based automated information extraction from building energy conservation codes. *Automation in Construction,* 74: 103-117, 2017.

[15] Williams T. P. and Betak J. F. *Identifying Themes in Railroad Equipment Accidents Using Text Mining and Text Visualization*. International Conference on Transportation and Development 2016. 2016.

[16] Lv X. and El-Gohary N. Text analytics for supporting stakeholder opinion mining for large-scale highway projects. *Procedia Engineering,* 145: 518-524, 2016.

[17] Zhao D., McCoy A. P., Kleiner B. M., Du J., and Smith-Jackson T. L. Decision-making chains in electrical safety for construction workers. *Journal of Construction Engineering and Management,* 142(1): 04015055, 2016.

[18] Mostafa K., Attalla A., and Hegazy T. Data mining of school inspection reports to identify the assets with top renewal priority. *Journal of Building Engineering,* 41: 102404, 2021.

[19] Statistics Canada. Age of Public Infrastructure: A Provincial Perspective. On-line: https://www150.statcan.gc.ca/n1/pub/11-621-m/11-621-m2008067-eng.htm, Accessed: Mar 08, 2022.

[20] Omar A. and Moselhi O. Condition Monitoring of Reinforced Concrete Bridge Decks: Current Practices and Future Perspectives. *Current Trends in Civil & Structural Engineering - CTCSE,* 8(4), 2022.

[21] Crisis and Disaster Management Research and Training Initiative at Syracuse University. Collapse of Concorde Boulevard Bridge. On-line: https://cdm.syr.edu/research/case-studies/, Accessed: Jan 30, 2023.

[22] Minnesota Legislature. Minneapolis Interstate 35W Bridge Collapse. On-line: https://www.lrl.mn.gov/guides/guides?issue=bridges, Accessed: Jan 30, 2023.

[23] Quebec Data Partnership. Structures Dataset. On-line: https://www.donneesquebec.ca/recherche/dataset/structure, Accessed: Sep 16, 2022.

[24] Ministry of Transport and Sustainable Mobility. Inventory and Inspection of Structures. On-line: https://www.transports.gouv.qc.ca/fr/projets-infrastructures/structures/Pages/inventaires-structures.aspx, Accessed: Sep 27, 2022.

[25] docparser. Extract Data From Your Business Documents. On-line: https://docparser.com/, Accessed: Oct 31, 2022.