Smart city: an advanced framework for analyzing public sentiment orientation toward recycled water

Mohamed Bahra, Abdelhadi Fennan

LIST Department of Computer Science, Faculty of Sciences and Techniques, Abdelmalek Essaadi University, Tangier, Morocco

Article Info	ABSTRACT
	The coronavirus pandemic of the past several years has had a profound impact on all aspects of life, including resource utilization. One notable example is the increased demand for freshwater, a lifeblood of our planet, on the other hand, the smart city vision aims to attain a smart water management goal by investing in innovative solutions such as recycled water systems. However, the problem lies in the public's sentiment and willingness to use this new resource which discourages investors and hinders the development of this field. Therefore, in our work, we applied sentiment analysis using an extended version of the fuzzy logic and neural network model from our previous work, to find out the general public opinion regarding recycled water and to assess the effects of sentiments on the public's readiness to use this resource. Our analysis was based on a dataset of over 1 million text content from 2013 to 2022. The results show, from spatio-temporal perspectives, that sentiment orientation and acceptance-behavior towards using recycled water have increased positively. Additionally, the public is more concerned in areas driven by the smart city vision than in areas of medium and low economic development, where investment in sensibilization campaigns is needed. *This is an open access article under the [CC BY-SA](#) license.*

Corresponding Author:

Mohamed Bahra
LIST Department of Computer Science, Faculty of Science and Techniques, Abdelmalek Essaadi University
Km 10, Ziaten, BP 416, Old Airport Road, Tangier, Morocco
Email: bahra002@gmail.com

1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) that we have witnessed and which was recognized as a worldwide pandemic on March 11, 2020, has drastically affected our everyday lives, especially due to the implementation of lockdowns as one of the main precautions used to curb the spread of the disease. Many sectors and resources have been immediately impacted, both in terms of consumption and production, such as energy, industry, and food supply, to name a few. Water consumption was not the exception as it was considered one of the most effective mitigation measures against COVID-19 transmission through handwashing [1], and sanitization. The United Nations Economic and Social Commission for Western Asia (ESCWA) has cited in its Policy Brief.5 that the household water demand will represent in the Arab region only, an average increase of 5%, a value equivalent to 4-5 million cubic meters per day due to COVID-19.

The global population's growth and increasing demands for safe water in agriculture, industry, and municipalities have amplified the need for freshwater, as noted in [2]. This escalating demand has led to a significant water supply-demand gap, a grave threat that [3] emphasizes cannot be mitigated solely by existing groundwater and water supply resources. Effective solutions demand comprehensive strategies, conservation, and innovative technology, necessitating collaborative efforts for sustainable water security.

Smart city on the other hand, aims through the usage of information and communication technologies to enhance people welfare [4], [5], and efficiently manage economy, living, mobility, people, governance, and environment that include waste water management [6]. To this end, governments and many world organizations have started giving more attention to water scarcity, naming the Organization for Economic Co-operation and Development (OECD), by working with regions and countries to reform water policies and exchange international community best practices, innovations, and effective approaches for better water management, as well as to raise awareness among citizens about this challenge and the importance of recycled water usage. But the issue is with the acceptance and the scale of usage of this new resource, which is still low even with the quoted quality standard reached by advanced purification technologies such as biofilm [7]. Many research and studies of people's behavior agreed that emotions, sentiment preferences and opinions directly affect a person's behavior towards a subject [8], [9], and this may be the case of public willingness to use recycled water.

Through our work, we aim first to enhance the approach used in our previous work [10] by including domain ontology to extract domain-specific data from social media platforms. Additionally, we added a sentiment word embedding layer that use fuzzy logic linguistic functions [11] to the word embedding layer used by long-shirt-term-memory (LSTM) model. Also, we provide a good explanation of each step of our proposed framework to help future researchers who want to apply fuzzy logic and deep learning in sentiment analysis. Second, we wanted to explore the impact of public sentiment towards recycled water reuse, by benefiting from the immense data published through social media platforms, which gives people the freedom of expression, anonymously and without any external pressure or social barrier.

Our remaining work is structured as follows: The background research for our study is described in section 2, the key concepts used in our work are presented in section 3, and our sentiment analysis system is described in section 4, followed by section 5, in which we present the results and discuss public sentiment orientation, and finally, we conclude our work and highlight our perspectives.

2.   RELATED WORK
2.1. Water scarcity and recycled water alternative

Addressing water scarcity has long been, and continues to be, an area of research and development, due to it is connection to the global vision of transitioning to green energy [12], challenges posed by climate change, and the increasing demand caused by the rising global population, to name a few. Srinivasan *et al.,* [13] examined the causes and nature of the world's water dilemma, by applying a qualitative comparison analysis (QCA) to 22 case studies conducted in different regions, their results show that water crisis issues are grouped into six syndromes that falls into "demand changes", "supply changes", "governance systems", and "infrastructure/technology" categories. Mancuso *et al.*, [14] discuss the value of recycled water as a strategy for overcoming the problem of water scarcity, particularly as a supply for irrigation. Another research in [15] also draw attention on using reclaimed water for agriculture, however, the authors point to the impact of irrigation water on food crops and human health, due to the widespread consumption of raw fruits and vegetables, which can cause public health risks, and increase the doubt about using this resource. The COVID-19 pandemic has also influenced people's behavior and increased public sensitivity to alleged health dangers [16], [17]. Vakula and Kolli [18] emphasize the importance of utilizing recycled water in the development of smart cities and outlined the stages of treatment for recycled water. They also suggest incorporating sensors to automate the process at each stage, resulting in more efficient water management.

2.2. Impact of sentiment preferences on recycled water use

Emotions and sentiment preferences are known to be factors that directly affect people's behavior [19]. Accepting the usage of reclaimed water by citizens requires from governments to sensibilize and involve the public through promotion methods that describe the processes and technologies used in this field. To do so, it is crucial to understand the current state of public sentiment orientation towards recycled water. In both [20], [21], Nkhoma *et al.* analyze the potential permissive factors for treated water and the effect of the 'yuck factor' or the disgust emotion related to it, their finding show that ethnicity, education level and the disgust factor in public psyche influence the number of recycled water supporters. Disgust sensitivity and psychological aspect were addressed in [22], based on the analysis of two surveys in which question-answer methods were used, their results show that people's intuition and feeling can contradict their rational self-interest, which makes recycled water acceptance a psychological issue. Another work by Leong [23] supports the previously mentioned work by using the Q technique to look at how emotions affect drinking recycled water, his findings show that narratives including disgust, anger and fear, emphasize the negativity bubble related to recycled water use, and these emotions can affect ecological policy making and discourage

stakeholders. In [24] a sentiment analysis was applied on text content from social media for exploring public's attention and current status regarding the adoption of recycled water use in China.

2.3. Methods used for sentiment analysis

The work by Medhat *et al.* [25] provide through a list of articles analysis, an exhaustive explanation of sentiment analysis field, including its methodologies, classification techniques, related fields, and the trend of researchers using these techniques. In [26], [27] Malviya *et al.* highlighted algorithms namely k-nearest neighbors (KNN), multilayer perceptron (MLP), naive Bayes (NB), and support-vector machine (SVM), when discussing machine learning techniques applied in sentiment analysis (SA), they also presented a comparison of the performance of these algorithms for SA tasks. Deep learning, as described in [28], is a more in-depth technique of machine learning that can also be applied to sentiment analysis [29]. In [30], the authors applied LSTM model for sentiment analysis on text reviews. They constructed a model consisting of a dense layer, an LSTM layer, and an embedding layer, which was then trained and tested using Amazon and the internet movie database (IMDB) datasets, and the findings indicate that the LSTM model had an accuracy of 85%. The work in [31] addresses the issue that most pre-trained word embeddings for sentiment analysis confront., i.e., two words with alike contexts can be close in term of linguistic properties in word embeddings, yet they could have opposite feelings (e.g., good and bad) [32], which reduces words discrimination and leads to poor performance of these models. The authors propose two approaches to tackle this problem. The first method entails adding sentiment polarity to already-trained word embeddings, while the second method is based on learning word embeddings from contexts, and consider contextual relationships and sentiment associations between words. Fu *et al.*, [33] also highlight the problem encountered by word embeddings, and propose a model that incorporates a sentiment lexicon into LSTM, adding sentiment information to word's representation. Fuzzy logic for sentiment analysis was proposed in [34], where the authors discuss the implementation of a model that can analyses social media content in order to capture users perceptions and opinions regarding products and services.

The application of sentiment analysis using social media content poses a challenge related to extracting domain-specific data. In [35], Wongthongtham and Salih address this task by proposing an ontology-based method to semantically analyze and extract social content, their experiments show that using ontology to capture domain knowledge yield better results. Some of the related studies cited in this work are compiled in Table 1.

Table 1. Summary of cited studies

Reference #	Proposed	Finding	Comments related to our work
Mancuso *et al.* [14]	Investigate the importance of using reclaimed water to cope with the problem of freshwater scarcity in the Mediterranean area.	Based on the information related to wastewater in the Mediterranean area, authors confirm the necessity of reclaimed water use to lower the pressure on freshwater.	The study ignores to highlight the importance of the population factor in determining whether or not recycled water resources will be accepted and used.
Nia *et al.* [16]	The investigation examines how psychological behavior is affected by anxiety and fear associated with COVID-19.	According to the research, individuals' psychological behavior. Is impacted by anxiety sentiments and fear of COVID-19 affect.	Highlights the importance of considering sentiment analysis in studies related to finding individual behaviors regarding a subject (e.g., recycled water).
Vakula and Kolli [18]	Treats the necessity of wastewater treatment for smart cities, and the usage of sensors for the automation of these treatments.	The proposition of using sensors in every stage of wastewater treatment and providing flowchart automation of the treatment plant.	Authors confirm the importance of using recycled water for the development of smart cities but lack to provide limitations that could be associated with people's acceptance of using this resource.
Rozin *et al.* [22]	Analyzing recycled water rejection based on psychological aspects i.e., disgust, contamination, and purification.	Accepting recycled water usage is a psychological matter, and most people can accept it, while a minority refuse this water resource because of the emotion of disgust and contamination.	The questionnaires used in the analysis were distributed to individuals in public spaces, which may have limited the respondents' freedom to express their true opinions and potentially influenced the results due to the presence of others.
Li *et al.* [24]	Using social media text content for text mining analysis to gather public opinion and attention on recycled water.	Results show that the overall sentiment regarding reclaimed water is positive, however, a part of the public lacks knowledge about it and still holds negative attitudes.	The analysis focused only on the country of China, whereas in our work, we provide a worldwide vision of the people's sentiment orientation toward recycled water and also provide a detailed and advanced framework for sentiment analysis that can aid the future researcher in the sentiment analysis field.
Murthy *et al.* [30]	The utilization of LSTM model for text reviews' sentiment analysis.	The study provides a detailed explanation of LSTM model and shows that the model gives better performance in sentiment analysis tasks with an accuracy of 85%.	The proposed model misses the implementation of the text cleaning phase, as the content shared on social networking platforms (e.g., Facebook, Twitter) is usually noised. Also, the word embedding used does not consider the sentiment factor in its word's representation.
Fu *et al.* [33]	Using a sentiment lexicon to extract and add sentiment information to word representation and feed the LSTM model with a combination of word embedding and its associated sentiment embedding.	Incorporating sentiment information into word embeddings improves the accuracy of the LSTM model for sentiment analysis tasks	We were inspired by this work to add sentiment word embedding to word embedding and we included fuzzy logic linguistic functions to extract the degree of sentiment expressed.

3.   MAIN CONCEPTS

3.1. Ontology

For any given data analysis to produce relevant and meaningful results, the quality of the data used is a major factor. In social media platforms, Twitter only takes a part of more than 12 Terabytes of data generated daily, which is immense and touches almost every domain. And here where ontology comes in handy. The definition of ontology has various meanings depending on the domain where it is solicited. In data science, the most cited definition of ontology is "An explicit specification of a conceptualization" [36]. The ontology's main purpose is to gather domain specifics concepts, objects, and entities and describe the relations between them in a way that form a domain knowledge base, shared, and understood by human and machine without semantic ambiguation.

3.2. Sentiment analysis

Sentiment analysis, often referred to as opinion mining [37], is a crucial facet of natural language processing. This research area aims to investigate human emotions present in text content by using different approaches and algorithms for extracting the sentiment expressed and categorize it as positive, negative, or neutral. In Figure 1, advanced methods employed by researchers and developers are depicted, illustrating their efforts to address the challenges inherent in sentiment analysis.

3.3. Fuzzy logic

As human being, our cognition and thinking process does not follow strict binary patterns, but rather, approximate reasoning, many decisions we make in our life are not of absolute value (i.e., true, or false). This idea was the major foundation for the establishment of fuzzy logic theory. In [38], Zadeh, who is acknowledged as the "father of fuzzy logic," established the idea of fuzzy sets, which expanded upon the traditional binary logic system of zero and one, into a set that takes into account the values in between. Fuzzy logic aims to mimic the human decision process through fuzzy sets and membership functions that help to quantify the degree of truthfulness and falseness [39]. By projecting this concept to the sentiment analysis field, fuzzy logic can be used to define the degree of sentiment expressed, i.e., moving from the absolute values "Good" or "Bad" to more sentiment degree values e.g., "Very good", "Good", "Neutral", "Bad", "Very bad". Forward we present some fuzzy logic components that we incorporate in our work.
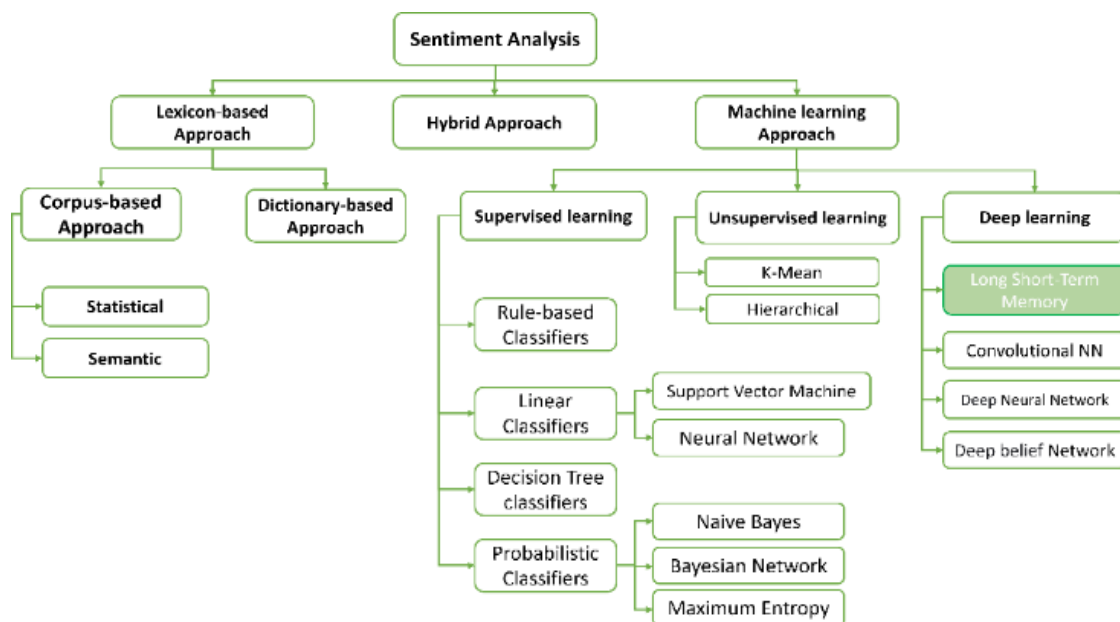
Figure 1. Well-known approaches used in sentiment analysis field

a. Fuzzy set: A fuzzy set, rooted in fuzzy set theory, represents a broader concept than a conventional set. It employs a membership function to establish a class of entities, wherein each element possesses a degree of belongingness to the class, ranging from zero to one. This nuanced approach allows for a more flexible and nuanced representation of uncertainty and vagueness in data and decision-making processes.

b. Membership function: It represents a characteristic function that define the membership level of an entity z in the class A belonging to the universe of discourse U. In an ordinary set, the characteristic function $f_A(z)$ will have only two values, which are 0 or 1, however in a fuzzy set, the function characteristic will be defined by (1).

$$\begin{cases} f_A(z) = 0, if\ z\ is\ not\ in\ A \\ f_A(z): U \rightarrow [0,1]; \quad f_A(z) = 1, if\ z\ is\ totaly\ in\ A \\ 0 < f_A(z) < 1, if\ z\ is\ partially\ in\ A \end{cases} \tag{1}$$

c. Linguistic variable: In his study [40], Zadeh examines and defines the linguistic variable as one with words or phrases as values as opposed to numbers. For a given 'Age' variable, the values will be represented by "very young", "young", "not young", "quite young", in place of number values e.g., 14, 20, 40, ... in our work, we incorporated this concept as a sentiment embedding feature.

d. Linguistic hedges: presented in [41], they represent terms that can be treated as operators for their operands. Terms such as "very", "quiet", "Not", "much", "more or less" can act on the fuzzy sets defining the meaning of the associated operands. Many operations are associated with linguistic hedges, such as accentuation, convex intensification, intersection, concentration, complementation, and dilation. In our analysis, we focused only on the last three operations.

e. Concentrator modifier: In the context of linguistic hedges, the hedge modifiers such as "very, extremely, positively" are identified as intensifiers that can reinforce the characterization of their operands. Applying a concentration operation to a fuzzy set A in the universe of discourse U, is denoted by CON(A)=A2 and defined in the (2):

$$\{ \ U_{CON(A)}(y) = U_A^2(y), \ y \in U \tag{2}$$

f. Complement modifier: correspond to negation operation. For a fuzzy set A in U, the complement of A is denoted by -A and defined by the (3):

$$\{ \ U_{-A}(y) = 1 - U_A(y), \ y \in U \tag{3}$$

g. Dilator modifier: also called, weakening modifiers, (e.g., "more or less", "negatively") represent the opposite effect of the concentrators, thus decreasing the characterization of their operands. According to this modifier, applying a dilation operator on a fuzzy set A, results in a fuzzy set indicated by DIL(A)=A$^{0.5}$, defined in (4):

$$\{ \ U_{DIL(A)}(y) = \sqrt{U_A(y)}, \ y \in U \tag{4}$$

3.4. Long shirt-term memory network (LSTM)

The long shirt-term memory (LSTM) is a type of recurrent neural network (RNN), developed by Hochreiter and Schmidhuber [42] in 1997 to address the exploding and vanishing gradients issues encountered by the RNN [43]. This network can hold information in the memory for a lengthy period using a "memory cell" which makes it a context-aware network that is suitable for problems such as speech recognition, machine translation, and natural language processing (e.g., sentiment analysis). The LSTM model consists of a cell state, also called "memory cell" that maintains data for an extended periods, and three gates i.e., input, forget, and output gates that regulate adding and removing data from the cell state [44].

4. PROPOSED FRAMEWORK

Our sentiment analysis framework consists of four principal modules, each of which is composed of multiple stages, organized in a specific pipeline: data extraction module, text pre-processing module, content vectorization, fuzzy sentiment scoring module, and modal training and analytics console module. An overview of our framework's architecture can be found in Figure 2. We shall outline each component that each module in our architecture is made up of in the ensuing subsection.

4.1. Data extraction

In our framework, the data extraction module is the part that focuses on gathering all the data we used for our sentiment analysis. Knowing that social big data is scattered and diversified, it is challenging to get fine and domain content specific. To this end, we incorporated a domain ontology component in which

we define the searching keywords related to recycled water. Figure 3 represents our recycled water domain ontology used to collect data from social media Twitter platform.

Forward, we used social media extractor component which is responsible for setting up the connection with the social media application programming interface (API). We used an academic research application for Twitter, which gives us free access to all Twitter APIs, including the full-archive search endpoint, which allows retrieving Twitter historical data. We customized the API request to focus on getting the following tweet attributes 'id' of the tweet, 'created_at', 'text', 'source', 'username', 'location', and 'country'. Table 2 shows some data samples. Finally, we stored the collected data in our database for further processing.
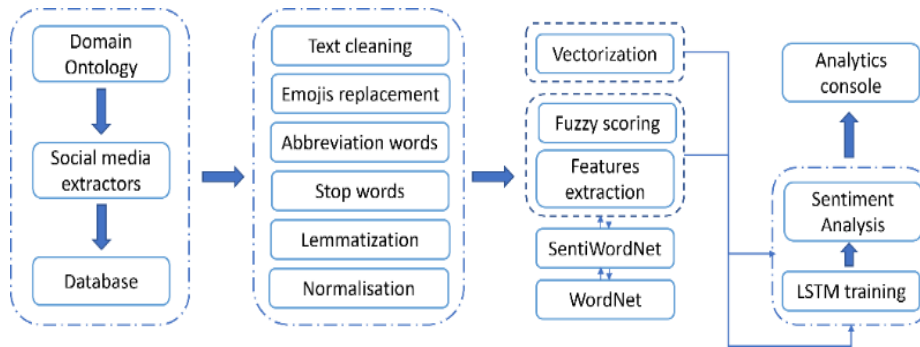
Figure 2. Global architecture of proposed framework
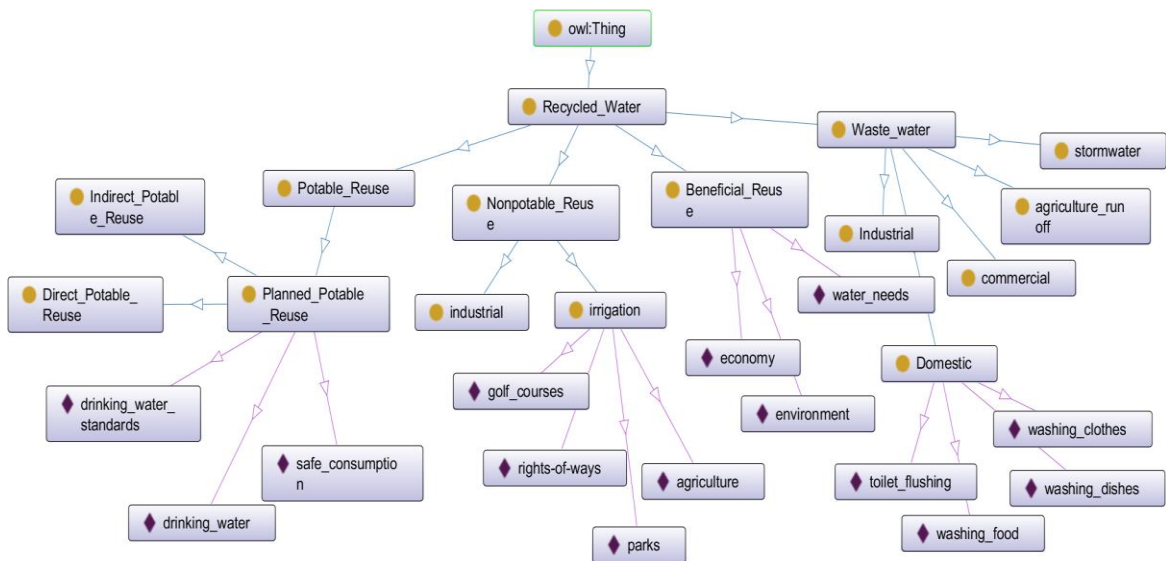
Figure 3. Recycled water domain ontology

Table 2. Samples of tweets collected using social media twitter API

Tweet Id	Username	Text
1572409035273699329	BurbankH2OPower	Today was the last day for the Recycled H2O to GO fill station! \n\nThank you to more than 100 residents who used recycled water to keep their plants healthy. We had over 500 visits with more than 25,000 gallons of recycled water distributed .\n\n#recycledwater #bwp #burbankh2opower https://t.co/qzWgUW7zMJ
1564788076160876544	ManchesterNews8	#British official: people should give up being disgusted & drink treated sewage water.\nhttps://t.co/Hy1EXnPFPg
666387525003407360	albertvilarino	Did you know that today, there are more than 100K wastewater treatment plants worldwide. #water #technology https://t.co/kQt3mieQXt

4.2. Text preprocessing

Users of social media platforms have the freedom of expression, reason for which the content generated is noisy and contain symbols, Uniform resource locator (URL), hashtags, and repeated letters, so for us to get better accurate results, we included in this module the most used techniques in natural language pre-processing that involves, cleaning the text from URLs, extra whitespaces, numbers and screen names, then, substitute emojis by their text representation, replacing abbreviation (e.g., gr8 replaced by Great), along with handling contraction (e.g., don't become "do not" hasn't become "has not"). We also proceed by removing stop words (e.g., on, the, and it.). Finally, we lowercase and lemmatize the outputted text content.

4.3. Features extraction and fuzzy sentiment scoring

In the process of developing our model, we prepared in this stage the dataset and features required for us to examine the public's sentiment orientation toward recycled water. We utilized the cleaned text from the previous stage as input for word embeddings in order to represent text vocabulary and capture words relations, syntactics and semantic information [45], for this matter, we exploited an open-source python library called Gensim [46] and Word2Vec [47] algorithm to vectorize our text and use the results as input weight for our model. However, as word embeddings helps learning words semantics and relations but not sentiment, we proceed by incorporation a fuzzy sentiment scoring as a sentiment embedding of words, denoted st, that will add up to word embeddings, denoted et (i.e., $wt=st \oplus et$), for sentiment-oriented text representation.

For fuzzy sentiment scoring calculation for a given text input, we start by tagging each word using Part-of-Speech tagger, and for every opinionated word found, we look in the lexical resource SentiWordNet [48] for the initial polarity value $\mu(s)$. If not present, we seek in WordNet (i.e., a large dictionary database) to get the first matched synonym and look back in SentiWordNet for the associated value. This initial score value can be modified using the fuzzy functions mentioned earlier, whether by inverting the value, increment it, or by decrementing it based on the existence of a compliment, concentrator, or dilator hedge, respectively. The final fuzzy word sentiment embedding is concatenated with its corresponding word embedding is used as the LSTM model's final input.

4.4. Model training and sentiment analysis

As we mentioned earlier, we used the LSTM algorithm to tackle sentiment analysis. And we used the IMDB dataset (i.e., a well-known movie reviews dataset used for natural language processing) for our model training and testing processes. Initially, we cleaned this dataset based on the same pipeline from our framework, then we used the fuzzy functions along with SentiWordNet and WordNet to construct word sentiment embeddings. Forward we apply word embeddings using Word2Vec and pass the results as input to the LSTM model. After we experimented with the network hyperparameters using a trial-and-error process, we fixed the dimension of word embeddings to 300 and set the max_features to 3,000 as it provides a good variety of captured keywords, also we set a spacialDropout1D layer to 0.3 to avoid overfitting, we used Adam as the optimizer. We experimented also batches size of 32 and 64, and we fixed the value to 32 as it requires less memory and yield good performance, and we choose the accuracy metric for performance evaluation. We trained our model on 66% of the IMDB reviews dataset, and we set the number of epochs to 15 because we notice that after this number the validation accuracy began to decrease and the model overfit. With this tunning we got a final accuracy of 92.3%. Lastly, we applied our model to the dataset we collected from Twitter API related to recycled water subject and the results are shown in the section below.

5. RESULTS AND DISCUSSION
5.1. Spatial results of public participation in recycled water subject

Our conducted sentiment analysis on social media content reveals a significant disparity among countries in terms of public engagement and interaction with the topic of recycled water, as illustrated in Figure 4. To further investigate this, we compared our findings with the geolocation data of smart cities worldwide, as depicted in Figure 5. This comparison allows us to observe the correlation between public participation in the recycled water discourse and countries with smart city initiatives.

The United States topped the list of countries that had the most engagement with the topic of recycled water, followed by Canada, which can be attributed to the growing challenges of freshwater scarcity faced by the country [49]. India ranked next, while African countries were found to have limited participation in discussions about recycling water. The comparison of our world map results with the finding of [50] presented in Figure 5 where the author reveals the geo-location places of smart and digital cities, shows that countries with more smart city vision and advanced digital implementation. The citizens are relatively active regarding the subject of recycled water.
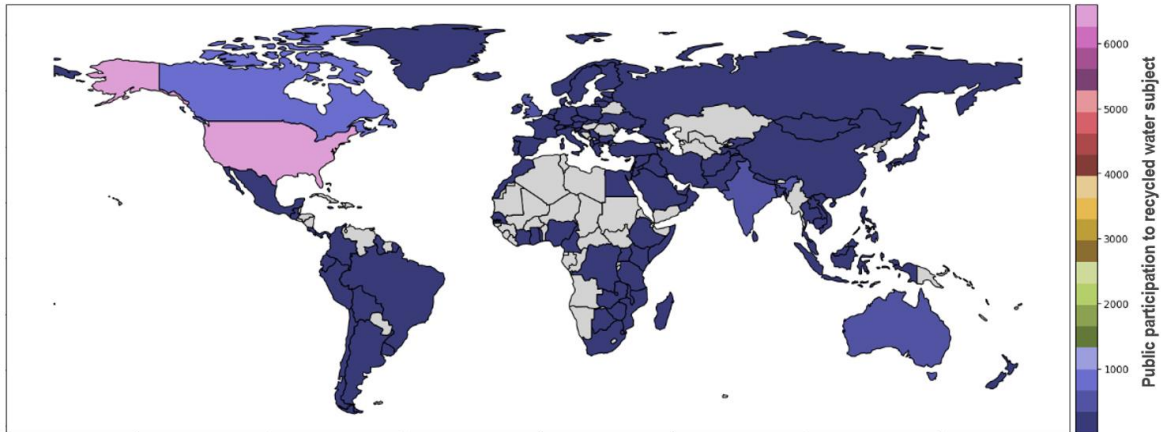
Figure 4. Results of public participation to recycled water

Figure 5. Digital and smart cities geo-location on the world map

5.2. Temporal results of public participation in recycled water subject

Utilizing data gathered from the Twitter platform spanning the years 2013 to 2022, we conducted an analysis of public engagement with the subject of recycled water. Our findings clearly indicate a substantial increase in public interest in this topic. This upward trend is visually evident in Figure 6, which illustrates the growing volume of publications and discussions related to recycled water over this extended timeframe.
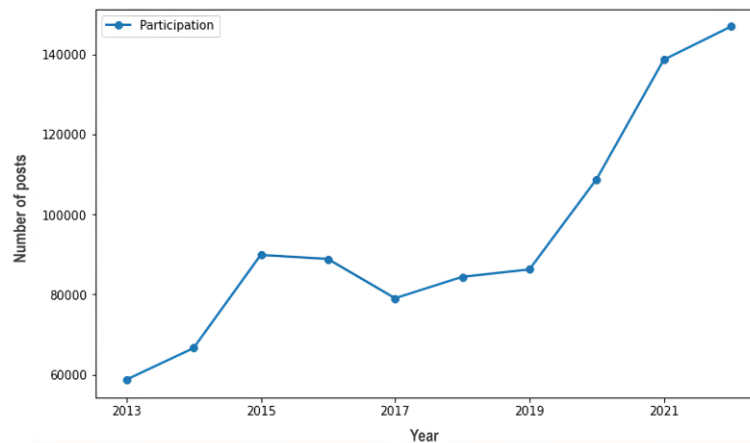
Figure 6. Public posts on Twitter regarding recycler water subject

5.3. Public sentiment orientation

The objective of our sentiment analysis in this paper is to extract and assess public sentiment toward recycled water. Additionally, we aim to gauge the public's willingness to engage in the development of their cities, aligning with the smart city vision for intelligent water management. Based on the results shown in Figure 7, we found that in countries with higher public participation in the recycled water subject, the number of posts expressing positive sentiment is relatively similar to the ones expressing negative sentiment. This suggests that the fear and negative opinions about this subject may not be as widespread as they appear. Additionally, as shown in Figure 8, we also noticed that the number of positive.
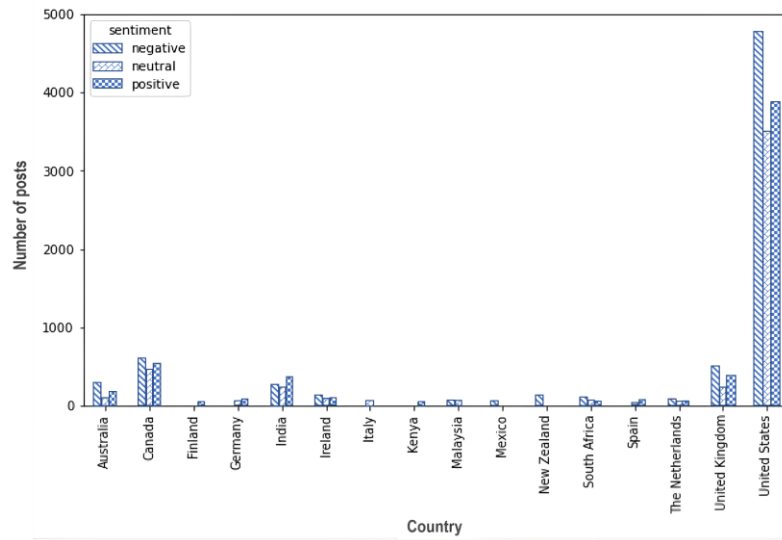
Figure 7. Public sentiment orientation by country
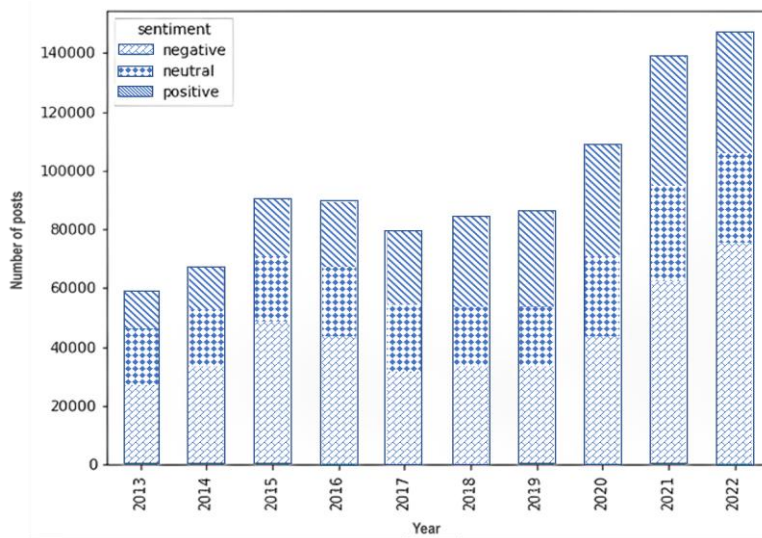
Figure 8. Public sentiment orientation by year

Sentiments towards recycled water have been on the rise over the years. In 2013, the percentage of positive posts among users was 21.85%, but by 2022, this percentage had increased to 27.87%. This represents a 6.02% increase between 2013 and 2022. These results suggest that people are becoming more aware of the challenges posed by water shortage, and the potential of recycled water to address this issue. Furthermore, it may serve as an incentive for governments and stakeholders to invest more in the execution of reclaimed water projects.

To better understand the public's perception of reclaimed water, we analyzed the most frequently used words in both positive and negative text content from our dataset. We found that the words "health", "worries", "safety", "radioactive", "disgust", "S**t", "Covid", and "worse" had the highest frequencies in negative text content. This suggests that the motivation behind recycled water disapproval by the public's is related mostly to the worries about safety, rather than disgust. In contrast, in positive text content, we found words such as "sensors", "enhance", "IoT", "data" and "forward" as shown in Figure 9, indicating that the public is optimistic and trusts the advancements in technology to improve recycled water solutions.

Figure 9. High frequency words from positive text content

6.   CONCLUSION

In our work, we focused on two main elements. Firstly, we aimed to enhance the sentiment analysis model we used in our previous papers by incorporating a domain ontology to collect domain-specific data, and using a combination of word embedding along with a sentiment word embedding. Also, we tried to give a step-by-step explanation of our framework to help future researchers who are concerned about utilizing fuzzy logic and deep learning for sentiment analysis. Secondly, we aimed through our paper to gain insight into the global public sentiment towards recycled water and its impact on citizen's acceptance and intension to use this new water resource, so we can help investors and governments direct their sensibilization campaigns towards removing the 'Yuck sentiment' barrier attached to recycled water and attain the smart water management goals.

Our findings reveal that public engagement is more pronounced in countries with advanced economic development and active smart city initiatives. Notably, we observed a rising trend in the percentage of positive sentiment related to recycled water over the years. Simultaneously, the growing concern among the public is oriented more towards the safety of this new resource in comparison to the disgust sentiment, which can highlight that overcoming the "yuck factor" barrier is possible through advancements in research, development, and the implementation of smart city visions for intelligent water management systems.

In our work, we encountered certain limitations, including the time-intensive nature of data collection from the Twitter API and model training. Additionally, we did not incorporate demographic information linked to Twitter user accounts, such as educational background, age, and gender, into our research. Including these details could enhance the analysis and provide deeper insights into public awareness of the potential of recycled water.

In our future research, we intend to delve into the technologies employed in the recycled water domain. Our goal is to assess whether citizens possess sufficient knowledge about these technologies, their processes, and their efficacy. Additionally, we plan to incorporate demographic data into our analysis of public acceptance of reclaimed water use. By doing so, we aim to enable governments and companies to tailor their awareness campaigns more effectively, targeting specific demographic groups and fostering increased investment in recycled water solutions.

REFERENCES

[1]   H. B. Tabish and C. H. Basch, "Back to the basics: Hand washing is public health 101 and it works to slow down the spread of viruses," *Infection, Disease and Health*, vol. 25, no. 4, pp. 319–320, Nov. 2020, doi: 10.1016/j.idh.2020.08.001.
[2]   R. K. Chakraborti, J. Kaur, and H. Kaur, "Water shortage challenges and a way forward in India," *Journal AWWA*, vol. 111, no. 5, pp. 42–49, May 2019, doi: 10.1002/awwa.1289.

[3]   V. Lazarova and T. Asano, "Challenges of sustainable irrigation with recycled water," in *Water Reuse for Irrigation*, CRC Press, 2004, pp. 1–30.
[4]   R. P. Dameri, "Using ICT in smart city," *Smart City Implementation: Creating Economic and Public Value in Innovative Urban Systems*, pp. 45–65, 2017.
[5]   H. Yeh, "The effects of successful ICT-based smart city services: From citizens' perspectives," *Government Information Quarterly*, vol. 34, no. 3, pp. 556–565, Sep. 2017, doi: 10.1016/j.giq.2017.05.001.
[6]   A. Taouraout, A. Chahlaoui, D. Belghyti, M. Najy, and R. Sammoudi, "How to improve wastewater treatment in smart city," in *Innovations in Smart Cities Applications Edition 3*, Springer International Publishing, 2020, pp. 1141–1153.
[7]   J. Yang, M. Monnot, L. Ercolei, and P. Moulin, "Membrane-based processes used in municipal wastewater treatment for water reuse: state-of-the-art and performance analysis," *Membranes*, vol. 10, no. 6, Jun. 2020, doi: 10.3390/membranes10060131.
[8]   M. Zouaoui, G. Nouyrigat, and F. Beer, "How does investor sentiment affect stock market crises? evidence from panel data," *Financial Review*, vol. 46, no. 4, pp. 723–747, Nov. 2011, doi: 10.1111/j.1540-6288.2011.00318.x.
[9]   H. K. Sul, A. R. Dennis, and L. (Ivy) Yuan, "Trading on Twitter: using social media sentiment to predict stock returns," *Decision Sciences*, vol. 48, no. 3, pp. 454–488, Jun. 2017, doi: 10.1111/deci.12229.
[10] B. Mohamed, H. Haytam, and F. Abdelhadi, "Applying fuzzy logic and neural network in sentiment analysis for fake news detection: case of Covid-19," in *Studies in Computational Intelligence*, Springer International Publishing, 2022, pp. 387–400.
[11] K. Indhuja and R. P. C. Reghu, "Fuzzy logic based sentiment analysis of product review documents," in *2014 First International Conference on Computational Systems and Communications (ICCSC)*, Dec. 2014, pp. 18–22, doi: 10.1109/COMPSC.2014.7032613.
[12] D. Elliott, "Renewable energy and sustainable futures," *Futures*, vol. 32, no. 3–4, pp. 261–274, Apr. 2000, doi: 10.1016/S0016-3287(99)00096-8.
[13] V. Srinivasan, E. F. Lambin, S. M. Gorelick, B. H. Thompson, and S. Rozelle, "The nature and causes of the global water crisis: Syndromes from a meta-analysis of coupled human-water studies," *Water Resources Research*, vol. 48, no. 10, Oct. 2012, doi: 10.1029/2011WR011087.
[14] G. Mancuso, S. Lavrnić, and A. Toscano, "Reclaimed water to face agricultural water scarcity in the Mediterranean area: An overview using sustainable development goals preliminary data," in *Wastewater treatment and Reuse Present and future perspectives in technological developments and management issues*, Elsevier, 2020, pp. 113–143.
[15] A. P. Rao, J. Patel, and A. K. Pradhan, "Application of alternative sources of water in agricultural food production-current trends and future prospects," *Current Opinion in Food Science*, vol. 47, Oct. 2022, doi: 10.1016/j.cofs.2022.100877.
[16] H. Sharif Nia *et al.*, "A predictive study between anxiety and fear of COVID-19 with psychological behavior response: the mediation role of perceived stress," *Frontiers in Psychiatry*, vol. 13, Mar. 2022, doi: 10.3389/fpsyt.2022.851212.
[17] C. A. Harper, L. P. Satchell, D. Fido, and R. D. Latzman, "Functional fear predicts public health compliance in the COVID-19 pandemic," *International Journal of Mental Health and Addiction*, vol. 19, no. 5, pp. 1875–1888, Oct. 2021, doi: 10.1007/s11469-020-00281-5.
[18] D. Vakula and Y. K. Kolli, "Waste water management for smart cities," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, Dec. 2017, pp. 275–279, doi: 10.1109/ISS1.2017.8389414.
[19] E. Affuso and K. D. Lahtinen, "Social media sentiment and market behavior," *Empirical Economics*, vol. 57, no. 1, pp. 105–127, Jul. 2019, doi: 10.1007/s00181-018-1430-y.
[20] P. R. Nkhoma, K. Alsharif, E. Ananga, M. Eduful, and M. Acheampong, "Recycled water reuse: what factors affect public acceptance?," *Environmental Conservation*, vol. 48, no. 4, pp. 278–286, Dec. 2021, doi: 10.1017/S037689292100031X.
[21] L. Garcia-Cuerva, E. Z. Berglund, and A. R. Binder, "Public perceptions of water shortages, conservation behaviors, and support for water reuse in the U.S.," *Resources, Conservation and Recycling*, vol. 113, pp. 106–115, Oct. 2016, doi: 10.1016/j.resconrec.2016.06.006.
[22] P. Rozin, B. Haddad, C. Nemeroff, and P. Slovic, "Psychological aspects of the rejection of recycled water: Contamination, purification and disgust," *Judgment and Decision Making*, vol. 10, no. 1, pp. 50–63, Jan. 2015, doi: 10.1017/S193029750000317X.
[23] C. Leong, "The role of emotions in drinking recycled water," *Water*, vol. 8, no. 11, Nov. 2016, doi: 10.3390/w8110548.
[24] L. Li, X. Liu, and X. Zhang, "Public attention and sentiment of recycled water: Evidence from social media text mining in China," *Journal of Cleaner Production*, vol. 303, Jun. 2021, doi: 10.1016/j.jclepro.2021.126814.
[25] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.
[26] S. Malviya, A. K. Tiwari, R. Srivastava, and V. Tiwari, "Machine learning techniques for sentiment analysis: A review," *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, vol. 12, no. 2, pp. 72–78, 2020.
[27] P. Baid, A. Gupta, and N. Chaplot, "Sentiment analysis of movie reviews using machine learning techniques," *International Journal of Computer Applications*, vol. 179, no. 7, pp. 45–49, Dec. 2017, doi: 10.5120/ijca2017916005.
[28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[29] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 4, Jul. 2018, doi: 10.1002/widm.1253.
[30] G. S. N. Murthy, S. R. Allu, B. Andhavarapu, M. Bagadi, and M. Belusonti, "Text based sentiment analysis using LSTM," *International Journal of Engineering Research and*, vol. 9, no. 5, May 2020, doi: 10.17577/IJERTV9IS050290.
[31] B. Naderalvojoud and E. A. Sezer, "Sentiment aware word embeddings using refinement and senti-contextualized learning approach," *Neurocomputing*, vol. 405, pp. 149–160, Sep. 2020, doi: 10.1016/j.neucom.2020.03.094.
[32] L.-C. Yu, J. Wang, K. R. Lai, and X. Zhang, "Refining word embeddings using intensity scores for sentiment analysis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 671–681, Mar. 2018, doi: 10.1109/TASLP.2017.2788182.
[33] X. Fu, J. Yang, J. Li, M. Fang, and H. Wang, "Lexicon-enhanced LSTM with attention for general sentiment analysis," *IEEE Access*, vol. 6, pp. 71884–71891, 2018, doi: 10.1109/ACCESS.2018.2878425.
[34] K. Howells and A. Ertugan, "Applying fuzzy logic for sentiment analysis of social media network data in marketing," *Procedia Computer Science*, vol. 120, pp. 664–670, 2017, doi: 10.1016/j.procs.2017.11.293.
[35] P. Wongthongtham and B. A. Salih, "Ontology-based approach for identifying the credibility domain in social big data," *Journal of Organizational Computing and Electronic Commerce*, vol. 28, no. 4, pp. 354–377, Oct. 2018, doi: 10.1080/10919392.2018.1517481.
[36] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, Jun. 1993, doi: 10.1006/knac.1993.1008.

[37] X. Fang and J. Zhan, "Sentiment analysis using product review data," *Journal of Big Data*, vol. 2, no. 1, Dec. 2015, doi: 10.1186/s40537-015-0015-2.

[38] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.

[39] L. A. Zadeh, "Fuzzy logic-a personal perspective," *Fuzzy Sets and Systems*, vol. 281, pp. 4–20, Dec. 2015, doi: 10.1016/j.fss.2015.05.009.

[40] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning-I," *Information Sciences*, vol. 8, no. 3, pp. 199–249, 1975, doi: 10.1016/0020-0255(75)90036-5.

[41] L. A. Zadeh, "A fuzzy-set-theoretic interpretation of linguistic hedges," *Journal of Cybernetics*, vol. 2, no. 3, pp. 4–34, Jan. 1972, doi: 10.1080/01969727208542910.

[42] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[43] R. Grosse, "Lecture 15: Exploding and vanishing gradients," *University of Toronto Computer Science*, 2017.

[44] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019, doi: 10.1162/neco_a_01199.

[45] F. Almeida and G. Xexéo, "Word embeddings: A survey," *arXiv preprint arXiv:1901.09069*, Jan. 2019.

[46] R. Řehůřek and P. Sojka, "Gensim: Topic modelling for humans." https://radimrehurek.com/gensim/ (accessed Dec. 23, 2022).

[47] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, Jan. 2013.

[48] S. Baccianella, A. Esuli, and F. Sebastiani, "Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining," in *Lrec*, 2010, pp. 2200–2204.

[49] K. Banks, "Canada's troubled waters," 2021. https://www.universityaffairs.ca/features/feature-article/canadas-troubled-waters/ (accessed Dec. 29, 2022).

[50] A. Agha, "A stakeholder based assessment of developing country challenges and solutions in smart mobility within the smart city framework: a case of lahore," *University of Cambridge: Cambridge, UK*, 2016.

BIOGRAPHIES OF AUTHORS

Mohamed Bahra 🆔 📚 SC ⬡ a software engineer and currently in the last year as a Ph.D student, his field of research is related to big data, smart city, and data science. He teached data science modules as a temporary teacher during his period of Ph.D. research, also he worked as a software engineer at the Ministry of Interior of Morocco, and he supervised students for the realization of their end-of-study projects at the Computer Science Department of Abdelmalek Essadi University, Morocco. His skills are related to sentiment analysis, information extraction, social network analysis, semantic web, ontologies, natural language processing, machine and deep learning. He can be contacted at email: bahra002@gmail.com and mohamed.bahra@etu.uae.ac.ma. His profile can be found at https://www.researchgate.net/profile/Mohamed-Bahra.

Abdelhadi Fennan 🆔 📚 SC ⬡ he is currently works at the Computer Science Department of Abdelmalek Essaâdi University, Morocco. Abdelhadi does research in artificial intelligence, information systems, and software engineering. His disciplines are related to artificial intelligence, software engineering, information systems, higher education, educational technology, curriculum theory, and his skills are of the fields of e-learning, online learning, software engineering, artificial intelligence, semantic web, sentiment analysis, homogenization, technology enhanced learning, and competitive intelligence. He can be contacted at email: afennan@gmail.com. His profile can be found at https://www.researchgate.net/profile/Abdelhadi-Fennan.