INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

WILEY

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Intl. Trans. in Op. Res. 0 (2024) 1–32
DOI: 10.1111/itor.13568

Mood and emotion assessment for risk reduction of pandemic spread through passenger air transport: a DSS applied to the COVID-19 in the case of Spain

J. Aguarón[a,*], A. Altuzarra[a], R. Aznar[b], M.T. Escobar[a], A. Jiménez-Martín[c], A. Mateos[c], A. Moreno-Díaz[c], J.M. Moreno-Jiménez[a], C. Moreno-Loscertales[d], V. Muerza[a], J. Navarro[a], A. Sarango[c], A. Turón[a] and L.G. Vargas[e]

[a]*Grupo Decisión Multicriterio Zaragoza (GDMZ), Departamento de Economía Aplicada, Facultad de Economía y Empresa, Universidad de Zaragoza, Gran Vía, 2. 50005-Zaragoza, Spain*
[b]*Medicina Familiar y Comunitaria, Salud Aragón., Ejea, Zaragoza, Spain*
[c]*Grupo de Análisis de Decisiones y Estadística (GADE), Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Campus de Montegancedo S/N, Boadilla del Monte 28660, Spain*
[d]*Instituto de Investigación Sanitaria de Aragón, Servicio de Neurología del Hospital Miguel Servet, Zaragoza, Spain*
[e]*Katz Graduate School of Business, University of Pittsburgh, Pittsburgh, PA 15260, USA*
E-mail: aguaron@unizar.es[Aguarón]; altuzarr@unizar.es[Altuzarra]; raznarsa@salud.aragon.es[Aznar]; mescobar@unizar.es[Escobar]; antonio.jimenez@upm.es[Jiménez-Martín]; amateos@fi.upm.es[Mateos]; amoreno@fi.upm.es[Moreno-Díaz]; moreno@unizar.es[Moreno-Jiménez]; cmoreno@salud.aragon.es[Moreno-Loscertales]; vmuerza@unizar.es[Muerza]; jnavarrol@unizar.es[Navarro]; ayrton.sarango@alumnos.upm.es[Sarango]; turon@unizar.es[Turón]; vargas@katz.pitt.edu[Vargas]

Received 28 February 2024; received in revised form 26 August 2024; accepted 11 October 2024

Abstract

This paper presents a decision support system (DSS) for sentiment analysis of Spanish texts based on lexicons. The information provided by this DSS, named Spanish Sentiment Analysis-DSS (SSA-DSS), is employed to assess the social impacts considered in an external software module (RRPS-PAT) centered on risk reduction of pandemic spread through passenger air transport. RRPS-PAT is a complex multiobjective optimization module simultaneously addressing different conflicting objectives, including epidemiological, economic, and social aspects. This allows more effective and realistic decisions to be made. The specificity and novelty of the problem suggest the use of lexicon-based approaches because there is no prior information about the problem to train machine learning–based approaches. The SSA-DSS covers the entire process from the incorporation of texts, particularly tweets, to be analyzed, the application of preprocessing and cleaning tools, the selection of lexicons (general, context, and emoji lexicons) to be used and their possible modification, to the visualization of results and their exportation to other software tools. This paper contemplates, apart from the RRPS-PAT module, the connection with a social network analysis tool (Gephi) that complements the information provided by SSA-DSS with the identification of social leaders. The usefulness

*Corresponding author.

and functionalities of SSA-DSS are illustrated by means of an example related to the evolution of societal mood in Spain during the COVID-19 pandemic.

Keywords: decision support system; risk reduction; pandemic spread; sentiment analysis; lexicon

1. Introduction

Passenger air transport is one, if not the most, important pathway for the rapid spread of pandemics, as highlighted by the COVID-19 pandemic (Vicente and Mateos, 2020; Wu et al., 2020; Pang et al., 2021). To reduce the negative effects that pandemics have on society, it is necessary to develop analytic and Information and Communication Technology (ICT) tools that make it possible to (i) predict the potential risks arising from pandemics for the different actors involved (society, the general public, airlines, workers providing the service, etc.); (ii) build early warning systems that detect the appearance of pandemics as quickly as possible; (iii) establish immediate response procedures that prevent, or at least help to control, the spread of pandemics; (iv) set up communication and collaboration mechanisms that extract the knowledge derived from the crisis situation for its storage and future exploitation; and (v) formulate actions that promptly revert or improve the situation wherever possible.

In consideration of the first point (i), the global report published by the World Health Organization (World Health Organization, 2022) explores potential future pandemics and epidemics, analyzing risks and providing recommendations to enhance global preparedness and response. Chen and Wang (2023, 2024) discuss the effectiveness of air travel policies in responding to emerging infectious disease pandemics. Lee et al. (2023) review tools, models, and frameworks for infectious disease emergency preparedness, focusing on COVID-19. They provide a detailed analysis of literature since 2020 and offer guidance on pandemic preparedness and response. Zhang et al. (2024) introduce a machine learning–based tool for predicting outbreak risks in various contexts, using diverse data to provide early alerts and enhance risk mitigation during pandemics.

Regarding the second point (ii), a great deal of research on the construction of early warning systems has been published in the scientific literature. *Dynamic networks* (Chen et al., 2020; Dong et al., 2021) and *network analysis* (Watts and Strogatz, 1998; So et al., 2020; de Souza et al., 2021) are two of the most popular methodologies proposed for the detection of early warning signals based on confirmed daily cases of COVID-19 and air traffic information. A comparative analysis carried out in a European context using data from the World Health Organization and Flightradar24 between February 15 to May 1, 2020, concluded that various combinations of network analysis–based methods with different adjacency matrices (including air traffic) provide good results and are able to predict the evolution of the pandemic around 20 days in advance based on air passenger traffic (Fragua et al., 2023).

This paper focuses on immediate response air traffic management procedures that prevent, or at least help to control, the spread of pandemic point (iii) and on mechanisms for extracting knowledge from the crisis point (iv). In this context, our main objective will be the study and evaluation of social mood within the social impacts contemplated in reducing the risk of pandemic spread through passenger air transport. The strategy chosen to reduce this risk is the closure of airports or connections between airports.

Air traffic management can help to prevent and control the worldwide spread of pandemics, especially during the early stages. A complete shutdown of air traffic could slow the spread, but the global impacts of such a measure are potentially enormous, both economically and socially. During the COVID-19 pandemic, different countries temporarily suspended flights to and from other countries depending on the status of the pandemic in those countries. However, pandemic incidence rates are not homogeneous across countries, and these decisions may cause regions with low epidemiological incidence rates to be negatively affected.

Therefore, it would appear to be necessary to make intelligent decisions at a lower level (air connections between airports) in order to integrate the epidemiological risk of importing the pandemic and the economic and social impacts associated with the decisions taken on air traffic into our formal models. This is a multiobjective optimization problem with conflicting objectives.

This problem was first addressed in Jiménez-Martín et al. (2023), describing a methodology for assessing the risk of pandemics imported via airports due to international air traffic. The economic impacts considered are losses caused to the airlines operating the canceled flights, the destination airports (loss of airport taxes), and the regions in which the airports are located (loss of tourism and business travel). With social impacts, account is taken of the loss of connectivity within the international air transport network and the percentage of passengers affected because of canceled traffic. Both social impacts were taken as objectives measured in percentage terms. As far as we know, no previous work related to the influence of social mood on these decisions has been published. The analysis of sentiment that the closure of air connections causes in the population is an innovative and relevant idea that improves the acceptance and effectiveness of policies and contributes to a more holistic and humane management of the crisis. The gap in the literature and the development of a software tool to support decision-making and knowledge extraction in this context, are the main issues contemplated in this work.

Before presenting the DSS developed for evaluating social mood, Section 2 introduces the software module for risk reduction of pandemic spread through passenger air transport (RRPS-PAT). RRPS-PAT (Peña et al., 2024) is a complex multiobjective optimization module that simultaneously considers conflicting objectives, including epidemiological, economic, and social aspects. The functionalities provided by this software include the quantification of these and other impacts that have been introduced into the analysis of the problem, the incorporation of decision-maker preferences, optimization problem solving by means of metaheuristics, and the visualization and sharing of the results.

The RRPS-PAT module will be integrated with the SSA-DSS (Spanish Sentiment Analysis-Decision Support System) presented in Section 4. This DSS provides information about the mood and emotions (sentiment analysis) of society and its influence on the social impacts considered in the RRPS-PAT. Once a decision has been made, the RRPS-PAT software calls the SSA-DSS to monitor the tweets from that moment onwards. The analysis provides a valence reflecting the social impact of the decision and this information is sent back to the RRPS-PAT software when a new decision needs to be taken.

The consideration of social impacts assessed through affective aspects (mood sentiments and emotions), particularly the possibility of understanding and incorporating the mood of the population regarding possible air traffic opening/closure measures into the decision-making process, is a recent subject for the scientific literature, and it is, therefore, one of the key issues of the proposal. Although emotions and mood have been examined in other fields such as marketing, finance, and crisis management (Kaiser and Kuckertz, 2023), we are not aware that this has been done in air traffic management during pandemics nor is there a DSS specifically designed for this purpose.

The relevance of incorporating this objective is due to (i) social acceptance decisions affecting mobility and connectivity have a great emotional impact on the population and the inclusion of feelings increases the acceptance of the decisions taken; (ii) effective communication and crisis management understanding the feelings of the population helps design more effective communication strategies, reducing panic and resistance which translates into transparency and empathy; (iii) psychological impact and well-being decisions that minimize the negative impact on mental health and emotional well-being contribute to a faster and more sustained recovery.

The study of the evolution of the mood (sentiment) of society and the citizens' emotions, through the analysis of the content of texts (from different sources, such as social networks, newspapers, and so on), including audio recordings transcribed into texts, makes it possible to promptly identify moments of joy and sadness, trust and disgust, anticipation and surprise, and anger and fear; this helps to determine the most appropriate time to introduce measures that are not usually liked by citizens.

Methodologically, sentiment analysis of social media involves a series of approaches that extract opinions, feelings, and thoughts from users by means of natural language processing (NLP) procedures (Hasan et al., 2019). There are two main approaches in sentiment analysis (Verma and Thakur, 2018; Catelli et al., 2022): The first is based on machine learning techniques, in particular large language models (LLMs), and the second is based on lexicons. The first approach is more efficient but requires prior information to train the machine learning models. The second is less efficient and is usually recommended in novel situations for which prior information is not available.

Due to its (i) simplicity and speed, (ii) interpretability, (iii) consistency, (iv) efficiency, (v) adaptability, (vi) facility of using and updating, and, finally, (vii) the specificity of the considered problem (incorporate social mood when measuring social impacts in pandemic spread through passenger air transport) and the inexistence of prior information to train LLM methodologies, the sentiment analysis based on lexicon approach will be examined in this work.

The sentiment analysis of texts based on lexicons has been recently considered within the field of pandemic management (Manguri et al., 2020; Garcia and Berton, 2021; Turón et al., 2023; Navarro et al., 2024). It has been applied to the study of the evolution of the mood of Spanish society in response to different public policies and health decisions made by respective governments (national or regional) in relation to COVID-19 vaccination in Spain (Turón et al., 2023). In this case, by analyzing comments on Twitter (now X), it was possible to identify changes in trends in the sentiment of the population at different moments in time. A similar evolution of mood has been obtained by analyzing tweets issued exclusively by social leaders (and their Twitter Communities) identified according to their influence on the rest of the participants in the debate (Navarro et al., 2024).

The effectiveness of topic identification and sentiment analysis applied to COVID-19 tweets was also compared and discussed in two different languages (English and Portuguese) in Manguri et al. (2020). A sentiment analysis applied to the week of greatest worldwide spread of COVID-19 (April 9 to 14, 2020) revealed that people's daily reactions are clearly influenced by the tweets of individuals, government organizations, and communication agencies (Garcia and Berton, 2021).

Other applications of sentiment analysis based on lexicons to study pandemics are in publications related to the Middle East respiratory syndrome (Do et al., 2016), influenza A (Jurek-Loughrey et al., 2015; Wang and Vergeer, 2024), pandemics events (Qin and Ronchieri, 2022), HIV (Lohmann et al., 2018), and e-cigarette/vaping associated lung injury (EVALI) outbreaks (Kasson et al., 2021).

Sentiment analysis has been applied to (nonpandemic) disaster management: Navarro et al. (2023) analyzed press articles published during 2021 on the eruption of the Cumbre Vieja volcano (La Palma, Canary Islands) and public decision-making over different temporal horizons. The paper also includes the design of improved strategies for disaster risks reduction (DRR). Further DRR works refer to: disaster response and recovery (Ragini et al., 2018); the Syrian chemical attack (Bashir et al., 2021); risk detection through crisis information (Rexiline Ragini et al., 2018); and, the social impact of Hurricane Dorian (Badmus, 2020).

As already mentioned, the present study focuses on a specific disaster: the spread of pandemics through air transport. The work analyzes the influence of the mood of the population on the measurement of social impacts and exploitation in decision-making to reduce or mitigate the risk of the spread of the pandemic. Despite the specificity of the problem (closure of connections at airports), many of the ideas put forward can be extended and/or applied in a natural manner to other disasters.

The last point (v), of the five used for establishing the analytical and technological tools necessary to reduce negative effects of pandemics in RRPS-PAT, centers on the formulation of "actions that promptly revert or improve the situation wherever possible." As in other works in other domains (Gao et al., 2020; Cinelli et al., 2020; Samuel et al., 2020), the initial proposal of the RRPS-PAT software has now incorporated sentiment analysis to help authorities design more effective public policies. In addition to the previous uses, the knowledge extracted when analyzing the state of mind of society (Section 2) helps to set the minimum and maximum connectivity thresholds proposed by the political authorities and to plan public air transport interventions in light of social concerns.

After introducing some basic ideas about the relevance of affective aspects in decision-making, this paper describes the integration of the risk analysis module (RRPS-PAT) with the SSA-DSS and details how this software operates. The SSA-DSS was initially designed for the analysis of tweets using Spanish lexicons, but it can be applied to any other types of texts (messages, comments on messages, news in the press, technical or specialized reports, and so on). This tool imports, preprocesses, prepares, and analyzes tweets using lexicons. Additionally, this software includes different tools for the visualization and exportation of the results, and their analysis (employing Gephi, open-source and free visualization and exploration software for all kinds of graphs and networks; https://gephi.org), which aims to identify the social leaders and their communities.

The remainder of the paper is structured as follows: Section 2 provides additional information on the decision support system for air traffic management to control the spread of pandemics; Section 3 introduces some basic ideas on the relevance of affective aspects in decision-making; Section 4 describes the main features of the SSA-DSS, its implementation, and integration with RRPS-PAT (Subsection 4.1), it further presents the SSA-DSS interface and its connection with Gephi (Subsection 4.2); Section 5 includes an illustrative case study of the SSA-DSS and the evolution of the social mood in Spain during the COVID-19 pandemic; finally, Section 6 discusses relevant conclusions.

2. Key aspects for controlling the spread of pandemics through passenger air traffic

The problem of the selective opening/closure of passenger air traffic to control the spread of pandemics is a multiobjective optimization problem, the epidemiological, economic, and social impacts associated with the decisions taken on air traffic must be simultaneously dealt with. As these are usually conflicting criteria, a compromise on their achievement levels must be sought. This problem was first addressed in Jiménez-Martín et al. (2023), where the use of the NSGA-III metaheuristic was proposed to develop an approximation of the Pareto frontier, then expert preferences were incorporated to derive a compromise solution. The proposed methodology was illustrated at a national (Spain) level accounting for 1196 flights corresponding to 569 international connections arriving at the 38 Spanish airports on September 24, 2020. The quantification of the epidemiological, economic, and social impacts was improved by the authors, and new impacts were included.

In terms of the epidemiological impact, a methodology to quantify the risk of imported pandemics at airports was first proposed by García-Moreno et al. (2021) based on the susceptible-infectious-recovered (SIR) method (Kermack and Mckendrick, 1927) to quantify the risk of in-flight virus transmission. This model was later improved by means of (i) the incorporation of the risk of virus transmission at airports based on deep learning and computer vision techniques (Rodríguez-Escabias, 2023); (ii) a more refined computation of the probability of passengers arriving at an airport being infected based on population mobility information provided by the Spanish Ministry of Transport, Mobility, and Urban Agenda sourced from the geolocation of mobile terminals (Carbonell, 2022); and (iii) the use of the different variants of the SIR method, such as the susceptible-exposed-infectious-recovered (SEIR), susceptible-exposed-infectious-recovered-susceptible (SEIRS) (Wang et al., 2020), susceptible-infected-susceptible (Ceria et al., 2021), or adaptive SIRS (aSIRS). Information on global pandemic incidence from John Hopkins University was used to identify the incidence rate in the catchment area of departure airports.

Regarding economic impact, airlines, airports, and the catchment area of the destination airports are taken into account. First, the loss of airline revenues due to flight cancellations is considered. To do this, flight information from Flightradar and aircraft occupancy data from the International Transport Association (IATA) were used. The percentage of passengers affected by the cancellations of each airline should be minimized, as should the standard deviation of these percentages, in order to uniformly distribute, insofar as possible, the damage across airlines. Airports also suffer economic losses as a result of the closure of air traffic. Specifically, the destination airports lose the airport taxes for canceled flights. These taxes depend on the aircraft's weight, with some airports increasing the amount to be paid depending on how loud the aircraft is and the time of landing. This information is available at www.aena.es/es/aerolineas/tarifas.html. The loss of sales at airport stores and restaurants has not been considered, since revenue from airport sales is mainly associated with departure airports where passengers spend more time and is negligible at destination airports. Finally, the catchment areas of the destination airports are also affected by flight cancellations. Specifically, many of the passengers on the canceled flights are traveling for tourism or business, which would have resulted in accommodation and catering revenues in the region in which the airport is located. Information from the Spanish Institute of Statistics (INE) on the average length of stay and average daily expenditure of both tourist and business visitors by country of origin has been taken into account to quantify this impact. This impact on the catchment area of the destination airports was incorporated into the problem-solving procedure by Peña et al. (2024). The damage has been also uniformly distributed both across airports and catchment areas, respectively, by minimizing the respective standard deviations of the individual impacts.

Regarding the social impact, the loss of connectivity to airports as a result of flight cancellations (which can lead to missed connections) was measured by Jiménez-Martín et al. (2023). To do this, the percentage of lost connectivity was multiplied by the connectivity index of the departure airport defined in IATA's Air Connectivity Document (2020). The percentage connectivity losses of all airports are then added together. In Peña et al. (2024), the so-called WELLBY concept was proposed for incorporation into the analysis. WELLBY is a subjective measure of the well-being of individuals (Frijters and Krekel, 2021). WELLBYs are cardinal measures. Therefore, every country can aggregate and compare its scores, even at different moments in time. The formula and methodological details on how this measure is reported in Frijters and Sanchis (2024), and the WELLBY World Database was recently created and updated by means of Economics @Intelligence (Economics @Intelligence, 2022). Finally, some sociopolitical decisions could restrict the problem-solving space under consideration. For example, a minimum level of air traffic could be established with certain EU countries or capitals or the rest of the world in the period under analysis. These types of decisions can be easily modeled and incorporated into the problem-solving procedure, as explained in Peña et al. (2024). The scope of measuring social impact includes an accurate assessment of the loss of air connectivity and a comprehensive understanding of individual well-being through WELLBY. However, limitations include dependence on accurate data, variability in well-being perception, potential biases in regulatory decisions, and the need for adaptable models. Comprehensive and continuous evaluation is crucial to address these challenges and improve the management of the social impact of disruptions in air connectivity.

All the information involved in the epidemiological, economic, and social impacts has been incorporated from the respective sources into a Neo4J graph-oriented database. Besides, a decision support system for risk reduction of pandemic spread in passenger air transport (RRPS-PAT) was developed in order to define the specific scenario for analysis (time period, the percentage of targeted risk reduction, political constraints, and decision-maker preference elicitation) and to visualize the results as maps and tables. NSGA-III was replaced by the *binary particle swarm optimization* (BPSO) metaheuristic in Peña et al. (2024), an extension of the traditional *particle swarm optimization* algorithm to binary problems (Nguyen et al., 2021), inspired by the social behavior of flocks of birds and schools of fish, to solve the corresponding multiobjective optimization problem.

However, as cited above, the initial version of RRPS-PAT does not account for emotional aspects of the decision-making problem, only objective impacts have been incorporated into this software tool. To resolve this limitation, the SSA-DSS has been integrated into the software. The new version of RRPS-PAT adds the emotional impact of decisions made about air traffic on the population by analyzing information from social networks or comments on related news in newspapers. In this way, given that decisions on air traffic will be made periodically, we will be able to take into account the impact of periodic air traffic decision-making on the mood of the population.

In addition to setting up connectivity thresholds proposed by political authorities, the explicit consideration of citizen sentiment is, as is the case when incorporating the opinion of all relevant actors in the resolution of a problem, a fundamental necessity for the social acceptance of the proposed measures and citizen collaboration in implementation.

3. Reason and emotion in decision-making: sentiment analysis

Decision-making is one of the essential activities of the human species and reflects its degree of evolution, knowledge, and freedom (Moreno-Jiménez and Vargas, 2018). There has been much debate about the dilemma between the two types of decisions traditionally associated with human behavior: analytical and intuitive.

Rational or analytical decisions, made with the left side of the brain, reflect the characteristics of the scientific method in each era (Vargas et al., 2023; Arduin, 2023). They tend to be slow to take and execute because they (i) systematically analyze the existing (internal and external) information, the alternatives, and their consequences according to different indicators, and so on; (ii) establish one or more criteria to guide decision-making; (iii) determine the problem-solving process followed for each type of considered problem; and (iv) validate and calibrate the formal model before its practical application. In this context, the problems can be classified (Ishizaka and Nemery, 2013; Zamarrón-Mieza et al., 2017) as the choice of the best alternative ($P.\alpha$ problems) or the k-best alternatives ($P.\alpha[k]$ problems); identification and assignment to groups ($P.\beta$ problems), including classification and segmentation; ranking of the alternatives ($P.\gamma$ problems); description of the problem ($P.\delta$ problems) and extraction of knowledge from the resolution process ($P.\kappa$ problems).

Emotional or intuitive decisions, made by the right side of the brain, reflect the evolution of living systems up until the current *Homo sapiens sapiens*. They are quick to take and execute in order to guarantee the subsistence of the species. Intuition guides decisions aligning them with our authentic selves. Intuitive decisions use knowledge beyond logical reasoning.

Traditionally, science only considered the objective and tangible associated with the left side of our brain, leaving aside the subjective and intangible associated with the right side. As a result of the work of knowledge psychologists and behaviorists, as of the last quarter of the 20th century, the importance of affective aspects (mood, emotions, attitude, personality, and motivation) in decision-making processes has begun to become clear (Lins et al., 2023). Furthermore, it is being highlighted that the best decisions are made by combining the rational and emotional parts of our brains (Kahneman, 2011).

This paper focuses on the analysis of mood and emotions in Spanish texts using lexicons. The Merriam-Webster dictionary defines mood as a conscious state of mind or predominant emotion. According to Antonio Damasio, emotions are instinctive reactions of our body, determined by environmental stimuli and derived from the development of biological regulation (Damasio, 1996). Emotions arise as a response from the unconscious mind and offer valuable information about our inner state and perception of the world. They can be classified according to their intensity, from most to least intense, as primary, secondary, and tertiary. Robert Plutchik (Mohsin and Beltiukov, 2019) separated emotions into eight broad and opposing (positive and negative) categories (joy and sadness, anticipation and surprise, trust and disgust, and anger and fear).

Automatic text processing to identify and categorize opinions and attitudes (sentiment analysis) can be carried out by means of two different approaches (Balazs and Velásquez, 2016): the unsupervised lexicon-based approach, based on semantic knowledge of words; and the supervised learning–based approach, which relies on statistical techniques. There is a third, more recent strategy, known as the concept-based approach, whose core are ontologies.

Lexicon-based models are preferred when the datasets are small, and the available computational resources are limited (Catelli et al., 2022). Supervised learning models perform well for the specific domain on which they have been trained but have important limitations for analyzing different domains or new topics. A combination of both techniques (hybrids: lexicon and machine learning–based models) improves the classification, detection, and measurement of sentiment at the concept level, providing highly accurate results (D'Andrea et al., 2015). In our case, to improve the effectiveness of the text analysis, the information obtained using the lexicon-based approach will be employed in the near future to train LLM methodologies (hybrid models).

4. The Spanish Sentiment Analysis-Decision Support System

4.1. Presentation of the SSA-DSS

The SSA-DSS is a Spanish lexicon-based sentiment analysis tool which was developed as part of the multidisciplinary project "Participación Ciudadana Cognitiva y Decisiones Públicas. Aplicaciones Sociosanitarias" (Ref.: LMP21-35), funded by the Government of Aragon.

The tool originally focused on the evolution of the mood of citizens (Turón et al., 2023) and the influence of social leaders (Navarro et al., 2024) during the COVID vaccination program in Spain. A three-step methodological procedure (Navarro et al., 2024) was employed: (i) corpus determination (data collection, data preprocessing, and geolocation of the authors); (ii) the social mood evolution (social network analysis, sentiment analysis, and a mood evolution matrix); and, (iii) graphic visualization (community detection, leader identification, path, and Fourier graphs).

Three lexicons can be used in text analysis: a general lexicon, a context lexicon, and an emoji lexicon. The Spanish version of the National Research Council Canada (NRC) (Mohammad and Turney, 2010, 2013) or AFINN (Nielsen, 2011) lexicons are provided as a general lexicon. The context lexicon includes specific words that are strongly related to the problem, with the corresponding positive or negative valences. During text analysis, the system first queries the context lexicon and then, if the word is not found, the general lexicon. Recognizing the importance of emojis in modern communication, we have integrated the emoji lexicon proposed by Godard and Holtzman (2022) into our text analysis tool. This lexicon is based on the original NRC Emotion Lexicon, allowing us to assign specific emotions to different emojis, from which we can analyze how these symbols are used to express feelings in digital communication. Research shows that emojis (like words) can convey a variety of emotions, and their inclusion in text analysis can significantly improve the accuracy of sentiment assessment. The integration allows us to assign specific emotions to emojis, thereby expanding our ability to analyze texts more comprehensively. By considering both words and emojis, our tool offers a deeper and more accurate analysis of the emotions and feelings expressed, better capturing the complexity of digital language.

The SSA-DSS was initially designed for the analysis of tweets using Spanish lexicons, but it can be applied to any other types of texts (messages, comments on messages, news in the press, technical or specialized reports, and so on). The tool imports, preprocesses, prepares, and analyzes tweets using lexicons. Additionally, this software includes tools for the visualization of the results and their exportation. From a technological point of view, a thorough study of the working environment was conducted before commencing the code implementation in order to ensure efficient and effective sentiment analysis. Python was chosen as the programming language, leveraging the flexibility and user-friendliness of Jupyter Notebooks in Anaconda. In this environment, we were able to divide development into stages and cells, facilitating project management and collaboration among team members.

The implementation relied on a carefully selected set of libraries and software packages. Data preprocessing was undertaken with `Spacy`, `Re`, and `NLTK`; `Pandas` for data manipulation and management; and, `Plotly` for detailed graph visualization. The database manager chosen was `MySQL`, using the corresponding driver for the Python connection. The `requests` package was used for interaction with the Twitter application programming interface (API) to gain access to necessary tweets. Various frameworks were evaluated for graphical interfaces. The decision was made to use Django due to its robustness, flexibility, and ability to develop complex web applications. This choice enabled the creation of an intuitive and user-friendly interface, allowing users to conduct sentiment analysis quickly and efficiently.

The tool was developed for both Linux/UNIX and Windows operating systems, ensuring compatibility and versatility across different platforms. The carefully selected multidisciplinary approach resulted in the creation of a comprehensive and robust sentiment analysis tool capable of meeting users' needs. The software runs on servers at the University of Zaragoza (Spain) and is currently at a beta level of development. Once completed, the intention is to offer an open basic version and market an advanced version, offering a free version for the higher education institutions.

The SSA-DSS could be very useful in many situations within the field of pandemic management, including the problem of air traffic management with respect to pandemic propagation. For example, it has been applied to the study of the evolution of the mood (sentiment) of Spanish society in response to public decision-making by respective governments (national and regional) in relation to COVID-19 vaccination in Spain, where it was possible to identify clear changes in trends in the sentiment of the population as a result of decision-making at different moments. We analyzed comments on Twitter, assessing social mood, and social support for the policies and health decisions on vaccination (Turón et al., 2023) in addition to analyzing tweets issued exclusively by social leaders (and their Twitter Communities), which were identified according to their influence on the rest of the participants in the debate (Navarro et al., 2024). The tool was also applied to press articles published about the eruption of the Cumbre Vieja volcano on the island of La Palma (Canary Islands) in 2021, guiding public decision-making over different temporal horizons, including the design of improved strategies in the risk reduction domain (Navarro et al., 2023).

Methodologically, sentiment analysis in social media involves a series of approaches that extract opinions, feelings, and thoughts from users by means of NLP procedures (Hasan et al., 2019). These approaches, especially those applied to Twitter exchanges, are growing rapidly to become a useful method for gaining insight into social mood and public opinion on any situation or event that affects society, example being COVID-19 (Manguri et al., 2020; Garcia and Berton, 2021; Turón et al., 2023; Navarro et al., 2024) and EVALI outbreaks (Kasson et al., 2021); hazard crisis responses (Behl et al., 2021); the impact of Brexit (Haider et al., 2020); the UK-EU referendum (Agarwal et al., 2018); general or presidential elections in USA (Somula et al., 2020; Chaudhry et al., 2021), India (Sharma and Ghose, 2020), and Indonesia (Budiharto and Meiliana, 2018); the Syrian refugee crisis (ztrk and Ayvaz, 2018); chemical attack (Bashir et al., 2021); the social impact of Hurrican Dorian (Badmus, 2020); and, extremism in social media (Asif et al., 2020).

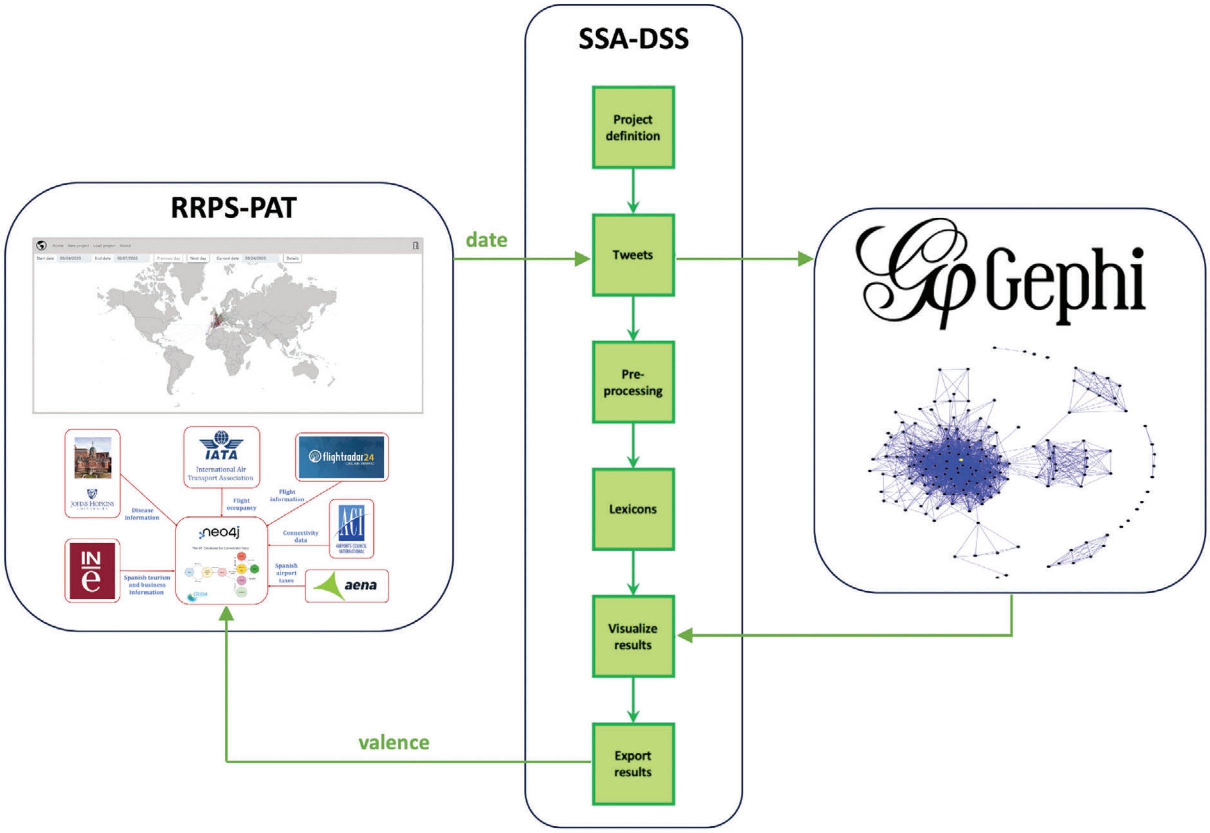

Fig. 1. Integration of SSA-DSS with the RRPS-PAT module and Gephi.

The SSA-DSS and RRPS-PAT systems (see Fig. 1) are integrated as follows: the percentage of canceled flights output as a result of the passenger air traffic management decision made by the RRPS-PAT system is saved, and the decision-making date is sent to the SSA-DSS. The SSA-DSS analyzes the tweets (using words and/or hashtags related to flight cancellations) over a time period ranging from the above date until a new decision is made. However, the time span can be shortened if the number of tweets drops over the time period. The accumulated valence of the tweets is then computed and sent back to the RRPS-PAT system. The percentage of canceled flights and the corresponding valence (impact on social sentiment) of the last and previous decisions are used to define a decreasing monotone function. This function is employed to predict the impact on population sentiment of any percentage of canceled flights associated with the solutions evaluated in the execution of the BPSO metaheuristic by the RRPS-PAT system, and a new decision is made.

The function range is (0,100), that is, the percentage of canceled flights and its values are in $(-1, 1)$ representing the corresponding valences. Its shape will be updated in line with the decisions made by RRPS-PAT. A linear decreasing function is initially taken when no decisions by RRPS-PAT are available. Once the RRPS-PAT is executed, different points representing the corresponding percentage of canceled flights and the associated valence become available. The function is built with a linear regression model. Other regression models can be considered if the linear condition is not met.

In the RRPS-PAT software, one of the attributes that measure social impact is the valence provided by the SSA-DSS. This attribute is combined with the others using weights derived from the decision-maker's preferences, following the additive model with ordinal information and the sum reciprocal method for weight calculation. The decision-maker must indicate the ordering, according to their preferences from highest to lowest, of the economic, social, and health impacts. Subsequently, they must order the corresponding attributes for each of the impacts. For example, with social impact, which consists of four attributes (passengers who cannot travel, lost connectivity, WELLBY, and valence provided by the SSA-DSS), the decision-maker must indicate their preferences among the four attributes by ordering them from most to least preferred. Once the decision-maker's preferences are obtained, the weights for each level of the objective hierarchy are determined, and the weights of all attributes are calculated by multiplying the weights from the root to the leaves in the objective hierarchy.

From a practical point of view, the interaction between the actors involved in the resolution process is as follows: the politician responsible for the decision asks the analyst who manages the RRPS-PAT to suggest the best decision regarding the opening and closing of connections and airports. Using the mathematical model (without using SSA-DSS), the analyst will suggest the measure that should be taken. Once the measure is applied, a debate will take place on social media (general or tailored to the problem) and, from the numerical and graphical analysis of the messages, social mood, communities, and their leaders are estimated using SSA-DSS and Gephi. With this information, the RRPS-PAT analyst will suggest new measures that already reflect the social mood, and so on until the politician feels that the process has concluded.

4.2. *SSA-DSS functionalities*

This subsection details the different functionalities of the software (SSA-DSS) and its connection with Gephi, which pinpoint the communities used to identify the leaders of social networks on the understanding that in many situations their behavior can reflect the feelings of society.

To access the SSA-DSS software, the user must first log in to its home page (see Fig. 2). The main menu of this page includes the options: Home, My Profile, New Project, Load Project, Text Analysis, About, and Logout, where the user can, respectively, view the welcome screen, modify the user profile, create or upload a project, directly analyze a text, read about the software, and log out. Users can directly paste a Spanish text from any source into the Text Analysis option and analyze its valence by using the Spanish version of the NRC lexicon. Moreover, the values associated with Plutchik's eight emotion categories (joy and sadness, anticipation and surprise, trust and disgust, and anger and fear) are also output (Mohsin and Beltiukov, 2019).

The New Project option is a six-phase procedure: Phase 1: project definition; Phase 2: tweet importation; Phase 3: text preprocessing; Phase 4: lexicon selection; Phase 5: result visualization/modification; and Phase 6: storage/exportation.

Phase 1 (Project definition) includes the following functionalities: new project opening, project title, and description.

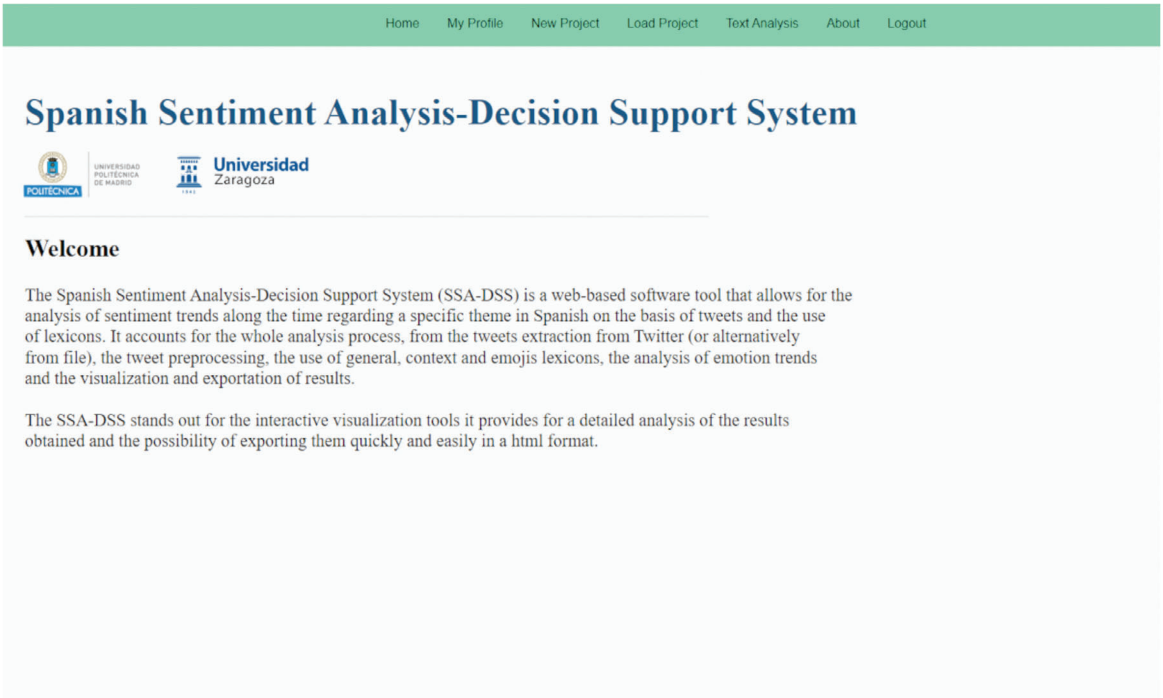

Fig. 2. Home page.

Phase 2 (Tweet importation). First, the users are asked if they want to query Twitter or directly load tweets from a file (they have previously downloaded from Twitter) (see Fig. 3a). If the users want to make a query, then they have to provide the parameters of the query in a similar way to Twitter (see Fig. 3b), including the search options "all of these words," "this exact phase," "any of these words," "none of these words," and "these hashtags" and the respective time period (dd/mm/yyyy). The system connects to the Twitter database via the API provided by Twitter.

Phase 3 (Text preprocessing) is related to tweet preprocessing before analysis. Ten preprocessing functions are available, which the user can select using the respective checkboxes (see Fig. 4). These preprocessing functions include the possibility of removing links or URLs, numbers, punctuation marks (commas, question marks, and so on), stop words, small words, and extra whitespaces, but also of normalizing chars and laughter by replacing repeated characters with a single repetition and the different variants of laughter, such as "hahaha" or "hehehehe" by the token "haha," respectively; and reducing words to their base or root form using the lemmization tool, which facilitates searches of the respective lexicon.

Phase 4 is associated with lexicon selection (see Fig. 5). Three lexicons can be used in tweet analysis: a general lexicon, a context lexicon, and an emoji lexicon. Users can choose a general lexicon, the Spanish version of the NRC or AFINN lexicons, or can load one from a file. An emoji lexicon is also available in the system. The user decides whether or not to use this lexicon or alternatively load one from the file. The context lexicon includes specific words that are strongly related to the analyzed problem, with the corresponding positive or negative valences. During tweet analysis, the system first queries the context lexicon and then, if the word is not found, the general lexicon.

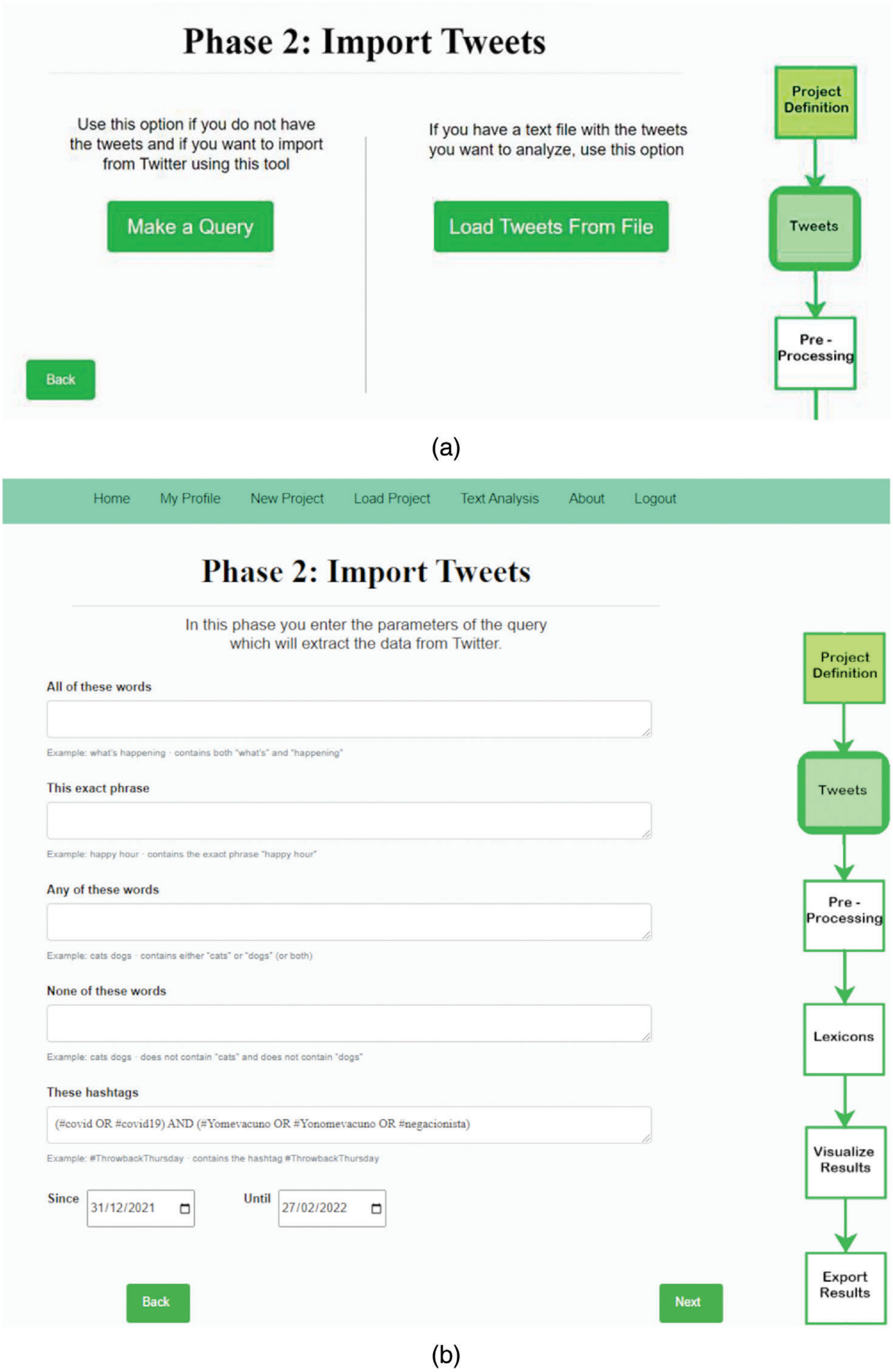

Fig. 3. (a) Tweets importation. (b) Making a query to Twitter.

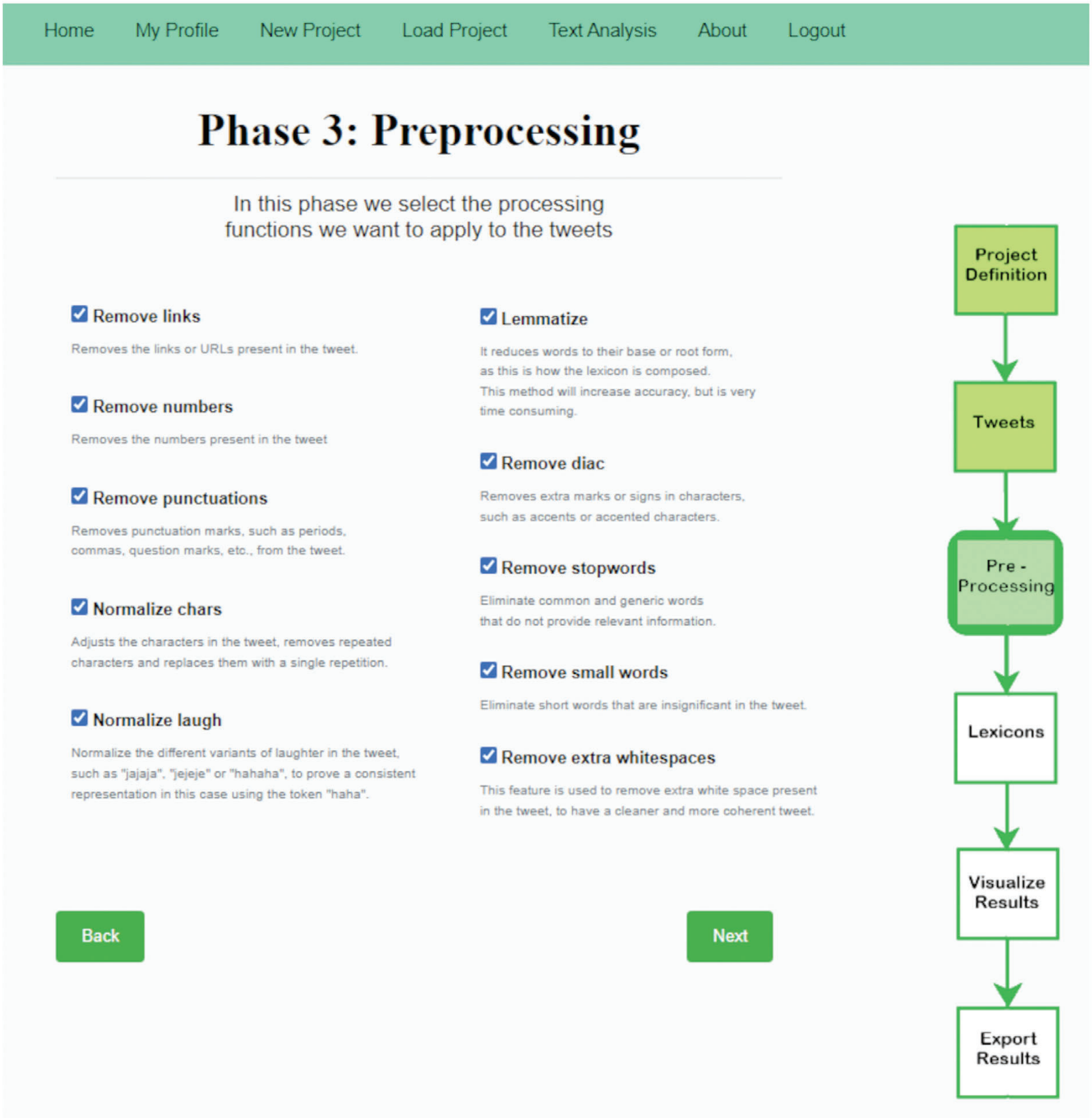

Fig. 4. Tweet preprocessing functions.

Note that the system users can create/modify a context lexicon, including the addition/removal of words and the corresponding valences, or valence updates. The context lexicon will be saved with the whole project. When the user clicks on the "Next" button, the tweets will be analyzed and then the result visualization options will be displayed.

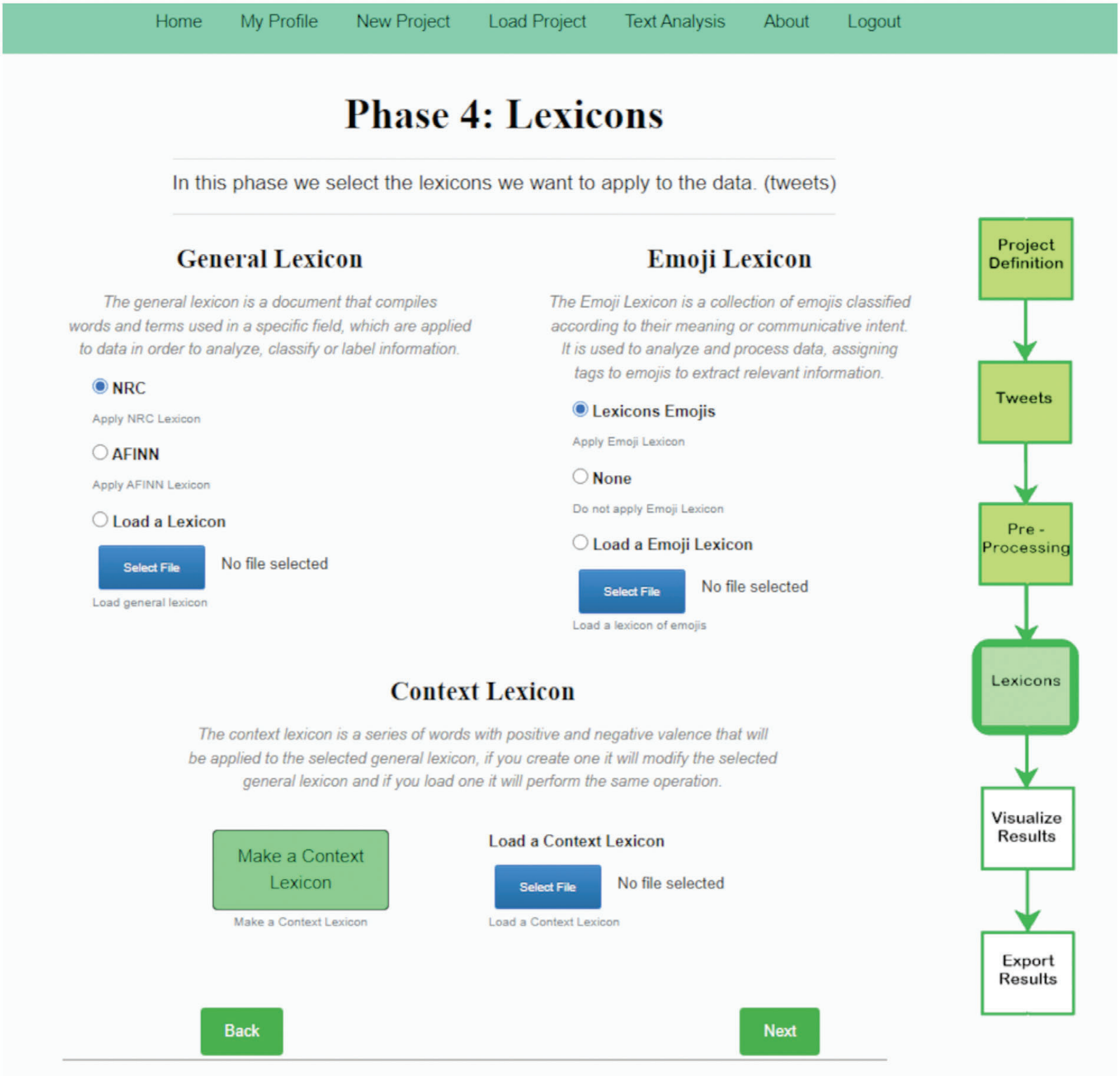

Fig. 5. Selection of general, context, and emoji lexicons.

In Phase 5, the results of the tweet analysis can be visualized by means of the following plots (see Fig. 6):

- *Tweet frequency*: This graph offers a visual representation of activity over time by illustrating the number of tweets published within a specific time period. It provides insights into the temporal dynamics of tweet frequency.

Fig. 6. Results visualization.

- *Sentiment evolution*: This visualization displays the evolution of sentiment expressed in tweets over time. It allows for a comprehensive understanding of emotional trends within the community, offering insights into changing sentiments. Fourier plot trajectory represents emotional valence versus the percentage of tweets (including tweets date).
- *Sentiment heat map*: This map shows the distribution of positive sentiment based on the time of day and day of the week. It provides a detailed view of when positive emotions are predominantly expressed and helps to identify temporal patterns of positive sentiment.
- *Most common words in tweets*: This graph offers a quick snapshot of the prevailing topics and subjects in the dataset and is useful for understanding the key themes discussed.
- *Wordcloud*: This visually striking graph represents the most relevant and frequent words in tweets. The size of each word is proportional to its frequency, providing a clear and graphical representation of prominent terms in the dataset.
- *Relationship between sentiment and tweet length*: Analyzing the relationship between sentiment and tweet length reveals patterns with respect to how sentiment varies with the length of tweets and provides insights into the emotional nuances associated with different message lengths.
- *Heatmap distribution of tweets by day and hour*: This map displays the concentration of tweets based on the day of the week and time of day, offering insights into the most active times and providing a temporal understanding of tweet distribution.
- *Positive feelings by users*: This visualization provides an overview of positive sentiments expressed by different users in their tweets, and it is useful for identifying influential or positive voices within the dataset.
- *Tweets by popularity*: This plot shows the distribution of tweets based on popularity or level of interaction, which is useful for identifying key or viral content within the dataset.
- *Tweets by popularity—top 10 users by number of Tweets*: This chart highlights the top 10 users with the highest tweet activity, offering a quick view of the most active contributors in terms of sheer volume.
- *Tweets by popularity—top 10 users by sentimental value*: Focusing on emotionally charged tweets, this graph identifies the top 10 users with the highest sentimental value, showcasing the most influential voices in terms of sentiment.
- *Confidence intervals*: This graph provides confidence intervals for different statistics or measures calculated from tweets, allowing for a more robust interpretation of the results, and offering a statistical perspective on data reliability.
- *Distribution of negative interventions by users over time*: Each point in this scatter plot represents an intervention by users with predominantly negative sentiments. The horizontal axis shows the intervention's timestamp, while the vertical axis indicates the associated sentiment. Each user is depicted with their own set of points, allowing observation of the variation in negative interventions over time for each individual. The dynamically adjusted point size based on the absolute sentiment value visually represents the intensity of negativity in each intervention.
- *Distribution of positive interventions by users over time*: Similarly, this scatter plot focuses on users with predominantly positive sentiments. Each point represents an intervention, with the horizontal axis displaying the timestamp and the vertical axis indicating the associated sentiment value. Each user has their own set of points, enabling observation of the variation in positive interventions over time. The dynamically adjusted point size based on the absolute sentiment value visually represents the intensity of positivity in each intervention. These temporal graphical

representations offer a unique perspective on the distribution of emotional tones over time for the analyzed users.

- *Users with the most negative sentiments*: This plot shows the top 10 users with the most negative tweets, providing a quick overview of negativity distribution in the dataset.
- *Users with the most positive sentiments*: This graph shows the top 10 users with the most positive tweets, providing a quick overview of positivity distribution in the dataset.

In Phase 6, the results and the whole project can be saved to a file. In the project file, information about the Twitter query, preprocessing functions used, and lexicons will be saved, together with the corresponding graphs. Graphs can also be saved into an html file for visualization. Moreover, the information can be exported into a file for further analysis in Gephi.

By exporting data to Gephi, we can identify leaders and extract arguments proposed in the debate. To this end, network influence is equated with a volume of shared content. Therefore, an influence network is built linking users A and B with directed arcs on the basis of the relationship $A\ \mathcal{R}\ B \Leftrightarrow B$ *retweets a message posted by* A, where each arc has a weight equal to the number of tweets posted by A that B has retweeted. The network is exported to a .gdf file that can be analyzed with Gephi software. This software applies the clustering algorithm mentioned below and can also be used for data visualization.

Leader identification is a two-stage process: (i) a community detection algorithm (Jacomy et al., 2014) is applied by Gephi to extract the clusters of users more connected with each other and ranking the users within each cluster by in-degree (number of times their messages are retweeted); (ii) community values are then returned to SSA-DSS, where the emotional valence of each user is computed taking into consideration the global polarity of the users' messages and the number of times each message has been retweeted (Navarro et al., 2024):

$$v_i = \sum_{k=1}^{n_i} w_{ik}(a_{ik}^{+} - a_{ik}^{-}),$$

where w_{ik} is the number of times that tweet k by user i has been retweeted and a_{ik}^{+} and a_{ik}^{-} are the emotional valences of the users' positive and negative tweets, respectively. The global polarity is then the net balance of the polarities of all messages posted by the user (i.e., all positive minus all negative values that output the sentiment analysis).

Users with higher in-degree and emotional valence values within each cluster are tagged as *leaders*. Once the leaders have been identified, the system calculates the correlation between the average emotional valences of the messages of the leaders and the whole population in a period of time around the main hotspots identified by the Fourier plot, that is, the moments in time when in which the mood of citizens has undergone a notable change in trend. The correlation between both series indicates the extent to which the leaders' messages can be considered representative of societal mood. Hence, the main arguments of the debate can be extracted from their posts, without having to analyze the full conversation.

Despite the original contributions of the present work that have already been noted, the research suffers from some limitations, mainly in the generation and integration of specific Spanish lexicons. Although the presence of Spanish on the Internet is high and increases every day, the generation and improvement of lexicons in Spanish does not advance as quickly as lexicons in the English language. Furthermore, the mass acquisition of textual data is not easy and may present certain biases, especially when acquired from social networks—their implementation in society is not total, and therefore it is difficult to evaluate their representativeness. In general, social media continues to be preferred by younger, IT-savvy members of society. Added to this is that non-human chatters (bots) which are often difficult to identify abound on social networks. Their presence produces an additional bias in the data whose effect is not easily quantifiable. In addition, social media does not usually provide reliable information that helps to filter messages coming exclusively from a specific geographical area.

Finally, the system allows for a better understanding of the problem, the opinions of stakeholders, and the facts that influence the mood of society and takes them all into account when making decisions that usually need to be quick and critical. Most political leaders are increasingly turning to social media to disseminate information about ongoing pandemics, natural disasters, response plans, and public health measures and connect with citizens (Rufai and Bunce, 2020), orienting the conversation towards the supportive mood needed to implement the corresponding public health or DRR strategies; therefore, identifying and monitoring those social leaders whose opinions most closely reflect the needs or demands of society will contribute to making public policies and decisions more realistic and effective.

To conclude Section 4.1, it is worth mentioning that the idea of incorporating emotions into the topic (RRPS-PAT), as with the management of most disasters, is to consider the emotional state of the society affected by the disaster when making decisions about it. The consideration and incorporation of the mood of society in formal models for disaster management usually leads to greater acceptance of the decisions made and greater involvement of citizens in their implementation; and this is particularly important when reaction time is vital. In terms of the methodology followed, the use of lexicons is motivated by the usual lack of appropriate information about the disaster to train the LLM methodologies. It is important to remember that the information extracted from previous disasters is strongly influenced by the changing boundary conditions.

The SSA-DSS is a Spanish language tool for evaluating the emotions and mood of society, and it integrates three lexicons (general, contextual, and emojis). As well as the options for managing all types of texts, the tool has numerous interactive graphics; it provides a Fourier graph that identifies key moments (milestones) in the mood throughout the period; through the use of artificial intelligence (ChatGPT in our case), it facilitates the extraction and dissemination of knowledge derived from the study by allowing the selection of the texts from the period set around the milestones, the identification of the arguments that characterize that period and the events that determine them. Similarly, the SSA-DSS is integrated with an external Gephi module that supplies the communities in which the individuals can be grouped and identifies the leaders of those communities. This information is especially relevant because when urgent decisions have to be taken, in many cases, it is enough to follow what the leaders of the communities say rather than analyze the group as a whole; something that is much cheaper and faster.

5. Illustrative example

Although the RRPS-PAT system is still under development, a working prototype version is available for Spanish airports (Peña et al., 2024). During the COVID-19 pandemic, the decisions taken on Spanish international air traffic were nothing like the ones that the RRPS-PAT system can take, since the only decisions taken were to reduce international air traffic by a certain percentage at a given time (70% reduction on March 27, 2020) or to suspend traffic with a country with a very high incidence rate (for instance, with the United Kingdom on December 22, 2020). Thus, this illustrative example focuses on the presentation of the SSA-DSS and its connection with Gephi. The COVID-19 pandemic is used as it had huge economic, social, and health impacts. Specifically, the opinions and prejudices of the general public, as well as of the political leaders, concerning the vaccination process in Spain, were reflected in social networks and had a significant impact on the progression towards achieving vaccination targets (Sattar and Arifuzzaman, 2021; Yousefinaghani et al., 2021).

Data were collected from the Twitter database using the Twitter API v2 (Navarro et al., 2024) and directly loaded from a file to SSA-DSS software following Fig. 3a. The search period ranged from February 27, 2020, to December 31, 2021, that is, from the beginning of the pandemic until the end of the main phase of the vaccination process in Spain.

The search key was built from the following hashtags in Spanish: #*covid;* #*covid19;* #*Yomevacuno* (I am getting vaccinated); #*Yonomevacuno* (I am not getting vaccinated); #*Negacionista* (denialist). Using the `httr` R package, the key string used to query the database was

```
(covid OR covid19) AND (Yomevacuno OR Yonomevacuno OR negacionista)
```

respectively, referring to COVID-19 and to the pro- and antivaccine positions. Only tweets written in Spanish and sent from Spain were included. This resulted in a dataset of 98,197 tweets.

The protocol described in Subsection 4.2 was applied to these 98,197 tweets, obtaining the graphical results explained in Subsection 4.2. Let us now detail the contributions of this research to the knowledge area of sentiment analysis.

Of the three lexicons (NRC, AFINN, or own lexicon) shown in Fig. 5, we chose the NRC Word-Emotion Association Lexicon Version 0.92 (Mohammad and Turney, 2013; Mohammad, 2016; Bravo-Marquez et al., 2016) as it has the largest number of words (14,182), making it perfect to analyze a new topic such as the COVID-19 pandemic. For the very same reason, we did not select any emoji lexicon. The Spanish version of NRC is a list of Spanish words and their associations with eight basic emotions (anger, anticipation, confidence, disgust, fear, joy, sadness, and surprise) and two sentiments (negative and positive). For each individual tweet, we calculated the valence, that is, the difference between the number of positive and negative words as well as the number of words associated with each of the above emotions and sentiments.

Regarding sentiment evolution over time, Fig. 7 shows the distribution of the positive and negative valences extracted from a DataFrame of tweets—the sentiment values are plotted against the corresponding timestamps. The chart includes two sets of markers: one for positive valence and another for negative valence. Positive valences are represented by markers above the x-axis, and negative valences are represented by markers below the x-axis. The x-axis represents the date, while the y-axis represents the sentiment value. The sentiment values are differentiated by their sign (positive or negative), providing a visual representation of changing sentiments over the specified time.

This chart can also be used to zoom into certain time periods where there has been a strategic political or health decision that has had a relevant impact on the evolution of social mood. For example, the zoom shown in Fig. 7, at the highest valence value, coincides with the start of vaccination in Spain with the Pfizer-BioNTech COVID-19 vaccine and the approval of the Moderna COVID-19 (MD) vaccine by the European Medicines Agency (EMA).

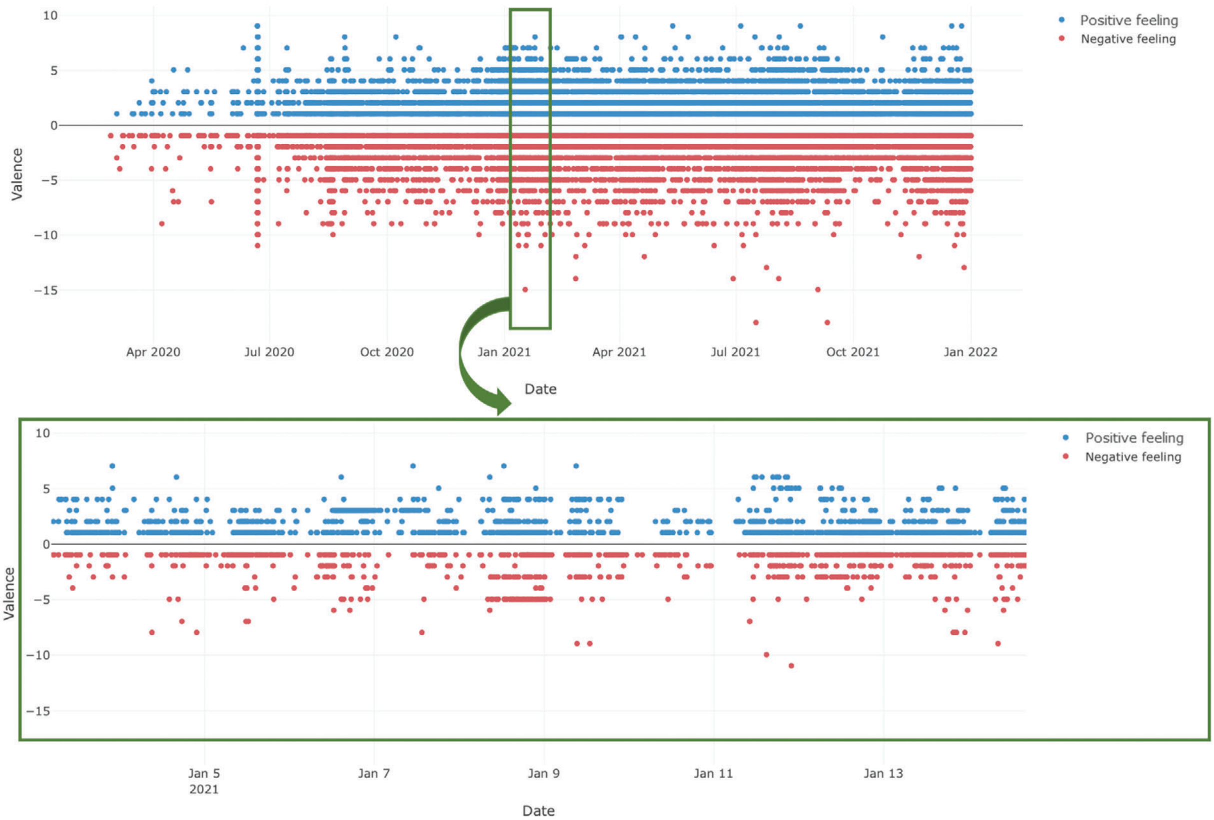

Fig. 7. Sentiment evolution over time.

Figure 8 illustrates the distribution of tweets with positive valence over time, showing the relationship between tweet creation dates, usernames, and valence, using color and size as additional indicators. The color intensity and size of the points further highlight the variations in the valence across different tweets. The generated graph is a scatter plot where each point represents a tweet, the x-axis corresponds to the timestamp of when the tweet was created, and the y-axis corresponds to the *username* of the tweet. The color of the points is mapped to the positive values of valence, using a color scale ranging from white to blue according to intensity valence. The size of the points is also determined by the valence, providing an additional visual dimension to the plot. Figure 8 can also be used to zoom into certain time periods where there has been a relevant impact on social mood. The zoom shown in Fig. 8, coincides with the same period of the start of vaccination in Spain.

Figure 9 plots the distribution of users' positive and negative tweets over time. As in the above two figures, it can be used to zoom into periods with relevant impact on vaccination. The zoom shown in Fig. 9 corresponds to the same positive period related to the start of vaccination in Spain.

Figure 10 shows the relationship between the length of tweets ("tweet_length") and their valence. Each point on the plot represents a tweet, with the x-axis indicating the length of the tweet and the y-axis representing the sentiment score. The color and intensity of each point correspond to its sentiment valence.

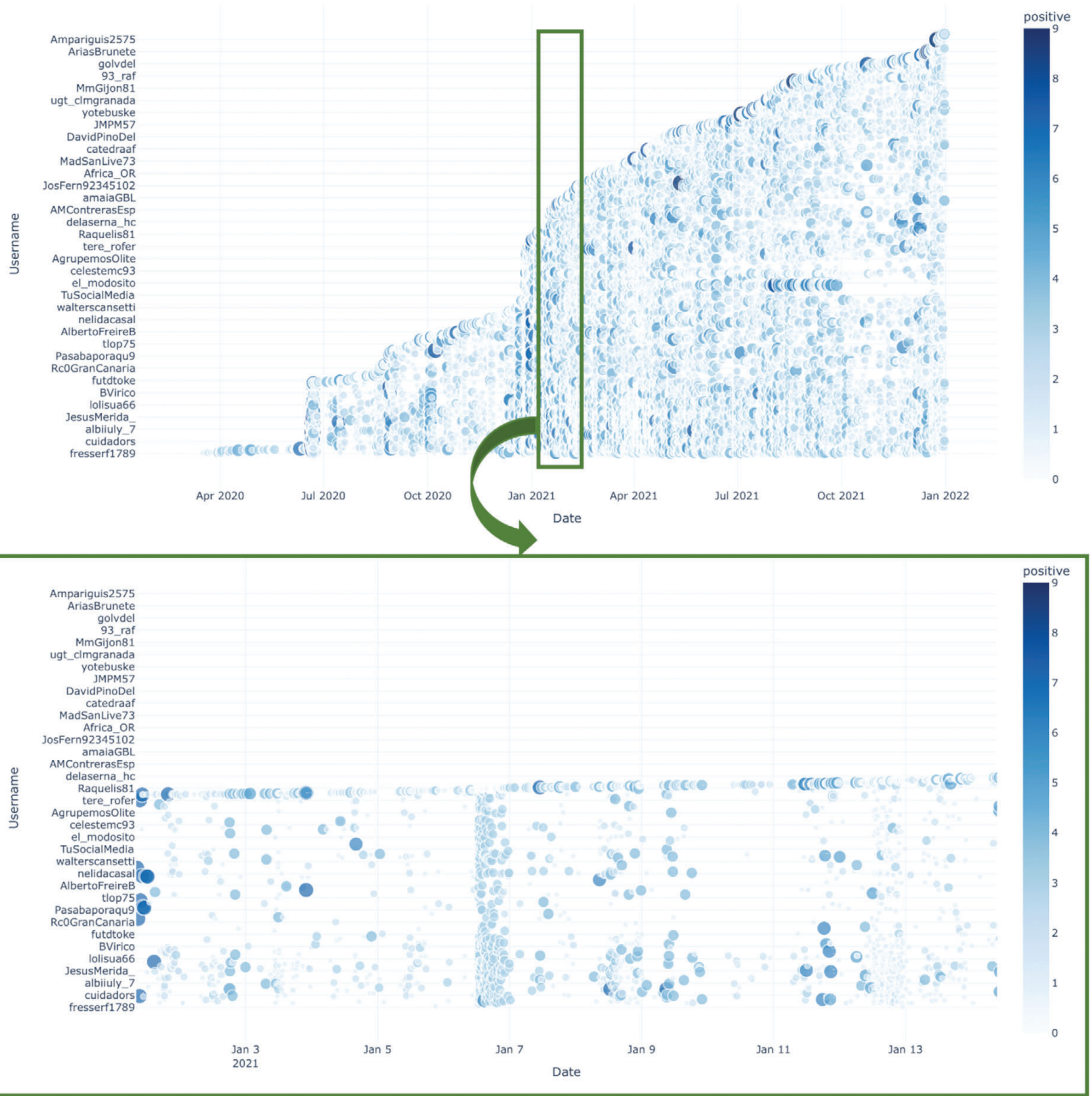

Fig. 8. Tweets by date.

Figure 11 shows the trajectory of the Fourier plot representing the emotional valence versus the percentage of tweets (date of tweets). From this analysis, we can see how the social mood of Spanish people changed over time, and whereas, by positioning the cursor, we can also pinpoint local hotspots and areas of trend change related to relevant news and political decisions. The plot starts with negative valences between 06/20/2020 and 06/24/2020 and then it increases with the release of Spanish vaccination strategy between 12/12/2020 and 12/14/2020 to reach the highest valence value between 01/04/2021 and 01/06/2021, corresponding to the start of vaccination in

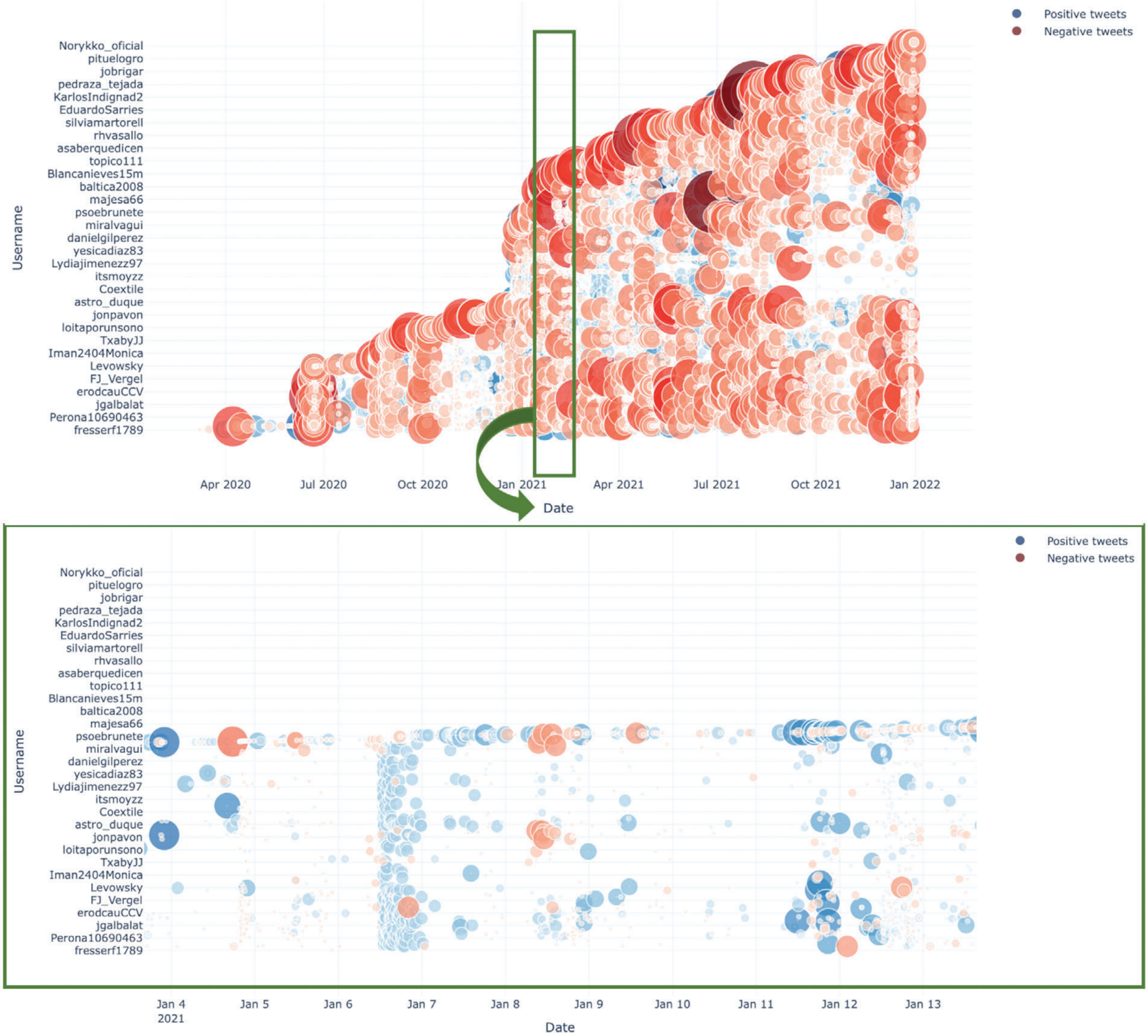

Fig. 9. Distribution of tweets.

Spain with the Pfizer-BioNTech COVID-19 vaccine and the approval of the Moderna COVID-19 (MD) vaccine by the EMA. Later, we find that the largest fluctuations occurred between March 27, 2021, and May 15, 2021, due to discordant health decisions on the Astra Zeneca (AZ) vaccine. Finally, the lowest valence value was found between October 16, 2021, and October 20, 2021, corresponding to the announcement of the need for booster doses and the debate on mandatory vaccination.

Once the most notable milestones in the evolution of valence have been identified, the tweets corresponding to a reduced period of days in the environment of the milestone are extracted, for example, ± three days, and artificial intelligence (ChatGPT with appropriate prompts) is used to extract

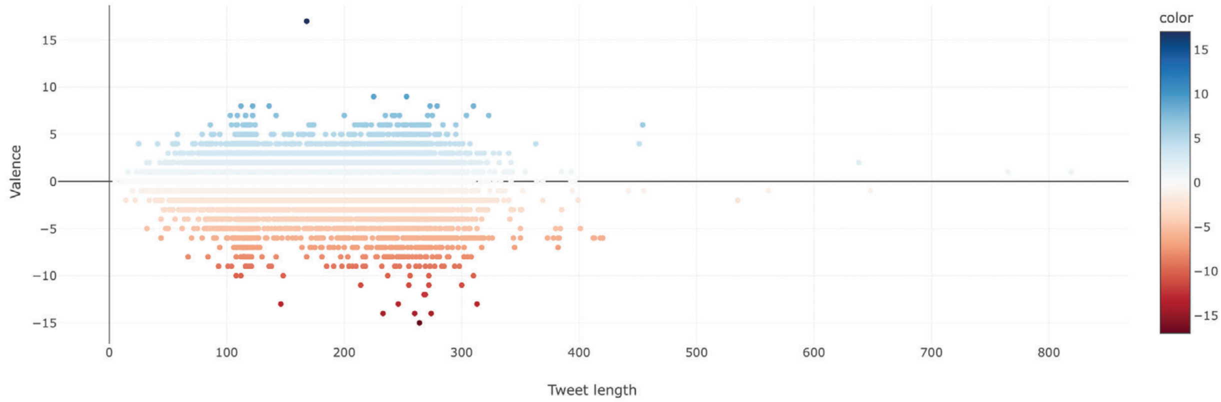

Fig. 10. Valence and length of tweets.

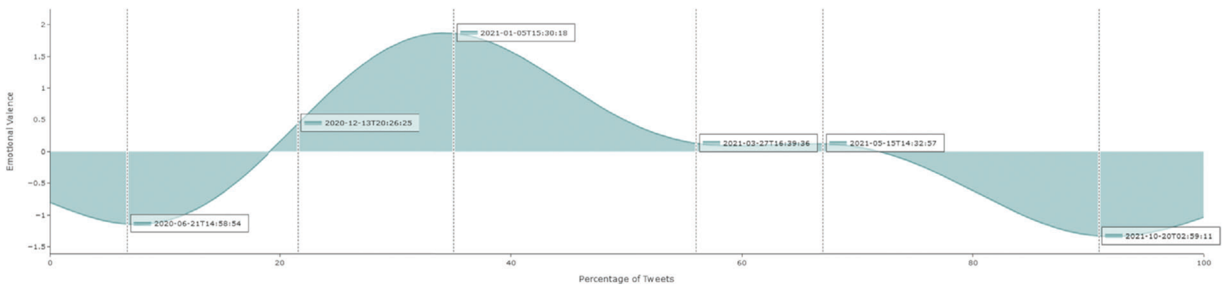

Fig. 11. Fourier plot trajectory.

the arguments that justify what happened at the milestone and the events that led to it. These arguments are validated with the experience and knowledge of the actors involved in solving the problem to extract relevant knowledge for future decisions.

As mentioned in Section 4, there are several limitations to this project that should be taken into account. First, the exact extent to which Twitter (X) represents the total population is unknown. In Spain, for instance, the number of users is just over 4 million, of a population of 48 million; the digital divide also skews user demographics. Second, in the example, it was necessary to differentiate texts written in Spanish by Spanish citizens from those written by Spanish-speaking users from other countries. We registered the information that the users themselves gave in the “location” field of the tweet.

As this field contains unstructured, textual information that the user can fill in arbitrarily, geolocation was obtained by searching the content of the field on the OpenStreetMap platform and selecting the location of the first result (the most “popular”). Not all users complete the field, and we were only able to derive the user’s country for a small number of them. It is also worth mentioning that in the example we could not use a contextual or an emoji lexicon, due to the lack of them. It is also worth mentioning that in the example we were not able to use a contextual or an emoji lexicon, due to none being available.

6. Conclusions

The connectivity and dynamism of our society favor the rapid spread of pandemics and the threat posed by their negative impacts. One, perhaps the most important, factor that contributes most to the rapid spread of pandemics is passenger air transport (PAT). Based primarily on COVID-19 experience, this topic is the subject of numerous studies focused on the development of mechanisms for risk estimation, early warning, immediate response, communication and collaboration, and rapid recovery.

An approach for immediate response procedures was first addressed in Jiménez-Martín et al. (2023). In addition to offering a methodology to assess the risk of pandemic spread through PAT, a series of epidemiological, economic, and social impacts that must be assessed and considered when taking the most appropriate measures were identified. They have also formulated a DSS (RRPS-PAT) to reduce the risk of pandemic spread (RRPS) through PAT, by opening and closing airports. Consideration of societal mood has been proposed, for the first time as far as we know, to assess the social impacts of the measures output by RRPS-PAT. The identification of citizens' emotions and the estimation of collective sentiment are carried out using the SSA-DSS tool. This tool is the result of a multidisciplinary project funded by the Government of Aragon, on which we have been working for two years.

The SSA-DSS covers the process from the incorporation of texts—particularly of tweets—to be analyzed, the application of preprocessing and cleaning tools, the selection of lexicons (general, context and emoji lexicons) to be used and their possible modification, to the visualization of results, and their export to other software tools. The possibility of incorporating general texts—not just tweets—for analysis, of using and customizing lexicons other than those initially offered in the program, and of having a notable set of formats for graphical visualization of the results are some of the notable features of SSA-DSS.

The SSA-DSS analyzes the societal mood over time periods starting as of the date a decision is made by the RRPS-PAT system on passengers' air traffic management involving a percentage of canceled flights. Then, the corresponding valences (impact on population emotion) of the last and previous decisions are used to define a function to predict the impact of any percentage of canceled flights associated with the solutions evaluated by the RRPS-PAT system on population emotion, and a new decision is taken.

This paper also discusses, apart from the RRPS-PAT module, the connection of the SSA-DSS with a social network analysis tool (Gephi) that complements its results by identifying the communities used to determine the social leaders and extracting the arguments that justify the different positions and decisions. This is of particular interest in many practical situations where it can be proven that the state of mind of society can be assimilated to that of the leaders of the different communities. It would be enough to use artificial intelligence to extract the arguments that characterize the milestones identified in the evolution of social sentiment and the events that have determined them. As the cognitive orientation claims, the consideration of social mood and the extraction and sharing of the knowledge derived from the resolution process increase the acceptance of public policies and the implication of citizens in their implementation. Although this option is now carried out outside the SSA-DSS, this is one of the most significant areas of improvement in this DSS.

In summary, this research underscores the critical importance of integrating advanced data analysis tools, such as SSA-DSS, with decision-making systems like RRPS-PAT to address the complex challenges posed by pandemics. By incorporating real-time sentiment analysis and social media insights, the proposed approach not only enhances the effectiveness of public health interventions but also ensures that the emotional and social impacts on the population are carefully monitored. The continuous improvement of these tools, including the integration of artificial intelligence to refine predictive and cognitive capabilities, represents a significant step forward in pandemic response strategies. Ultimately, this interdisciplinary effort contributes to more resilient and adaptive systems for managing global health crises in the future.

Based on the context, possible extensions to existing research and systems (RRPS-PAT and SSA-DSS) can be considered. In the case of the RRPS-PAT, these extensions could enhance the functionality, accuracy, and applicability of the systems for managing the spread of PAT and assessing societal mood:

- *Predictive analytics* use machine learning models to predict the impact of various PAT management decisions on pandemic spread and societal mood—these models can be trained on historical data to improve accuracy.
- *Geospatial analysis* integrates geospatial analysis tools to map the spread of the pandemic and the impact of PAT decisions on different regions—this can help in identifying hotspots and areas requiring immediate attention.
- *Automated reporting* implements features that generate comprehensive reports on the impact of PAT decisions and societal mood analysis—these reports can be customized for different stakeholders.
- *Policy simulation models* create simulation models that allow policymakers to test scenarios and their potential impact on pandemic spread and societal mood before implementing decisions—this can help in choosing the most effective measures.
- *Collaboration platforms* develop platforms for collaboration between government agencies, health-care providers, airlines, and other stakeholders to ensure a coordinated response to the pandemic.
- *International data sharing* facilitates comparison of the effectiveness of PAT management strategies across different countries—this can help in identifying best practices and areas for improvement.
- *Ethical AI* practices could ensure that the decisions made by the RRPS-PAT and SSA-DSS systems are fair and do not disproportionately impact certain groups.

For SSA-DSS, extensions could be aimed at

- *Real-time data integration*: The SSA-DSS tool could be enhanced by integrating real-time data sources such as news feeds, social media platforms (beyond Twitter), and public health databases—this could provide more timely and relevant information on societal mood and pandemic spread.
- *Exploitation of lexicons*: The inclusion of a discussion stage in the analysis of texts with a reduced number of relevant words (contextual lexicon) and relevant emojis (emojis lexicon) associated with the eight basic emotions and some others specific to the considered domain would provide a more effective exploitation of information and could lead to a more accurate (precise and unbiased) estimation of the social sentiment.

- *Integration of methodologies*: The integration of the information obtained when applying the lexicon-based approach to DRR problems into the LLM methodologies in mixed models would provide a more realistic and effective text analysis—this would help the use of the LLM methodologies in the DRR domain.
- *Extraction of knowledge*: Formalizing and integrating the use of the ChatGPT into the SSA-DSS would enable the extraction of the arguments that explain what happened in a specific period of time and the events that determine them—this would increase the potential of the SSA-DSS tool.

To conclude, it is worth noting that since the World Health Organization declared the recent Mpox outbreak to be a global public health emergency, several countries have been considering intensifying health controls at their main international airports. As the data from a similar outbreak in 2022 demonstrates, the risk of importing the virus is indeed significant (Kinoshita et al., 2023). This highlights the relevance and currency of our work.

Acknowledgments

This study was supported by the Regional Government of Aragón and the FEDER project "Participación Ciudadana Cognitiva y Decisiones Públicas. Aplicaciones Sociosanitarias" (ref. LMP35-21); the Spanish Ministry of Science and Innovation projects "Sistema de Ayuda a la Decisión basado en Aprendizaje Estadístico y Optimización en Redes. Aplicaciones a la Propagación de Pandemias a través del Transporte" (ref. PID2021-122209OB-C31), "Codecisión Cognitiva y Colaborativa. Aplicaciones en Salud" (ref. PID2022-139863OB-I00), and "Red Temática de Decisión Multicriterio" (ref. RED2022-134540-T).

References

Agarwal, A., Singh, R., Toshniwal, D., 2018. Geospatial sentiment analysis using Twitter data for UK-EU referendum. *Journal of Information and Optimization Sciences* 39, 1, 303–317.

Arduin, P.E., 2023. A cognitive approach to the decision to trust or distrust phishing emails. *International Transactions in Operational Research* 30, 3, 1263–1298.

Asif, M., Ishtiaq, A., Ahmad, H., Aljuaid, H., Shah, J., 2020. Sentiment analysis of extremism in social media from textual information. *Telematics and Informatics* 48, 101345.

Badmus, M.O., 2020. When the storm is over: sentiments, communities and information flow in the aftermath of hurricane Dorian. *International Journal of Disaster Risk Reduction* 47, 101645.

Balazs, J.A., Velásquez, J.D., 2016. Opinion mining and information fusion: a survey. *Information Fusion* 27, 95–110.

Bashir, S., Bano, S., Shueb, S., Gul, S., Mir, A.A., Ashraf, R., Shakeela, Noor, N., 2021. Twitter chirps for Syrian people: sentiment analysis of tweets related to Syria chemical attack. *International Journal of Disaster Risk Reduction* 62, 102397.

Behl, S., Rao, A., Aggarwal, S., Chadha, S., Pannu, H., 2021. Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises. *International Journal of Disaster Risk Reduction* 55, 102101.

Bravo-Marquez, F., Frank, E., Mohammad, S.M., Pfahringer, B., 2016. Determining word–emotion associations from tweets by multi-label classification. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI) Omaha, NE*. IEEE Computer Society, Washington, DC, pp. 536–539.

Budiharto, W., Meiliana, M., 2018. Prediction and analysis of Indonesia presidential election from Twitter using sentiment analysis. *Journal of Big Data* 5, 1, 51.

Carbonell, J., 2022. Análisis de la relación de la propagación del COVID-19 con la movilidad de origen y destino aeropuertos. MSc. final project, Universidad Politécnica de Madrid.

Catelli, R., Pelosi, S., Esposito, M., 2022. Lexicon-based vs. Bert-based sentiment analysis: a comparative study in Italian. *Electronics* 11, 374.

Ceria, A., Kstler, K., Gobardhan, R., Wang, H., 2021. Modeling airport congestion contagion by heterogeneous SIS epidemic spreading on airline networks. *PLoS ONE* 16, 1, 1–17.

Chaudhry, H.N., Javed, Y., Kulsoom, F., Mehmood, Z., Khan, Z.I., Shoaib, U., Janjua, S.H., 2021. Sentiment analysis of before and after elections: Twitter data of U.S. election 2020. *Electronics* 10, 17.

Chen, H., Wang, L., 2023. Pandemic preparedness and response: the role of air travel policies. *Public Health Reports* 138, 2, 145–155.

Chen, H., Wang, L., 2024. Air travel policies and pandemic response: evaluating effectiveness in the context of emerging infectious diseases. *Public Health Reports* 139, 2, 158–170.

Chen, Y., Yang, K., Xie, J., Xie, R., Liu, Z., Liu, R., Chen, P., 2020. Detecting the outbreak of influenza based on the shortest path of dynamic city network. *PeerJ* 8, e9432.

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C., Brugnoli, E., Schmidt, A., Zola, P., Zollo, F., Scala, A., 2020. The COVID-19 social media infodemic. *Scientific Reports* 10, 16598.

Damasio, A.R., 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 351, 1346, 1413–1420.

D'Andrea, A., Ferri, F., Grifoni, P., Guzzo, T., 2015. Approaches, tools and applications for sentiment analysis implementation. *International Journal of Computer Applications* 125, 26–33.

de Souza, D.B., da Cunha, J.T.S., dos Santos, E.F., Correia, J.B., da Silva, H.P., de Lima Filho, J.L., Albuquerque, J., Santos, F.A.N., 2021. Using discrete Ricci curvatures to infer COVID-19 epidemic network fragility and systemic risk. *Journal of Statistical Mechanics: Theory and Experiment* 2021, 5, 053501.

Do, H.J., Lim, C.G., Kim, Y.J., Choi, H.J., 2016. Analyzing emotions in Twitter during a crisis: a case study of the 2015 Middle East respiratory syndrome outbreak in Korea. In *2016 International Conference on Big Data and Smart Computing (BigComp)*, Hong Kong, China, 2016. IEEE, Piscataway, NJ, pp. 415–418.

Dong, M., Zhang, X., Yang, K., Liu, R., Chen, P., 2021. Forecasting the COVID-19 transmission in Italy based on the minimum spanning tree of dynamic region network. *PeerJ* 9, e11603.

Economics @Intelligence, 2022. Technological Solution, Digital Innovation Hub, Universidad Politécnica de Madrid, Madrid.

Fragua, A., Jiménez-Martín, A., Mateos, A., 2023. Complex network analysis techniques for the early detection of the outbreak of pandemics transmitted through air traffic. *Scientific Reports* 13, 18174.

Frijters, P., Krekel, C., 2021. *A Handbook for Wellbeing Policy-Making: History, Theory, Measurement, Implementation, and Examples*. Oxford University Press, New York.

Frijters, P., Sanchis, R.G., 2024. WELLBY. In Brockmann, H. and Fernandez-Urbano, R. (eds), *Encyclopedia of Happiness, Quality of Life and Subjective Wellbeing*. Edward Elgar Publishing, Cheltenham, UK.

Gao, J., Zheng, P., Jia, Y., Chen, H., Mao, Y., Chen, S., Wang, Y., Fu, H.H., Dai, J., 2020. Mental health problems and social media exposure during COVID-19 outbreak. *PLoS ONE* 15, e0231924.

Garcia, K., Berton, L., 2021. Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing* 101, 107057.

García-Moreno, J., Poveda, J., Villasante, Ó., Sánchez-Escalonilla, P., Caballero, A., Cestero, E., Lorenzo-Redondo, R., 2021. On-line platform for the short-term prediction of risk of expansion of epidemics: proof-of-concept based on COVID-19 evolution. Paper presented at the 14th USA/Europe Air Traffic Management Research and Development Seminar, Virtual, ATM 2021.

Godard, R., Holtzman, S., 2022. The multidimensional lexicon of emojis: a new tool to assess the emotional content of emojis. *Frontiers in Psychology* 13, 921388.

Haider, S., Ilyas, W., Soomro, Z., Anwar, A., Yaqub, U., 2020. Analyzing Brexit's impact using sentiment analysis and topic modeling on Twitter discussion. In *The 21st Annual International Conference on Digital Government Research*, Association for Computing Machinery Press, New York, pp. 1–16.

Hasan, M.R., Maliha, M., Arifuzzaman, M., 2019. Sentiment analysis with NLP on Twitter data. In *2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2)*, Rajshahi, Bangladesh, 2019. IEEE, Piscataway, NJ, pp. 1–4.

IATAs Air Connectivity Document, 2020. Measuring the connections that drive economic growth. https://www.iata.org/en/iata-repository/publications/economic-reports/air-connectivity-measuring-the-connections-that-drive-economic-growth/

Ishizaka, A., Nemery, P., 2013. *Multi-Criteria Decision Analysis: Methods and Software*. Wiley, Hoboken, NJ.

Jacomy, M., Venturini, T., Heymann, S., Bastian, M., 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE* 9, e98679.

Jiménez-Martín, A., Mateos, A., Peña, G.A., Moreno, A., 2023. A multi-objective approach to deal with international airspace closure/opening in Spain in an early-stage pandemic situation. In *2023 9th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, Piscataway, NJ, pp. 1062–1067.

Jurek-Loughrey, A., Mulvenna, M., Bi, Y., 2015. Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics* 4, 9.

Kahneman, D., 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.

Kaiser, M., Kuckertz, A., 2023. Emotional robustness in times of crisis: the effects of venture funding on the digital communication styles of entrepreneurs. *Journal of Small Business and Enterprise Development* 30, 4, 828–850.

Kasson, E., Singh, A.K., Huang, M., Wu, D., Cavazos-Rehg, P., 2021. Using a mixed methods approach to identify public perception of vaping risks and overall health outcomes on Twitter during the 2019 EVALI outbreak. *International Journal of Medical Informatics* 155, 104574.

Kermack, W.O., Mckendrick, À.G., 1927. A contribution to the mathematical theory of epidemics. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences* 115, 700–721.

Kinoshita, R., Sassa, M., Otake, S., Yoshimatsu, F., Shi, S., Ueno, R., Suzuki, M., Yoneoka, D., 2023. Impact of airline network on the global importation risk of Mpox, 2022. *Epidemiology and Infection* 151, e60.

Lee, J.M., Jansen, R., Sanderson, K.E., Guerra, F., Keller-Olaman, S., Murti, M., O'Sullivan, T.L., Law, M.P., Schwartz, B., Bourns, L.E., Khan, Y., 2023. Public health emergency preparedness for infectious disease emergencies: a scoping review of recent evidence. *BMC Public Health* 23, 1, 420.

Lins, M.E., Pamplona, L., Lins, A.E., Lyra, K., 2023. Metacognitive attitude for decision-making at a university hospital. *International Transactions in Operational Research* 30, 3, 1366–1386.

Lohmann, S., White, B.X., Zuo, Z., Chan, M.P.S., Morales, A., Li, B., Zhai, C., Albarracín, D., 2018. HIV messaging on Twitter: an analysis of current practice and data-driven recommendations. *AIDS* 32, 18, 2799–2805.

Manguri, K., Ramadhan, R., Mohammed Amin, P., 2020. Twitter sentiment analysis on worldwide COVID-19 outbreaks. *Kurdistan Journal of Applied Research* 5, 54–65.

Mohammad, S., Turney, P., 2010. Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. Association for Computational Linguistics, Stroudsburg, PA, pp. 26–34.

Mohammad, S.M., 2016. Sentiment analysis: detecting valence, emotions, and other affectual states from text. In Meiselman, H.L. (ed.) *Emotion Measurement*. Woodhead Publishing, Sawston, UK, pp. 201–237.

Mohammad, S.M., Turney, P.D., 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* 29, 3, 436–465.

Mohsin, M.A., Beltiukov, A., 2019. Summarizing emotions from text using Plutchiks wheel of emotions. In *Proceedings of the 7th Scientific Conference on Information Technologies for Intelligent Decision Making Support (ITIDS 2019)*, Atlantis Press, Amsterdam, the Netherlands, pp. 291–294.

Moreno-Jiménez, J.M., Vargas, L., 2018. Cognitive multiple criteria decision making and the legacy of the analytic hierarchy process. *Estudios de Economía Aplicada* 36, 67–80.

Navarro, J., Aguarón, J., Moreno-Jiménez, J.M., Turón, A., 2024. Social mood during the COVID-19 vaccination process in Spain. A sentiment analysis of tweets and social network leaders. *Heliyon* 10, 1, e23958.

Navarro, J., Piña, J.U., Mas, F.M., Lahoz-Beltra, R., 2023. Press media impact of the Cumbre Vieja volcano activity in the Island of La Palma (Canary Islands): a machine learning and sentiment analysis of the news published during the volcanic eruption of 2021. *International Journal of Disaster Risk Reduction* 91, 103694.

Nguyen, B., Xue, B., Andreae, P., Zhang, M., 2021. A new binary particle swarm optimization approach: momentum and dynamic balance between exploration and exploitation. *IEEE Transactions on Cybernetics* 51, 2, 589–603.

Nielsen, F.Å., 2011. Afinn. Informatics and mathematical modelling. Technical University of Denmark, Lyngby, Denmark.

Pang, J.K., Jones, S.P., Waite, L.L., Olson, N.A., Armstrong, J.W., Atmur, R.J., Cummins, J.J., 2021. Probability and estimated risk of SARS-CoV-2 transmission in the air travel system. *Travel Medicine and Infectious Disease* 43, 102133.

Peña, G.A., Mateos, A., Jiménez-Martín, A., Sanchis, R.G., 2024. A decision support system for risk reduction in pandemic spread based on the management of passenger air traffic. *International Transactions in Operational Research* Under review.

Qin, Z., Ronchieri, E., 2022. Exploring pandemics events on Twitter by using sentiment analysis and topic modelling. *Applied Sciences* 12, 23.

Ragini, J.R., Anand, P.R., Bhaskar, V., 2018. Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management* 42, 13–24.

Rexiline Ragini, J., Rubesh Anand, P., Bhaskar, V., 2018. Mining crisis information: A strategic approach for detection of people at risk through social media analysis. *International Journal of Disaster Risk Reduction* 27, 556–566.

Rodríguez-Escabias, D., 2023. Aplicación de técnicas de visión por computador para medir el riesgo de contagio por virus en aeropuertos. MSc. final project, Universidad Politécnica de Madrid, Madrid.

Rufai, S.R., Bunce, C., 2020. World leaders usage of Twitter in response to the COVID-19 pandemic: a content analysis. *Journal of Public Health* 42, 3, 510–516.

Samuel, J., Ali, G.G.M.N., Rahman, M.M., Esawi, E., Samuel, Y., 2020. Covid-19 public sentiment insights and machine learning for tweets classification. *Information* 11, 6.

Sattar, N.S., Arifuzzaman, S., 2021. COVID-19 vaccination awareness and aftermath: public sentiment analysis on Twitter data and vaccinated population prediction in the USA. *Applied Sciences* 11, 13.

Sharma, A., Ghose, U., 2020. Sentimental analysis of Twitter data with respect to general elections in India. International Conference on Smart Sustainable Intelligent Computing and Applications under ICITETM2020. *Procedia Computer Science* 173, 325–334.

So, M.K.P., Chu, A.M.Y., Tiwari, A., Chan, J.N.L., 2020. On topological properties of COVID-19: predicting and controlling pandemic risk with network statistics. *Scientific Reports* 11, 5112.

Somula, R., Dinesh Kumar, K., Aravindharamanan, S., Govinda, K., 2020. Twitter sentiment analysis based on US presidential election 2016. In Satapathy, S.C., Bhateja, V., Mohanty, J.R., Udgata, S.K. (eds) *Smart Intelligent Computing and Applications*. Springer Singapore, Singapore, pp. 363–373.

Turón, A., Altuzarra, A., Moreno-Jiménez, J., Navarro, J., 2023. Evolution of social mood in Spain throughout the COVID-19 vaccination process: a machine learning approach to tweets analysis. *Public Health* 215, 83–90.

Vargas, L.G., Moreno-Loscertales, C., Moreno-Jiménez, J.M., 2023. Conflict resolution in the era of cognitive multi-criteria decision-making: an AHP-retributive approach. *International Transactions in Operational Research* 30, 3, 1453–1478.

Verma, B., Thakur, R.S., 2018. Sentiment analysis using lexicon and machine learning-based approaches: a survey. In Tiwari, B., Tiwari, V., Das, K.C., Mishra, D.K., Bansal, J.C. (eds), *Proceedings of International Conference on Recent Advancement on Computer and Communication*. Springer Singapore, Singapore, pp. 441–447.

Vicente, E., Mateos, A., 2020. Analizamos el papel de los vuelos internacionales en su propagación. *The Conversation*. doi: 10.1016/j.jairtraman.2020.101819.

Wang, H., Wang, Z., Dong, Y., Chang, R., Xu, C., Yu, X., Zhang, S., Tsamlag, L., Shang, M., Huang, J., Wang, Y., Xu, G., Shen, T., Zhang, X., Cai, Y., 2020. Phase-adjusted estimation of the number of Coronavirus disease 2019 cases in Wuhan, China. *Cell Discovery* 6, 1, 10.

Wang, X., Vergeer, M., 2024. Effect of social media posts on stock market during COVID-19 infodemic: an agenda diffusion approach. *SAGE Open*. https://doi.org/10.1177/21582440241227688

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world's networks. *Nature* 393, 6684, 440–442.

World Health Organization, 2022. Imagining the future of pandemics and epidemics: a 2022 perspective. World Health Organization, Geneva, Switzerland.

Wu, J.T., Leung, K., Leung, G.M., 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* 395, 10225, 689–697.

Yousefinaghani, S., Dara, R., Mubareka, S., Papadopoulos, A., Sharif, S., 2021. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. *International Journal of Infectious Diseases* 108, 256–262.

Zamarrón-Mieza, I., Yepes, V., Moreno-Jiménez, J.M., 2017. A systematic review of application of multi-criteria decision analysis for aging-dam management. *Journal of Cleaner Production* 147, 217–230.

Zhang, T., Rabhi, F., Chen, X., Young Paik, H., MacIntyre, C.R., 2024. A machine learning-based universal outbreak risk prediction tool. *Computers in Biology and Medicine* 169, 107876.

ztrk, N., Ayvaz, S., 2018. Sentiment analysis on Twitter: a text mining approach to the Syrian refugee crisis. *Telematics and Informatics* 35, 1, 136–147.