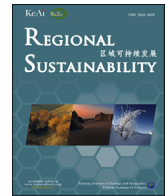




ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Regional Sustainability

journal homepage: www.keaipublishing.com/en/journals/regional-sustainability

Environmental complaint insights through text mining based on the driver, pressure, state, impact, and response (DPSIR) framework: Evidence from an Italian environmental agency

Fabiana Manservigi^{a,*}, Michele Banzi^a, Tomaso Tonelli^a, Paolo Veronesi^a,
Susanna Ricci^a, Damiano Distante^b, Stefano Faralli^c, Giuseppe Bortone^a

^a Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Via Po 5, Bologna, 40139, Italy

^b Department of Law and Economics, University of Rome Unitelma Sapienza, Piazza Sassari 4, Rome, 00161, Italy

^c Department of Computer Science, Sapienza University of Rome, Viale Regina Elena 295, Rome, 00161, Italy

ARTICLE INFO

Keywords:

Environmental complaints
Text mining approach
Term frequency-inverse document frequency (TF-IDF)
Driver, pressure, state, impact, and response (DPSIR) framework
Semantic network analysis
Regional agency for prevention
Environment and energy (arpae)

ABSTRACT

Individuals, local communities, environmental associations, private organizations, and public representatives and bodies may all be aggrieved by environmental problems concerning poor air quality, illegal waste disposal, water contamination, and general pollution. Environmental complaints represent the expressions of dissatisfaction with these issues. As the time-consuming of managing a large number of complaints, text mining may be useful for automatically extracting information on stakeholder priorities and concerns. The paper used text mining and semantic network analysis to crawl relevant keywords about environmental complaints from two online complaint submission systems: online claim submission system of Regional Agency for Prevention, Environment and Energy (Arpae) (“Contact Arpae”); and Arpae’s internal platform for environmental pollution (“Environmental incident reporting portal”) in the Emilia-Romagna Region, Italy. We evaluated the total of 2477 records and classified this information based on the claim topic (air pollution, water pollution, noise pollution, waste, odor, soil, weather-climate, sea-coast, and electromagnetic radiation) and geographical distribution. Then, this paper used natural language processing to extract keywords from the dataset, and classified keywords ranking higher in Term Frequency-Inverse Document Frequency (TF-IDF) based on the driver, pressure, state, impact, and response (DPSIR) framework. This study provided a systemic approach to understanding the interaction between people and environment in different geographical contexts and builds sustainable and healthy communities. The results showed that most complaints are from the public and associated with air pollution and odor. Factories (particularly foundries and ceramic industries) and farms are identified as the drivers of environmental issues. Citizen believed that environmental issues mainly affect human well-being. Moreover, the keywords of “odor”, “report”, “request”, “presence”, “municipality”, and “hours” were the most influential and meaningful concepts, as demonstrated by their high degree and betweenness centrality values. Keywords connecting odor (classified as impacts) and air pollution (classified as state) were the most important (such as “odor-burnt plastic” and “odor-acrid”). Complainants perceived odor annoyance as a primary environmental concern, possibly related to two main drivers: “odor-factory” and “odors-farms”. The proposed approach has several theoretical and practical implications: text mining may quickly and efficiently address citizen needs, providing the basis toward automating

* Corresponding author.

E-mail address: fmanservigi@arpae.it (F. Manservigi).

<https://doi.org/10.1016/j.regsus.2023.08.002>

Received 23 December 2022; Received in revised form 9 June 2023; Accepted 18 August 2023

(even partially) the complaint process; and the DPSIR framework might support the planning and organization of information and the identification of stakeholder concerns and priorities, as well as metrics and indicators for their assessment. Therefore, integration of the DPSIR framework with the text mining of environmental complaints might generate a comprehensive environmental knowledge base as a prerequisite for a wider exploitation of analysis to support decision-making processes and environmental management activities.

1. Introduction

The relationship between citizen and government has become a strategic focus area of modern public administrations. In this context, the interaction between these social actors is thought to provide a vector for the better provision of public service and efficiency in public institutions (Secchi, 2009; Ponte, 2015). Citizen may provide information to public administrations through different channels (e.g., online platforms, postal mail, call centers, and in-person services); and public administrations rely on such information to enhance their public services. Particularly, the modern citizen is online, which has created a culture of participation and involvement (e-governance), substituting for that of unilateral action. Recent research showed that governance supports transparency by improving administrative efficiency, promoting ethical behavior, and increasing trust and confidence in public institutions (Twizeyimana and Andersson, 2019). This new orientation towards citizen encourages public entities to improve and increase their access to service demands by implementing customer (citizen) relationship management (CRM) models, which aim at reducing operational costs, processing complaints, and shortening response times (Richter and Cornford, 2007; Chen, 2010). On the other hand, citizen is becoming more pro-active, demanding high-quality public services and a timely response rate from public authorities. In particular, complainants often expect public authorities to remediate the issues they raised, and this might be understood as a key driver of their satisfaction. In this sense, CRM models may contribute to changing the traditional and self-referential approach of government into one that is more citizen-centered with improved service quality and efficiency to meet the needs of citizen (Kannabiran et al., 2004; Al-Khouri, 2012). At the same time, the rapid development of information technology and growing environmental awareness of citizen have increased the number of environmental complaints, which represent the expressions of these grievances (such as poor air quality, illegal waste disposal, water contamination, and general pollution). A highly intelligent assessment of citizen's demands and complaints may contribute to detecting problems within a city or area, fostering dialogue between decision-makers, reallocating resources, improving service efficiency, enhancing service coordination, and predicting citizen's demand. Furthermore, the management of environmental complaints may also enhance the relationship between government and citizen, and increase citizen's satisfaction (Hartmann et al., 2017).

Rapid growth in the quantity of user-generated data makes big data analysis more necessary. Big data analysis provides a valuable tool for understanding realities by condensing a large amount of data into a small amount of information (Ghodousi et al., 2019). Recent progress in algorithms extends the application of big data analysis to even unstructured text data, including customer suggestions, complaints, and feedback (Katir et al., 2020; Lucini et al., 2020). As a result, many researches aim at using text mining approach to identify customer needs, in order to guide the development of market-oriented products (Zhan et al., 2009; Aguwa et al., 2017; Joung et al., 2019).

Since knowledge acquisition is the primary need for the continuous adjustment of programming and planning, government authorities must receive and redress grievances in a timely and efficient manner. In this sense, high-quality complaint-management systems may not only contribute to environmental compliance assurance, but also be considered as the drivers of “public value”, which can satisfy the actual and future needs of citizen (Deidda Gagliardo, 2002). From the perspective of environmental management, this means protecting the environment for its own sake and ensuring that the interests of future generations are considered through sustainable environmental solutions (Sami et al., 2021).

The primary step in this process is understanding the content of environmental complaints, which reflect local environmental problems or issues, with the expectation that complaint-handling body will intervene to address the situation. The driver, pressure, state, impact, and response (DPSIR) framework (Smeets and Weterings, 1999) is an important tool for understanding the interaction between people and environment, and identifying opportunities for developing sustainable and healthy communities. The framework recognizes that the state of environment (air, water and soil quality, nature, and landscape) is influenced by certain pressures, such as chemicals, waste, electromagnetic radiation, and noise. These pressures are influenced by economic and social driving forces (industry, population, and transport). Furthermore, the state or quality of the environment has impacts on health, well-being, businesses, and economy. For this reason, policy-makers may adopt different interventions or responses, such as controls, pressures, and actions, to mitigate impacts.

The DPSIR framework is based on the pressure-state-response framework initially proposed by Rapport and Friend (1979); it was adapted and largely promoted by the Organization for Economic Cooperation and Development (OECD) for environmental reporting (OECD, 1993). Several international organizations, including the United States Environmental Protection Agency (EPA, 1994), United Nations Environment Programme (UNEP, 1994), and the EU have adopted the framework, with the latter describing it as the most appropriate way of structuring environmental information (European Environmental Agency (EEA), 2005a). Indicators of the state and impact fall under the remit of the EEA (European Commission, 1999), which is required to communicate the state of the environment to policy-makers. The DPSIR framework is widely used to solve problem by both natural and social scientists, who further refine and apply the DPSIR framework and its derivatives in an on-going process tailored to a wide range of applications (Patrício et al., 2016).

The literature reports many cases, in which the application of the DPSIR framework highlighted essential connections and relations

to promote a more complete understanding of human-ecological systems (Cook et al., 2014). For instance, the framework has been used in several aquatic and marine ecosystem assessments to bridge the gap between science and decision-making (Gebremedhin et al., 2018; Mozumder et al., 2019), and it has also been applied to determine economic drivers and activities, contributing to coastal erosion and vulnerabilities (Tscherning et al., 2012; Lewison et al., 2016; Semeoshenkova et al., 2017). Mostly, the DPSIR framework has been used to develop indicators (Gabrielsen and Bosch, 2003); however, it has also been applied to organize the information contained in management plans and the communication between stakeholders (Khunanake et al., 2018). Finally, the framework has been used for new models and decision support tools to evaluate and compare decision outcomes (Cormier et al., 2013).

Decision-making happens “within the context of a social system that includes different levels of capacity, commitment, economics, political mandates and pressures, and cultural and traditional frameworks” (Loomis and Paterson, 2014). Thus, decision-making should be based on the best available evidence from a wide range of sources, including citizen reports. To support this, policies and processes must be comprehensively and systematically improved (Brewer, 2007). In particular, a complaint system that categorizes the contents of complaints according to the DPSIR framework might provide a systemic approach to responding to unknown or underappreciated drivers, pressures, and impacts (European Commission, Directorate-General for Environment, 2020), while providing crucial information for decision-makers to monitor policies and respond to unexpected events.

However, few studies have applied the DPSIR framework to categorize various perspectives and perceptions of stakeholders in order to foster a better understanding of sustainability challenges in different geographical contexts. A case study in Australia traced the link between local residents' perceptions of a range of water quality issues and the institutions responsible for managing the water sources drawing on the DPSIR framework to guide the data collection and establish response measures (Larson and Stone-Jovicich, 2011). Another study used focus group discussions and key informant interviews with the local community members to identify the drivers and impacts of land use or land cover and climate change on the environment, water, and livelihoods in the Lake Kyoga Basin (Uganda) (Obubu et al., 2022). In doing so, they gain insight into areas of convergence and divergences between residents' perceptions and institutional responsibilities and responses.

Although the DPSIR framework is a formal structure, some adaptations can be identified. Currently, there are at least 25 derivative schemes of the DPSIR framework and a widespread and increasing use of this framework to structure and analyze information for management and decision-making across ecosystems (Patrício et al., 2016). Since differences in terms and definitions may generate misunderstanding and ambiguity in the interpretation of the results, a declaration of the adopted terms is required (Bruno et al., 2020). The present study relies on the rigorous system of definitions developed by the EPA, which distinguishes among social, economic, and environmental factors. This decision is made to reduce confusion regarding the meaning of the DPSIR framework and ensure that the process could be easily replicated and applied to different systems, topics, and decision contexts. Fig. 1 presents the hierarchy of the DPSIR framework identified by the EPA and represents the basis for the keywords classification applied by the framework (Bradley, 2015).

The present study applied a text mining approach to identify the relevant keywords (denoting issues) within environmental complaints made by different stakeholders in nine provinces of the Emilia-Romagna Region, Italy. The keywords revealed by text mining approach might represent a starting point for finding repeated patterns in the environmental complaints, thus providing the basis for automated support of complaint processes. Second, the DPSIR framework was used to categorize complaints into driver, pressure, state, impact, and response, with the aims of supporting the planning and organization of information, and identifying stakeholder concerns and priorities, as well as measuring and evaluating indicators.

Overall, by applying a hybrid of powerful analytical capabilities based on the DPSIR framework, the present study proposed both quantitative and qualitative approaches to investigate pressing environmental complaints.

2. Materials and methods

This section reports the materials and methods employed in the present study. Fig. 2 shows the progressive steps taken to analyze the environmental complaints. The process began with the collection of data from the Regional Agency for Prevention, Environment, and Energy (Arpae) in Emilia-Romagna Region, Italy. Local challenges were identified by analyzing the location of each complaint and determining regional differences between stakeholder perceptions of environmental complaints.

2.1. Study area

Arpae is an environmental and technical body that supports local authorities to protect people health, ecosystems, and territorial safety in the Emilia-Romagna Region of Italy. Its activities cover all aspects of environment, including the monitoring of environmental components, the surveillance of human activities and their impacts, the assessment of environmental impacts on plans and projects, and the creation and management of environmental information systems. In addition, since 1 January 2016, Arpae became the permitting and inspection authority in the Emilia-Romagna Region. Typically, a permit refers to an environmental authorization, which establishes limits for pollutant emissions into air and water, and for the generation and management of waste, together with other environmental conditions that are specific to individual installation. Following receipt of a valid application, Arpae consults with other competent authorities to gather facts and opinions that might contribute to the determination of the application, including those of the public. This is consistent with initiative advocated by the Integrated Environmental Permitting Guidelines for Eastern Europe, Caucasus and Central Asia (EECCA) Countries to increase stakeholders' involvement in environmental affairs (OECD, 2005). Then, regular environmental inspections are conducted to verify compliance with environmental conditions. As the number of environmental complaints increases, the frequency of such inspections may increase.

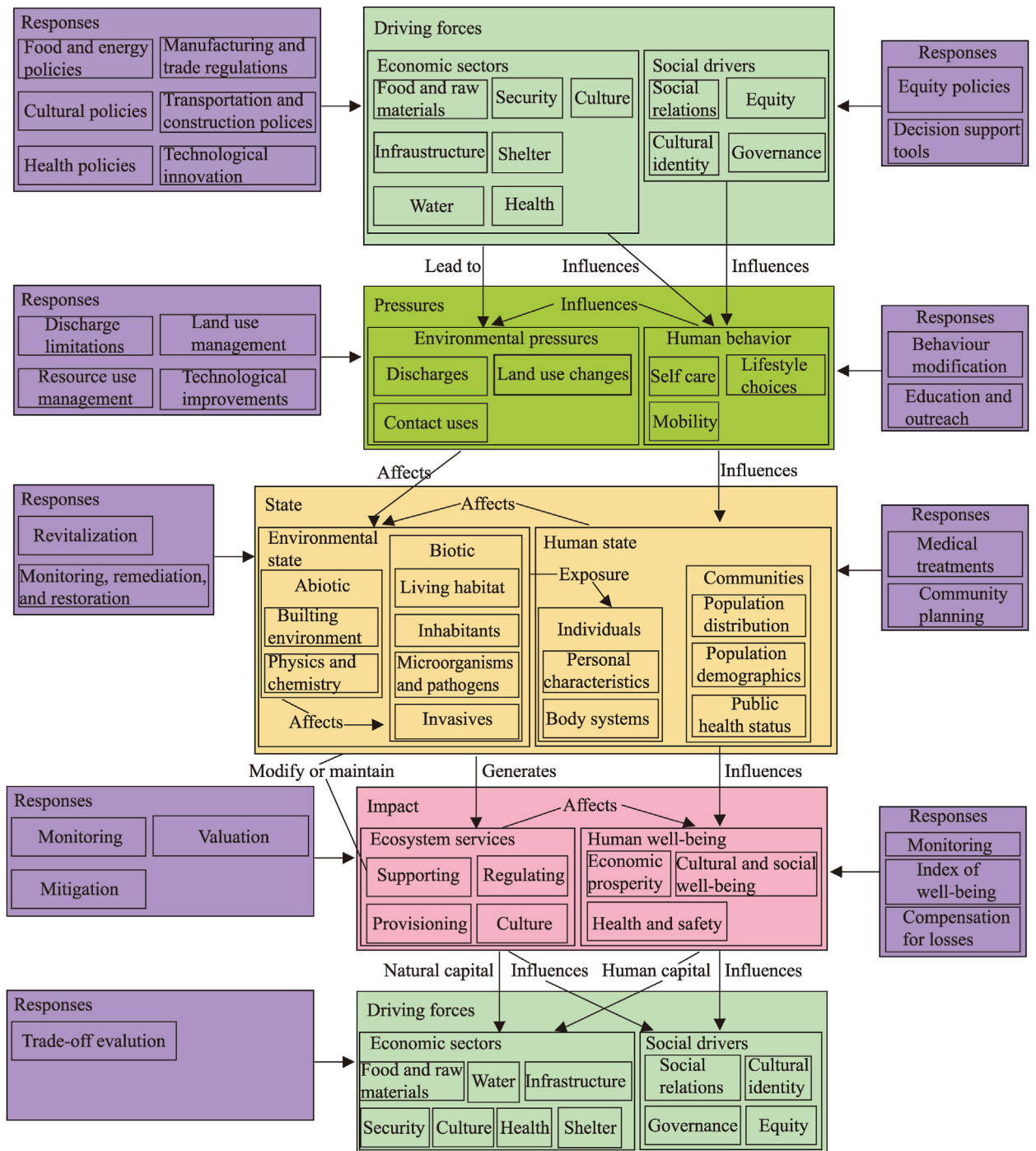


Fig. 1. Hierarchy of the driver, pressure, state, impact, and response (DPSIR) framework identified by Environmental Protection Agency (EPA).

The stakeholders include: (1) members of the public (including citizen associations) who may file complaints after discovering cases of non-compliance or maladministration and propose requests for information; (2) companies, as the main applicant for environmental authorizations; and (3) local authorities, who may request technical support or pass on complaints about matters falling outside their areas of competence.

The Emilia-Romagna Region is home to 4.40×10^6 residents, of whom 8.9% live in its capital of Bologna City (Italian National Statistical Institute (ISTAT), 2023). The region is divided into nine provinces: Piacenza, Parma, Reggio-Emilia, Modena, Bologna, Ferrara, Ravenna, Forlì-Cesena, and Rimini. The population density is 200 persons/km² in 2019, which is close to the national average.

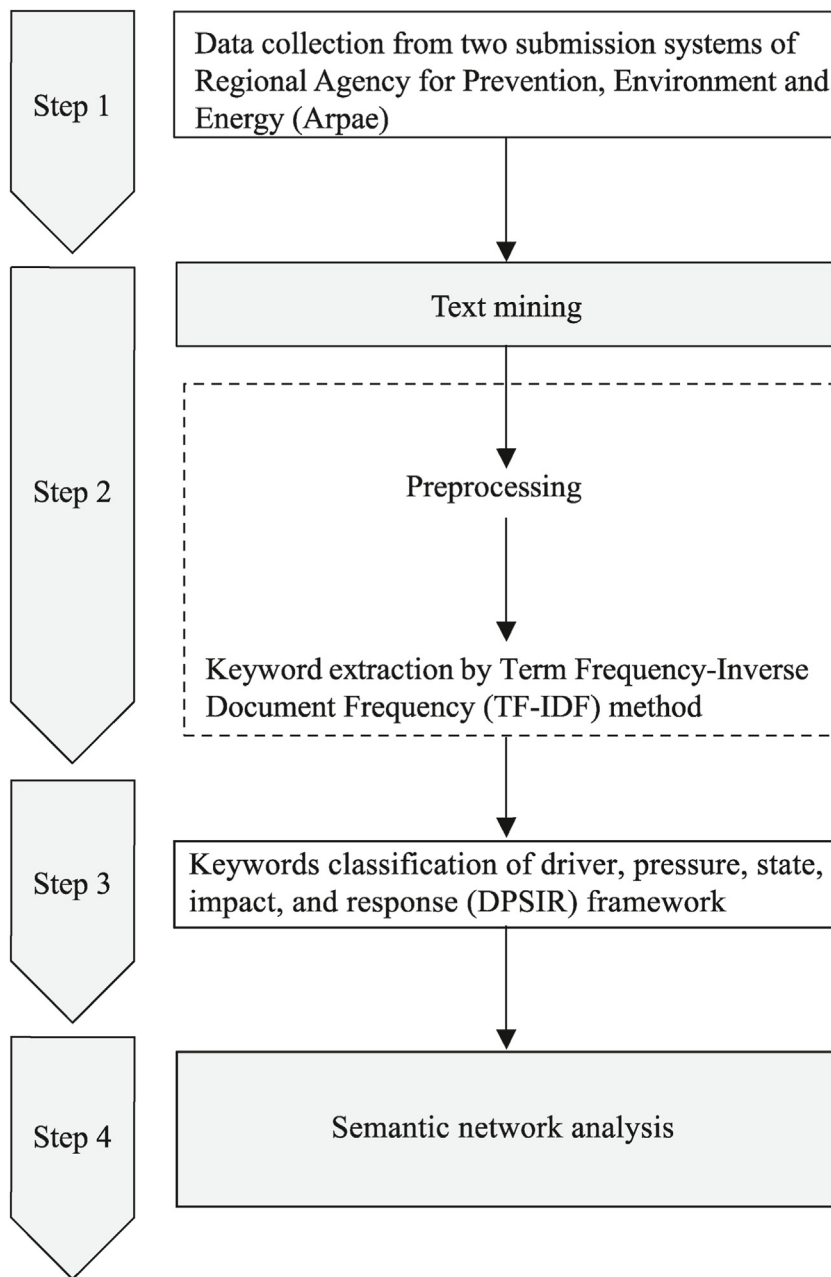


Fig. 2. Progressive steps assigned to analyze the environmental complaints.

Moreover, the population is evenly distributed. There is no dominant large city, but an axis of medium-sized cities along the Via Emilia (Italy), where two-thirds of the population and the majority of the industrial production are concentrated.

The Emilia-Romagna Region is well known for its automotive and motorcycle industries, food production, and ceramic industry, which have developed in the form of large and successful industrial districts spreading over the entire region. The manufacturing sector dominates the regional economy. As shown in Table 1, the biggest segments with manufacturing including: (1) food products, which are mainly distributed in Parma, Modena, Bologna, and Forlì-Cesena provinces; (2) tobacco products, which are only scattered in Bologna Province; (3) textile and wearing apparel industries, which are least distributed in Modena Province; (4) wood and products of wood and cork, straw, and plaiting materials, which are least distributed in Forlì-Cesena Province; (5) basic pharmaceutical products and pharmaceutical preparations, which are mainly distributed in Piacenza Province; (6) other non-metallic mineral products (including ceramics), mostly distributed in Ferrara and Forlì-Cesena provinces; and (7) machinery and equipment factories (including automotive and motorcycle industries), which are mainly distributed in Bologna, Reggio-Emilia, and Modena provinces.

Agriculture and farming (especially the pig and poultry sector) and related agro-industrial activities are critical for the Emilia-

Table 1
Socio-economic context of the Emilia-Romagna Region reported by the Italian national statistical institute (ISTAT).

		Piacenza	Parma	Reggio- Emilia	Modena	Bologna	Ferrara	Ravenna	Forli- Cesena	Rimini
Resident population		283650	450854	525155	702521	1011659	338477	385661	390868	338084
Total number of active enterprises		22498	37252	41987	59474	89695	23979	29475	33312	35093
Number of manufacturing enterprise	Food products	240	865	504	749	590	292	333	368	265
	Beverages	14	19	21	37	30	9	23	9	8
	Tobacco products	–	–	–	–	1	–	–	–	–
	Textiles	36	32	121	351	101	31	44	56	50
	Wearing apparel	63	147	490	1007	511	146	116	165	163
	Leather and related products	14	30	20	50	115	25	39	182	77
	Wood and products of wood and cork, except furniture, manufacture of articles of straw and plaiting materials	94	169	221	228	232	70	101	140	105
	Paper and paper products	10	35	36	65	56	19	15	38	10
	Printing and reproduction of recorded media	70	90	140	203	278	45	61	99	107
	Coke and refined petroleum products	–	1	1	2	1	–	1	–	–
	Chemicals and chemical products	23	43	45	69	127	17	36	27	28
	Basic pharmaceutical products and pharmaceutical preparations	2	4	2	8	11	1	–	2	3
	Rubber and plastic products	40	78	222	170	202	37	60	90	31
	Other non-metallic mineral products	58	84	213	279	120	54	115	101	89
	Basic metals	17	19	37	37	67	11	18	23	20
	Fabricated metal products, except machinery and equipment	512	922	1248	1452	1827	411	420	513	325
	Computer, electronic, and optical products	20	44	94	139	185	37	43	34	44
	Electrical equipment and of non-electric domestic appliances	54	82	184	190	269	58	49	67	67
	Machinery and equipment	194	459	558	678	790	150	188	167	134
	Motor vehicles, trailers, and semi-trailers	29	31	46	92	85	13	14	21	11
	Other transport equipment	12	15	19	24	59	13	18	47	23
	Furniture	39	84	139	122	159	29	42	289	117
	Other manufacturing	108	200	231	310	509	110	133	171	202
	Repair and installation of machinery and equipment	249	429	442	692	640	230	302	267	193

Note: These classifications are based on Nomenclature statistique des activités économiques dans la Communauté européenne (NACE) code, which is the statistical classification of economic activities in the European Community and is the subject of legislation at the European Union level.

Romagna Region. Looking more specifically at the agricultural sector, the most interesting crops are fruit (grapes), horticultural crops, tuber plants, legumes, wheat, maize, and sorghum. Nevertheless, the dominant activity is the tomato processing, with the Emilia-Romagna Region representing the second largest producer in Italy. Even regarding the dairy sector, the Emilia-Romagna Region is among the most productive regions in Italy, with regional cow milk production of approximately 1.77×10^6 t. Since the milk of the Emilia-Romagna Region is largely used to produce Parmigiano Reggiano and Grana Padano hard cheeses with relatively low whey, and the production of whey and buttermilk is also very high, approximately 1.45×10^6 t (Motola et al., 2009).

2.2. Customer relationship management in Regional Agency for Prevention, Environment and Energy (Arpae)

The Public Relations Office (Italian acronym for “Ufficio Relazioni con il Pubblico (URP)”) represents the entry point for all individuals who wish to connect with Arpae. It collects complaints, petitions, reports, suggestions, and claims aimed at facilitating the customer-agency relationship to guarantee the rights of information as enshrined in the Italian Law 241/90 on the transparency of the Public Administration, and other laws. The responsibilities of the URP are defined by Italian Law 150 of 7 June 2000 as “Discipline of the information and communication activities of the Public Administration”. Arpae has four URPs, one for each territorial area (west, central, metropolitan area, and east). Individuals can contact these URPs for the following reasons: requesting information on Arpae's services or environment; registering disservices and complaints; filing environmental incident reports; and citing events that may cause harm or potential harm to an environmental receptor (e.g., air, water, land, wildlife, and local habitat). In this case, the URP manually assigns the report to environmental prevention area of Arpae (territorial services) of the relevant territory covered by the report to address the problem.

The individual reporting a complaint, disservice, or environmental incident, or requesting information or data, typically need fill out online form titled “Contact Arpae”, which is available on Arpae website (www.arpae.it). Graphic elements help individuals navigate, and the form assists them structure their complaint to include all the relevant information with recourse to drop-down lists. In particular, individuals must complete the following fields: identity of the complainant, identification and location of the complaint, and the reason for complaint. When a complaint is filed, the system is configured to notify the URP operators by email. The URP operators take charge of the complaint within five working days. Currently, all the incoming user data are processed via a pipeline that is fully human-supervised by the URP operators. This costly and time-consuming effort includes human analysis by the means of reading text data. Subsequently, the URP operators process each claim or general request for information according to an own classification. Indeed, on the basis of human-supervised ticket tagging, the URP operators manually assign the request to a specific department of Arpae.

Additionally, Arpae provides a 24-h emergency service for environmental incidents throughout the region, concerning pollution phenomena that may cause serious and immediate damage to the environment. To activate environmental emergency intervention, individuals must contact Arpae by phone. Arpae' operators who receive the call then enter the data into the “Environmental incident reporting” portal to identify the individual making the claim and the characteristics of their report. The personal data provided by users are processed in accordance with regulation of the EU No. 2018/17251 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the institutions, bodies, offices, and agencies of the EU.

2.3. Data sources and processing

The proposed approach is based on a keyword extraction and semantic analysis of mined data regarding environmental complaints submitted to Arpae by different stakeholders (public, citizen associations, local authorities, and companies) in the Emilia-Romagna Region of Italy. Data were retrieved for one year (from 1 January 2021 to 31 December 2021) from two online complaint submission systems: Arpae's online claim submission system (“Contact Arpae”) and Arpae's internal platform for environmental pollution (“Environmental incident reporting portal”), which are managed by Arpae environmental technicians.

These sources represent the main communication channels between stakeholders and Arpae. Each record in the complaint dataset extracted from the “Contact Arpae” (in CSV format), which contained the date, complaint ID, province, municipality, personal contact information, user classification, claim topic, claim content, URP response, and response date (Table 2). Some of these fields were prompted by Arpae through drop-down menus within the contact form (e.g., user classification and claim topic). This online reporting platform was established to facilitate individuals' filing of complaints regarding environmental pollution and general information requests.

Each record in the complaint dataset extracted from the “Environmental incident reporting portal” (in CSV format) contained the date, complaint ID, province and municipality, exact location of the incident (address), position on the map (geolocalization), environmental incident classification, intervention level, territorial area responsible for managing the incident, and incident severity. These reports were filled in by Arpae' operators on the basis of calls.

In total, 732 valid records are from “Contact Arpae” and 1745 valid records are from the “Environmental incident reporting portal”. Records with missing or mismatched geographical information were excluded. Ultimately, 2477 records were evaluated (requests or claims that were unrelated to environmental issues were excluded). We classified environmental complaints based on the reported information, which includes environmental categories (air pollution, water pollution, noise pollution, waste, odor, soil, weather-climate, sea-coast, and electromagnetic radiation), stakeholders' claim topic, and the spatial distribution of complaints, using the Microsoft Power BI tool (Microsoft Power Platform, Washington, United States).

Appropriate precautions were taken to ensure anonymity by removing identifiable information and aggregating data by provinces or municipalities and stakeholder groups. The selected regular expression was developed to identify several identifiers including addresses,

Table 2
Example of complaint records from the online claim submission system “Contact Arpae”.

Field	Content
Date	15:04:15 on 24 June 2021
Complaint ID	64890
Province	Modena
Municipality	Maranello
Personal contact information	XXXX
User classification	Private
Claim topic	Odor
Claim content	I cannot open the window because the nearby factory smells like burning rubber at nighttime
Public Relations Office (Italian acronym for “Ufficio Relazioni con il Pubblico (URP) response	Your request has been taken care of by our territorial service and you will be contacted as soon as possible
Response date	09:10:30 on 25 June 2021

emails, phone numbers, and zip codes. A privacy statement explains how Arpae uses and shares users’ information and must be accepted before filling a request or complaint to the agency.

2.3.1. *Keywords extraction and the classification of the driver, pressure, state, impact, and response (DPSIR) framework*

Natural language processing was used to extract keywords from the datasets. For this task, the open-source tool of Orange (University of Ljubljana, Ljubljana, Slovenia) (Demsar et al., 2013) was employed to define the workflow depicted in Fig. 3. Each component (widget) embedded a data retrieval, preprocessing, visualization, modeling, or evaluation task. Data preprocessing was performed using the Orange widget Preprocess Text to Corpus which takes the text file in a specific data format as input. This preprocessing phase included transformation, tokenization, filtering, and normalization steps as follows:

Transformation means that transformed the input corpus to lowercase.

Tokenization broke up documents (words, n-grams, and phrases) into elements called tokens, which were used as input for text mining. For this process, “Regexp” was used to split the text by a given regex.

Filtering reduced the number of corpus features to those that were most significant by removing so-called stop words. A bespoke stop words dictionary was employed to filter out all terms considered irrelevant to the research topic.

Normalization further simplified the corpus features using a lemmatization process that involved a full morphological analysis to accurately identify the lemma of each word. Lemmatization removed inflection suffixes and compresses words into a lemma, defined as the canonical form of a term, and stripped of most conjugation, suffixes, and transformations.

N-grams ranging created n-grams from tokens, by means of a sliding window of size n. A range of one–two grams was used to allow for high granularity in the basic unit of data.

After preprocessing, a document-term matrix was built using the bag of word widget. Subsequently, data were input to the t-distributed Stochastic Neighbor Embedding method (t-SNE) widget to generate a document map. We selected a subset of documents relative to each environmental category, and extracted the first 10 keywords for each environmental category using the extract keyword widget based on the Term Frequency-Inverse Document Frequency (TF-IDF) (Eq. (1)) (Salton and Buckley, 1988). The TF-IDF computed the significance of a term t to document d by combining two scores (Eq. (2)), which is the frequency of term t in document d (where TF is the Term Frequency and IDF is the number of documents in the corpus containing t, regardless of the frequency) (Eq. (3)). If a term t appears frequently in a document d, but less frequently in other documents, the word can be considered more significant for distinguishing document d and expressing its core content.

$$TF-IDF = TF \times IDF, \tag{1}$$

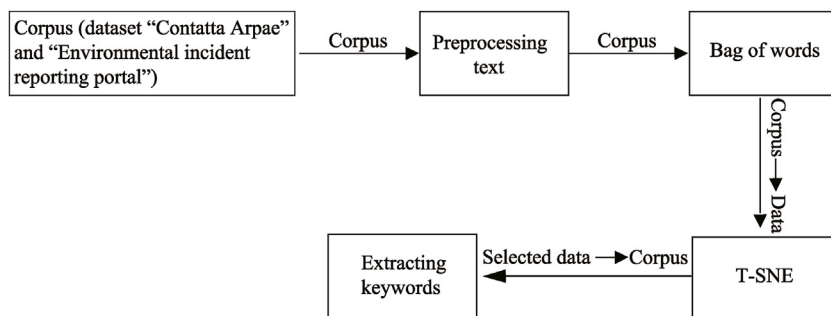


Fig. 3. Orange’s workflow for keyword extraction. T-SNE, t-distributed Stochastic Neighbor Embedding. We classified the keywords with a high TF-IDF values according to the categories and sub-categories of the DPSIR framework to define the interconnections between environmental related aspects within complaints (Bradley, 2015).

$$TF = \frac{f(t, d)}{|d|}, \tag{2}$$

$$IDF = \log\left(\frac{|D|}{|d/t \in d|}\right), \tag{3}$$

where $f(t, d)$ is the number of times term t appears in a document; D is the total number of documents; and $|d/t \in d|$ is the number of documents that include the term t .

2.3.2. Semantic network analysis of keywords in the DPSIR framework

Keywords were used to generate a semantic network illustrating the link between the DPSIR frameworks. For the semantic network analysis, text data were coded as nodes (words) and node pairs (word co-occurrences within a sentence or paragraph), which generate an $n \times n$ matrix.

No state-of-art ontology with explicit directed or undirected semantic relations was reported in the literature. The creation of such an ontology would require different approaches, possibly leveraging sophisticated ontology learning algorithms (Khadir et al., 2021) and human expertise. Instead, according to the distributional hypothesis (Harris, 1954), the present work only focused on co-occurrence graphs, which are basic knowledge that can be carried out, highlighting the semantic relations between the pairs of co-occurring words. Co-occurrence graphs were used to evaluate the centrality measures of semantic network analysis, providing a ground for further analysis.

In the present study, the semantic network of the corpus was created using the Orange widget Corpus to Network. We mined the top 160 nouns, adjectives, and verb keywords (with a frequency of 40 or more) according to a frequency analysis. Subsequently, the Orange widget Network Explorer was used to visualize the network graph (Fig. 4).

A one-way mode symmetric matrix was derived. At this stage, the Pajek tool (Batagelj and Mrvar, 1998) was used to identify the structure of the links between words and analyze centrality, thus quantifying the degree of relationships. This degree centrality was analyzed to identify keywords centered on the network among the extracted keywords; it indicated the number of nodes (keywords) connected to other nodes, with high values indicating many connective relationships between nodes. Additionally, betweenness centrality was calculated to identify the degree to which nodes are located between other network nodes. Keywords with greater betweenness centrality played an important role in facilitating connective flows within a network, thereby helping the network run seamlessly. For this reason, betweenness centrality was a useful metric for identifying central concepts in a corpus (Roche, 2011). Finally, 160 keywords were manually clustered into the five DPSIR framework categories. In the Pajek tool, we created random partition and selected five clusters and one-way mode. Then, we assigned cluster numbers to each vertex (1 = driver; 2 = pressure; 3 = state; 4 = impact; and 5 = response) on the basis of the DPSIR framework. We finally drew the network and relative partition, and observed the relationships among five categories of the DPSIR framework.

3. Results and discussion

3.1. Stakeholders' environmental complaints at the regional scale

We analyzed the differences in stakeholders' claim topic among the public, citizen associations, local authorities, and companies based on their self-classification through a drop-down menu on the online claim complaint system "Contact Arpae" and the filling out of the "Environmental incident reporting portal" managed by environmental technicians of Arpae. The public represented the main stakeholder group, showing 84.0% of all registered complaints, followed by companies (9.0%), and local authorities including municipalities, police department, fire department, and forest service (5.0%). Finally, 2.0% of the environmental complaints were registered by citizen associations.

Table 3 shows stakeholders' claim topic on environmental complaints. The public was more attentive to air pollution, which accounted for 32.8% of their registered complaints, followed by water and noise pollution, accounting for 17.5% and 16.8%, respectively. Companies demonstrated the most balanced focus, with 28.9% of their complaints focusing on water pollution, likely due to corporate requirements for sewage discharge standards, 12.4% referring to the weather-climate, and 11.9% pertaining to information requests related to environmental authorizations. Local authorities showed the greatest concern for noise pollution (34.4%), reflecting

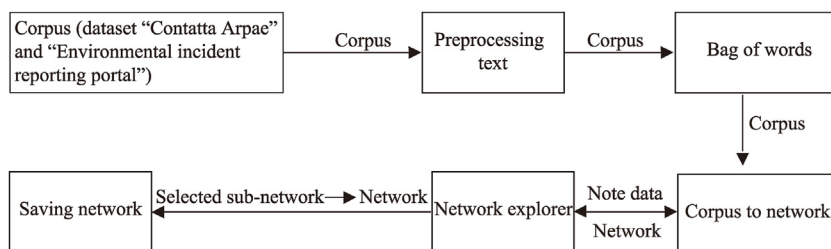


Fig. 4. Orange's workflow for the semantic network analysis.

Table 3
Stakeholders' claim topic on environmental complaints.

Claim topic	Percentage of claim topic to the total environmental complaints (%)			
	Public	Companies	Local authorities	Citizen associations
Air pollution	32.8	11.3	17.2	25.0
Water pollution	17.5	28.9	21.9	10.0
Noise pollution	16.8	2.6	34.4	20.0
Waste	8.7	10.3	6.2	7.5
Soil	7.4	4.1	3.9	2.5
Generic information	5.4	10.3	2.3	10.0
Electromagnetic radiation	4.3	6.2	12.5	10.0
Odor	4.0	1.5	1.6	10.0
Weather-climate	2.1	12.4	–	5.0
Environmental authorization	0.8	11.9	–	–
Sea-coast	0.2	0.5	–	–
Total	100.0	100.0	100.0	100.0

Note: means no data.

the fact that, they were the reference authorities in the Emilia-Romagna Region for noise pollution complaints. Indeed, the upon receipt of a complaint, municipalities opened formal administrative proceedings for the alleged acoustic pollution and request a phono-metric investigation from Arpa. The same applied for electromagnetic radiation, which accounted for 12.5% of local authorities' complaints. Local authorities also paid close attention to water (21.9%) and air (17.2%) pollution. Citizen associations represented people, groups, and organizations with environmental concerns in the local community. Although the proportion of environmental complaints from citizen associations was low, their grievances focused mainly on air (25.0%) and noise (20.0%) pollution, possibly reflecting local situations that require a response from Arpa.

3.2. Distribution of environmental complaints

The location and content of environmental complaints were analyzed to investigate spatial differences between stakeholder groups in their attention to different environmental complaints (Tables 4–7). With regard to the public, air pollution complaints were mostly associated with Modena Province, while noise and water pollution complaints were predominantly connected to Bologna and Parma provinces, respectively (Table 4). Companies located in Ravenna Province ranked water pollution as their first concern. In contrast, companies requesting data on the weather-climate or information related to environmental authorizations were mainly located in Bologna Province (Table 5). Most complaints forwarded by local authorities were from the eastern Emilia-Romagna Region, with Rimini, Forlì-Cesena, and Ravenna provinces submitting the most noise pollution complaints. Local authorities located in Rimini Province also gave the greatest attention to air and water pollution (Table 6). Citizen associations in Parma Province were most attentive to air pollution, while those in Modena Province expressed more grievances about odor. Of note, citizen associations were also concerned about electromagnetic radiation (Table 7).

3.3. Keywords of environmental complaints

The TF-IDF values were calculated to extract the most important keywords associated with eight categories of environmental complaints (air pollution, water pollution, waste, noise pollution, electromagnetic radiation, soil, odor, and weather-climate and sea-coast), indicating the most recurrent user-selected claim topics. Table 8 shows the top ten keywords for each category. Keywords

Table 4
Regional distribution of environmental complaints raised by the public.

Complaint topic	Number of environmental complaints									
	Bologna	Ferrara	Forlì-Cesena	Modena	Parma	Piacenza	Ravenna	Reggio-Emilia	Rimini	Total
Air pollution	75	69	41	216	57	60	68	57	22	665
Water pollution	51	21	35	57	69	29	38	43	11	354
Noise pollution	69	27	27	32	55	41	27	35	26	339
Waste	18	9	9	23	46	22	26	17	6	176
Soil	8	5	9	42	50	19	7	10	–	150
Odor	32	6	5	6	10	3	2	6	12	82
Generic information	10	6	11	32	17	4	9	17	3	109
Electromagnetic radiation	18	10	5	12	14	4	9	10	5	87
Weather-climate	13	1	4	11	2	3	4	3	–	41
Environmental authorization	3	1	2	1	1	2	3	3	–	16
Sea-coast	–	–	2	–	–	1	–	–	–	3
Total	297	155	150	432	321	188	193	201	85	2022

Note: means no data.

Table 5
Regional distribution of environmental complaints raised by the companies.

Complaint topic	Number of environmental complaints									
	Bologna	Ferrara	Forli-Cesena	Modena	Parma	Piacenza	Ravenna	Reggio-Emilia	Rimini	Total
Water pollution	10	6	3	4	7	5	12	6	1	56
Weather-climate	6	4	–	1	1	5	3	4	–	24
Environmental authorization	7	2	3	4	2	2	1	1	1	23
Air pollution	5	1	1	5	–	2	2	5	1	22
Generic information	5	1	3	3	2	1	2	1	2	20
Waste	3	1	2	–	5	1	3	2	3	20
Electromagnetic radiation	–	–	–	–	1	1	1	1	8	12
Soil	1	–	–	2	4	1	1	–	–	8
Noise pollution	1	–	–	2	–	–	–	1	1	5
Odor	–	–	–	3	–	–	–	–	–	3
Sea-coast	1	–	–	–	–	–	–	–	–	1
Total	39	15	12	24	22	18	25	21	19	194

Note: means no data.

Table 6
Regional distribution of environmental complaints raised by the local authorities.

Complaint topic	Number of environmental complaints									
	Bologna	Ferrara	Forli-Cesena	Modena	Parma	Piacenza	Ravenna	Rimini	Total	
Noise pollution	1	–	8	1	1	1	8	24	44	
Water pollution	6	1	10	2	1	1	2	5	28	
Air pollution	1	2	2	–	1	1	3	12	22	
Electromagnetic radiation	1	–	5	–	–	–	1	9	16	
Waste	1	–	3	1	–	–	3	–	8	
Soil	–	–	2	–	2	–	1	–	5	
Generic information	1	–	1	–	–	–	–	1	3	
Odor	–	1	1	–	–	–	–	–	2	
Total	11	4	32	4	5	3	18	51	128	

Note: means no data.

Table 7
Regional distribution of environmental complaints raised by citizen associations.

Complaint topic	Number of environmental complaints									
	Bologna	Ferrara	Forli-Cesena	Modena	Parma	Piacenza	Ravenna	Reggio-Emilia	Rimini	Total
Air pollution	–	2	1	1	5	–	1	–	–	10
Noise pollution	2	–	2	1	2	–	–	–	1	8
Electromagnetic radiation	2	–	1	–	–	–	–	–	1	4
Generic information	1	2	–	–	–	1	–	–	–	4
Odor	–	–	–	3	–	–	–	1	–	4
Water pollution	3	–	1	–	–	–	–	–	–	4
Waste	1	–	–	–	–	2	–	–	–	3
Weather-climate	1	–	1	–	–	–	–	–	–	2
Soil	–	–	–	–	–	–	–	1	–	1
Total	10	4	6	5	7	3	1	2	2	40

Note: means no data.

related to air pollution complaints included typical words such as “odor”, as well as sensory words such as “strong” (smell) and burnt plastic, and words indicating probable pollution sources (“factory”). Among the keywords of water pollution complaints, the pollution receptor “canal” ranked the first, followed by “foam”, “spillage”, and “stream”. Waste complaint keywords with the highest TF-IDF score were “littering”, “asbestos”, “material”, and “fire”. Within noise pollution complaints, the most important keywords were “request”, “factory”, “check”, and “trouble”. The words “house”, “condominium”, and “bar” also scored highly, mainly reflecting neighbor noise. The most critical keywords within electromagnetic radiation complaints comprised “request”, “radiobase stations”, “monitoring”, and “measurement”. For soil complaints, “spreading manure” and “sludge” from “farm” emerged as the most relevant. The keywords of odor complaints were strongly associated with words related to the duration of the annoyance (“hours”, “days”, and “night”) and the suspected cause, such as “spreading manure” from near “farm” and industrial complexes (as demonstrated by the keyword “factory”). Regarding the keywords of weather-climate and sea-coast complaints, citizen was most interested in receiving “data” (“rainfall data”).

Table 8

Top 10 keywords within each environmental complaint category according to term frequency-inverse document frequency (TF-IDF) score.

Rank	Air pollution		Water pollution		Waste		Noise pollution		Electromagnetic radiation		Soil		Odor		Weather-climate and sea-coast	
	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF	Keywords	TF-IDF
1	Odor	0.058	Canal	0.035	Littering	0.026	Request	0.038	Request	0.051	Spreading slurries	0.044	Hours	0.035	Data	0.052
2	Strong	0.029	Foam	0.023	Asbestos	0.024	Factory	0.015	Radiobase stations	0.027	Spreading manure	0.027	Strong smell	0.032	Website	0.030
3	Coming from	0.020	Spillage	0.020	Material	0.017	Check	0.015	Monitoring	0.024	Farm	0.024	Days	0.024	Rainfall data	0.020
4	Report	0.020	Stream	0.020	Fire	0.016	Trouble	0.014	Measurement	0.023	Zootechnical effluents	0.018	Air quality	0.023	Wind data	0.019
5	Air	0.020	Discharge	0.017	Request	0.013	Condominium	0.012	Mobile phone	0.022	Odor	0.018	Night	0.017	Information	0.016
6	Plastic	0.019	Bad smell	0.015	Action	0.013	House	0.012	Pollution	0.021	Spillage	0.017	Stink	0.017	Model	0.016
7	Burnt	0.019	Fish kills	0.014	Inspection	0.011	Bar	0.010	Cell tower	0.019	Request	0.016	Spreading manure	0.016	Daily data	0.013
8	Factory	0.018	Request	0.013	Police	0.010	Systems	0.010	Cab	0.017	Police	0.016	Evening	0.015	Annual data	0.013
9	Foundry cooperatives	0.017	River	0.012	Plastic	0.010	Verification	0.009	Radio-frequencies	0.016	Dung heap	0.013	Farming	0.015	Weather stations	0.013
10	Hours	0.017	Water discoloration	0.011	River	0.010	Inspection	0.009	Emissions	0.015	Sludge	0.011	Factory	0.014	Weather forecast	0.013

3.4. Keywords of the DPSIR framework

Keywords with TF-IDF value greater than 0.001 were manually grouped into the categories of driver, pressure, state, impact, and response for each environmental complaint. The following sections summarized the main environmental complaints on the basis of the keyword selection.

3.4.1. Air pollution and odor keywords

Economic activities such as factories were perceived as the most important causes of air pollution and odor. The distribution of air pollution complaints was closely related to population and specific industrial sectors. For example, the words “foundry cooperatives” and “ceramics” were most associated with the Modena Province, which has ceramic districts within its eight municipalities. According to environmental complaints, the agricultural sector was another important driver of air pollution. In particular, the “farm” sector (“cattle” and “dairy farm”) was considered as a key contributor to odor. The occurrence location of the word “farm” was distributed in Ferrara, Forlì-Cesena, Ravenna, Modena, Parma, and Rimini provinces, in decreasing order of relevance, which aligns with the widespread production of animal products (e.g., Parma ham, Bologna mortadella, and Parmigiano Reggiano cheese) in these provinces. Furthermore, complainants considered that “incinerators” and “biogas plant” is causally linked to air pollution. In general, individuals living close to foundries, ceramic factories, farms, incinerators, and biogas plants complained air pollution and odor. Pollutants induced changes to the environment and resulted from the operations of industry (“smoke”, “fumes”, “emissions”, and “dusts”), human activities (biomass “burning”), or agricultural practices (“spreading slurries” and “spreading manure”). Alterations in air quality (“burnt plastic”, “acrid smell”, “bitumen smell”, “chemical smell”, “ammonia smell”, and “unbreathable air”) were considered as the environmental triggers of “sore throat”, “headache”, “eye irritation”, and “respiratory problems”. Indeed, unpleasant odor could be a warning sign of potential risks to human health. Air pollution and odor also affected the quality of life, leading to “not sleeping” and “awakening”. As temperatures rose in summer, odor complaints gained in frequency, with individuals complaining of not being able to close their windows and having to avoid sitting outside. Furthermore, odor frequency (“hours”, “days”, “night”, and “evening”) was an important predictor of odor (Table 9).

3.4.2. Water pollution keywords

“Factory” and “farm” were considered as the main economic drivers of water pollution. The formation of “foam” in freshwater ecosystems such as “canal” and “river” was reported several times. Complainants often referred to land-spreading activities associated with “spreading slurries” and “gas oil spillage” as sources of human behavior pressure on water quality. Leaks or “spillage” could also emerge from machinery inside factories, or from breaks in containers and pipes holding dangerous liquids and gasses (“gas oil spilling”). Oil slicks eventually moved toward the shore, harming aquatic life (causing the death of “fish”) and damaging recreational areas. High quantities of nutrients in water from industrial crop fertilizers and animal waste might cause eutrophication and “algae” blooms that depleted oxygen levels in water and kill marine life. Among their cited impacts, complainants also reported “strong smell” and “nauseating smell” (Table 10).

3.4.3. Waste keywords

“Littering”, “asbestos”, “plastic”, “abandoned rubbers”, and “demolition waste” were the most frequent triggers of waste complaints. Similar to the water pollution complaints, waste complaints referred to altered ecosystems (“river” and “soil”), particularly in the event of “spillage”. A “bad smell” and perceived “danger” emerged as the main reasons of waste complaints (Table 11).

Table 9
Keywords and TF-IDF scores of air pollution and odor in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces	Factory (0.018), foundry cooperatives (0.017), farm (0.010), installation (0.006), fireplace (0.005), ceramics (0.004), yard (0.003), dairy farm (0.003), incinerator (0.002), biogas plant (0.002), and painting (0.001).
Pressure	Environmental pressures	Smoke (0.019), fumes (0.009), emissions (0.007), dusts (0.007), wind (0.006), pollution (0.004), gas fuels (0.004), hydrocarbons (0.003), sludge (0.003), and poultry manure (0.002).
	Human behavior pressures	Burning (0.005), spreading slurries (0.004), spreading manure (0.004), burning waste (0.004), trucks traffic (0.003), and spillage (0.001).
State	Environmental state (abiotic state-environment)	Burnt plastic (0.019), acrid smell (0.014), burning smell (0.012), unbreathable air (0.011), smelly fumes (0.009), bitumen smell (0.008), chemical smell (0.007), burnt rubber (0.005), fire (0.004), ammonia smell (0.003), smell of paint (0.003), acidic smell (0.002), and smog (0.001).
	Environmental state (biotic state-inhabitants)	Mosquitoes (0.003).
Impact	Human well-being (health and safety)	Strong smell (0.029), hours (0.017), stink (0.010), windows closed (0.007), unbearable odor (0.007), sore throat (0.006), unpleasant smell (0.006), nauseating smell (0.005), health (0.004), discomfort (0.003), headache (0.003), respiratory problems (0.002), eye irritation (0.002), and awakening (0.001).
Response	Driving forces-based responses	Request (0.006), check (0.004), intervention (0.003), inspection (0.003), data (0.003), mail feedback (0.002), police (0.002), controls (0.002), and compliance (0.001).
	Impact-based responses	Monitoring (0.001).

Table 10
Keywords and TF-IDF scores of water pollution in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces	Factory (0.007) and farm (0.004).
Pressure	Environmental pressures	Foam (0.023), wastewater discharge (0.017), drainage system (0.010), pollution (0.010), well (0.008), hydrocarbons (0.006), discharge in sewer (0.005), fumes (0.003), and detergent (0.002).
	Human behavior pressures	Spillage (0.020), spreading slurries (0.011), gas oil spilling (0.005), pipe burst (0.003), and accident (0.002).
State	Environmental state (abiotic state-living habitat)	Canal (0.035), stream (0.020), and river (0.012).
	Environmental state (abiotic state-environment)	Water discoloration (0.011), algae (0.002), and environmental degradation (0.002).
	Environmental state (biotic state-inhabitants)	Fish (0.014).
Impact	Human well-being (health and safety)	Bad smell (0.015), smelly (0.011), strong smell (0.005), and nauseating smell (0.004).
Response	Driving forces-based responses	Request (0.013), police (0.011), check (0.007), inspection (0.008), intervention (0.004), data (0.004) voluntary ecological guards (0.003), and information (0.003).
	Pressure-based responses	Waste management (0.004) and water treatment plant (0.003).
	Impact-based responses	Analyses (0.002).
	State-based responses	Remediation (0.010).

Table 11
Keywords and TF-IDF scores of waste in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	General driving force	Landfill (0.009).
Pressure	Environmental pressures (discharge)	Littering (0.026), asbestos (0.024), material (0.017), plastic (0.010), demolition waste (0.009), liquid waste (0.005), and oil in the soil (0.005)
	Human behavior pressures	Spillage (0.008), illegal dump (0.008), truck transport (0.007), incident (0.004), abandoned cars (0.004), waste burning (0.004), and abandoned rubbers (0.003).
State	Environmental state (abiotic state-environment)	Fire (0.016), environmental degradation (0.006), and contaminated soil (0.002).
	Environmental state (biotic state-living habitat)	Soil (0.014), river (0.010), and vegetation (0.004).
Impact	Human well-being (health and safety)	Danger (0.006), odor (0.004), and bad smell (0.003).
Response	Driving forces-based responses	Request (0.016), action (0.013), inspection (0.011), police (0.010), information (0.004), feedback (0.004), data (0.003), and waste disposal (0.009).
	Pressure-based responses	Uncontrolled disposal (0.007), waste recovery (0.005), waste transport (0.005), and waste management (0.005).

3.4.4. Noise pollution keywords

As expected, the major drivers for noise pollution were commercial establishments (“bar”, “hotel”, “restaurant”, and “supermarket”), as well as store animals (“barking of dogs”). Noise caused by loud “music” could be linked to commercial establishments, parks or playgrounds, streets or sidewalks, residences, and religious establishments (“bells”). Noise pollution associated with street traffic (“motorway”), transport (“train”), and construction (“yard”) presented annoyances and, at worst, impacted the quality of life. Heating, ventilation, and “air conditioning system” were reported as drivers of noise pollution. Noise, especially when persistent, or all day long or continuous could be distressing and reducing the quality of life (“trouble”, “discomfort”, and impact on “rest”) (Table 12).

3.4.5. Electromagnetic radiation keywords

Sources of electromagnetic radiation (“radiobase stations”, “mobile phone”, “antenna”, “cell tower”, and “cab”) were thought to expose local residents to low but continuous levels of electromagnetic radiation (Table 13). Complainants were often concerned about the health effects of this exposure. Noticeably, within the entire dataset, the word “health” was reported in 48 complaints, and more than 30.0% of these related to electromagnetic radiation. Concerns expressed in complaints reflected a high risk perception linked to this issue.

3.4.6. Soil keywords

The soil keywords showed many similarities with the water pollution keywords, except for the TF-IDF score ranking of some words. In particular, complainants attributed the main soil contamination pressures to “spreading slurries”, “spreading manure”, and “zootechnical effluents” (Table 14).

Table 12
Keywords and TF-IDF scores of noise pollution in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces	Factory (0.015), bar (0.010), installations (0.010), hotel (0.008), restaurant (0.007), yard (0.006), supermarket (0.005), traffic (0.004), air conditioning system (0.004), irrigation system (0.004), train (0.003), motorway (0.003), and filtering system (0.002).
Pressure State	Social driving forces	Barking of dogs (0.007), music (0.007), and bells (0.004).
	Environmental pressures	Vibrations (0.004) and emissions (0.004).
Impact	Environmental state (abiotic state physical and chemical)	Loud noise (0.006) and decibel (0.003).
	Human well-being (health and safety)	Trouble (0.014), condominium (0.012), house (0.012), night disturbances (0.009), all day long disturbances (0.009), disturbing noise (0.006), continuous noise (0.005), not sleeping (0.005), loud noises (0.005), discomfort (0.004), windows closed (0.004), unbearable noise (0.003), and rest (0.003).
Response	Driving forces-based responses	Request (0.038), check (0.015), verification (0.009), inspection (0.009), information (0.009), intervention (0.008), and investigations (0.007).
	Impact-based responses	Quiet (0.005) and surveys (0.005).

Table 13
Keywords and associated TF-IDF scores of electromagnetic radiation in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces (culture and infrastructure)	Radiobase stations (0.027), mobile phone (0.022), antenna (0.019), cell tower (0.019), cab (0.017), power line (0.009), radiobase stations (0.027), mobile phone (0.022), antenna (0.019), cell tower (0.019), and cab (0.017).
Pressure State	Environmental pressures	Pollution (0.021), radio-frequencies (0.016), emissions (0.015), field (0.012), and waves (0.012),
	Environmental state (abiotic state-physical and chemical)	High voltage (0.009).
Impact Response	Human well-being (health and safety)	Health (0.011), exposure (0.009), risks (0.005), and concern (0.002).
	Driving forces-based responses	Request (0.051), information (0.015), check (0.013), data (0.011), feedback (0.005), thematic maps (0.004), and appointment (0.002).
	Pressure-based responses	Monitoring (0.024), measurement (0.023), meters (0.008), distance (0.006), nearby (0.005), survey (0.006), and values (0.005).

Table 14
Keywords and associated TF-IDF scores of soil in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces	Farm (0.024), cattle farm (0.010), factory (0.009), stable (0.005), and pig farm (0.005).
Pressure	Human behavior pressures	Spreading slurries (0.044), spreading manure (0.027), spillage (0.017), dung heap (0.013), land spreading (0.011), tank (gas oil spillage) (0.010), burial (gas oil tank) (0.009), storage (0.009), oil spillage (0.007), substances (spillage) (0.005), jet of manure (0.005), and surface impoundment (0.004).
	Environmental pressures	Zootechnical effluents (0.018), sludge (0.011), gas oil (0.009), pollution (0.007), and fumes (0.004).
State	Environmental state (biotic state-living habitat)	Water contamination (by manure) (0.008), and stream contamination (by manure) (0.008).
	Environmental state (abiotic state-environment)	Fire (0.007) and paludification (0.005).
	Environmental state (biotic state-inhabitants)	Mosquitoes (0.011).
Impact	Human well-being (health and safety)	Odor (0.018), unpleasant smell (0.006), nauseating smell (0.006), strong smell (0.005) and discomfort (0.005).
Response	Driving forces-based responses	Request (0.016), police (0.016), control (0.011), intervention (0.011), and check (0.009).
	Pressure-based responses	Management (0.006) and compliance (distance) (0.006).
	Impact-based responses	Stop (0.005).

3.4.7. Keywords of weather-climate and sea-coast

With regard to weather-climate and sea coast issues, Arpa messages were not environmental complaints, but requests for “information” and “data”. Indeed, citizen mainly asked for data to support their personal “study” “research project”, “analysis”, and “insurance”. Among the impacts, we classified “modeling elaborations”, with regard to environmental impact assessments supporting ecosystem services (Table 15).

With regard to responses, complainants requested typical regulatory approaches in response to all environmental issues (“inspection”, “check”, and “controls”) or self-reporting to ensure compliance with regulations, with violations subject to penalties (fines and compensation payments). Additionally, some complainants suggested management actions such as the direct regulation of “waste disposal” and waste recovery and transport, or state-based responses such as the “remediation” of contaminated waters (e.g., rivers and aquifers), even if such measures were long-term and expensive and, in some cases, not even feasible.

Table 15
Keywords and associated TF-IDF scores of weather-climate and sea-coast in the DPSIR framework.

Category	Sub-category	Keywords (TF-IDF scores)
Driver	Economic driving forces (culture)	Study (0.012), research project (0.007), and analysis (0.005).
	Economic driving forces (security)	Insurance (0.006).
	Economic driving forces (infrastructure)	System (0.003).
State	Environmental state (abiotic state-physical and chemical)	High temperatures (0.012), hailstorm (0.008), and storm (0.007).
Impact	Ecosystem services (supporting)	Modeling elaborations (0.016)
Response	Driving forces-based responses	Data (0.052), web site (more user friendly) (0.030), rainfall data (0.020), wind data (0.019), and information (0.016).
	Pressure-based responses	Last data (0.015) and daily data (0.013).
	Impact-based Responses	Annual data (0.013), weather stations (data and localization) (0.013), weather forecast (0.013), temperature (data) (0.012), request (0.011), coordinates (0.008), maps (0.008), intensity (data rainfall or wind) (0.008), improvements in graphics (0.008), find (data) (0.007), receive (data) (0.006), know (data) (0.006), and answer (0.006).

3.5. Semantic network analysis of environmental complaints

Table 16 ranked the 30 most salient keywords according to both degree and betweenness centrality. The keywords of “odor”, “report”, “request”, “presence”, “municipality”, and “hours” were the most influential and meaningful concepts, as demonstrated by their high degree and betweenness centrality values. Thus, complainants were primarily concerned with odor, for which they sought a specific response (they made a “request” or submitted a “report”). The word “request” was extremely relevant to Arpa’s CRM, which aimed at responding quickly, succinctly, and accurately to complaints and requests for general information about policies, practices, and procedures. The word “municipality” represented another central hub, indicating two different semantic concepts: the first referring to the geographical content of the complaints (often very precise and local); and the second referring to the local authorities responsible for managing requests and complaints. In some cases, Arpa was not the competent authority to handle a complaint, and it instead delegated responsibility to different authorities or local municipalities. Duration or occurrence, in terms of the complaint topic, may be understood as a proxy for the magnitude of the concern raised (e.g., the keyword “hours” may have referred to the “presence” of a persistent or recurring compliance problem). Keywords such as “strong” and “odors” represented local hubs, with relatively high degree

Table 16
Most salient keywords according to degree and betweenness centrality.

Rank	keywords	Degree centrality	Keywords	Betweenness centrality
1	Odor*	69	Report	0.242
2	Report*	59	Odor	0.183
3	Presence*	40	Municipality	0.121
4	Municipality*	39	Private	0.089
5	Hours*	29	Presence	0.081
6	Air	27	Site**	0.066
7	Request	26	Request	0.062
8	Noise	25	House	0.054
9	Strong***	25	Noise	0.050
10	Coming from	24	Air	0.049
11	House	23	Hours	0.048
12	Present	21	Coming from	0.037
13	Factory	20	Present	0.035
14	Very much	20	Area	0.035
15	Area	20	Water	0.033
16	Odors***	19	Factory	0.032
17	Water	19	Data**	0.031
18	Days	19	Days	0.027
19	Pollution	17	Strong smell	0.025
20	Days	17	Strong	0.021
21	Site	16	Complainant	0.019
22	Data	16	Day	0.019
23	Canal	15	Very much	0.017
24	To smell	14	Canal	0.016
25	Complainant	14	Apartment	0.014
26	Private	13	Odors	0.013
27	Plastic	13	Wastes	0.012
28	Night	13	Pollution	0.012
29	Acoustic	12	Slurries	0.011
30	Acrid	11	River	0.001

Note: *, Hub (meaning circulation); **, bridging between communities; ***, local center (local hub).

centrality but low betweenness centrality. This suggested that these words were primarily related to odor and air pollution complaints. Finally, bridging concepts, with relatively high betweenness centrality but low degree centrality, included “site” and “data”. The keyword “site” referred to both the location of the event and Arpa website, where users searched for data or information regarding environmental topics. Furthermore, the keyword “data” were connected with many environmental issues (air and water quality, weather forecasts, rainfall, monitoring stations, etc.). The semantic network analysis provided a visual and quantitative method of representing and analyzing the hybrid connections among environmental driver, pressure, state, impact, and response. Indeed, environmental processes were visually interconnected and the all set of interactions were more diffuse than hierarchical (Cañas et al., 2012).

Table 17 reported the top 30 interrelationships between the environmental complaint keywords in the DPSIR framework. Keywords connecting odor (classified as impacts) and air pollution (classified as state) were the most important (such as “odor-burnt plastic” and “odor-acrid”). Complainants perceived odor annoyance as a primary environmental concern, possibly related to two main drivers: “odor-factory” and “odors-farms”. In addition, odor was related to pressure arising from industrial activities (“odor-smoke”) and agricultural practices (“odor-spreading manure” and “odor-spreading slurries”). It was noteworthy that, the grievances referred to an odor smelling like “burnt plastic”, might be related to specific productive sectors that potentially can produce olfactory annoyance such as ceramics with digital printing system. Indeed, olfactometric studies performed by Arpa, revealed that color application performed using an organic carrier, during the following firing, will turn into substances characterized by a low olfactory threshold, which are released through the ovens’ atmospheric emissions, often causing odorous issues (“burnt plastic” and “acrid”) in the surrounding area. For the Emilia-Romagna Region, such a manufacturing sector represented a key economic sector but, at the same time, might influence environmental pressures, particularly aldehydes emission, characterized by a high odorous charge (Capuano et al., 2018). Water complaints emphasized interconnections between the biotic state of the environment as a living habitat and pressures (“water-discharge”, “stream-foam”, “canal-spillage”, and “canal-foam”). The main water-related impacts on human well-being were identified as odor complaints: “water-smelly”, “smelly-fumes”, “foam-odor”, and “canal-odor”. Noise pollution complainants primarily requested intervention strategies to be executed by local authorities (“municipality-noise”), aimed at mitigating or eliminating the reported issue (“report-trouble”).

Table 17
Top 30 interrelationships between environmental complaint keywords in the DPSIR framework.

Category	Interrelationship	Weight
Impact-state	Odor-burnt plastic	65
	Odor-acrid	54
	Odor-air	36
	Strong smell-burnt plastic	18
Impact-response	Odor-report	21
Impact-driver	Odor-factory	18
	Odors-farms	13
Impact-pressure	Smelly-fumes	27
	Odor-smoke	13
	Odor-spreading manure	7
	Odor-spreading slurries	7
Pressure-driver	Pollution-installations	13
Pressure-response	Wastewater discharge-report	10
	Wastewater discharge-authorization	7
Pressure-impact	Foam-odor	8
	Pollution-noise	7
Response-impact	Report-noise	30
	Report-odors	14
	Municipality-noise	13
Response-state	Data-air	7
Response-pressure	Report-discharge	9
State-pressure	Water-discharge	19
	Stream-foam	13
	Canal-spillage	10
	Canal-foam	9
State-impact	Water-smelly	10
	Canal-odor	7
	Report-trouble	7
Driver-pressure	Installations-acoustic	12
	Factory-fumes	6

4. Theoretical and practical implications

The results of the present study have several theoretical and practical implications. In terms of the theoretical implications, the proposed methodology can be adopted to improve environmental management and engage citizen in policy-related activities.

First, keyword extraction applied to environmental complaints by TF-IDF helps to summarize the content of texts and recognize the main topics discussed. This is of relevance for public administration in the framework of the CRM.

Specifically, keyword extraction based on TF-IDF might provide an accurate entry point for the identification of proactive measures. For example, the present text mining revealed multiple reasons for air pollution or odor complaints, including facility emissions, waste management, livestock processing, and agriculture practices. Thus, local authorities responsible for regulating the impact of odor emissions (as well as Arpae) might define certain odor mitigation measures at the planning stage, including: (i) site assessment and building design for odor control; (ii) stock density planning; and (iii) the selection of different farms or fields at a geographical distance, to prevent the same receptors from being constantly affected and mitigate cumulative odor impacts from the routine spreading of farm waste. For these reasons, a visualizing map, positioning the location of complaints together with data collected by local authorities on a regular basis, might provide new and novel insights at a lower cost and with greater accuracy than new data collection approaches.

Furthermore, semantic network analysis might contribute to the identification of key nodes with a large number of connecting arcs, representing useful indicators that are likely to bear on a large number of issues (Niemeijer and de Groot, 2008). Practically, the identification of central hub keywords in a semantic network might support the selection and investigation of a set of keywords related to decision-making processes in the DPSIR framework. For example, the present analysis revealed that the central hub keyword “odor” was associated with the spreading of manure or slurries. Thus, relevant environmental complaints might relate to farming practices and management, including the use, production, and withdrawal or import of manure, as well as emissions indicators (ammonia, methane, and nitrous oxide) (FAO, 2003; European Environmental Agency (EEA), 2005b; OECD, 2006). Furthermore, because the DPSIR framework hinted at the dynamics of the system, it can provide the conceptual foundation for the development of mathematical simulation models for forecasting the effects of alternative decision-making scenarios on long-term sustainability and health of communities (Salem et al., 2021; Roy et al., 2022, 2023). Such analyses may support the development of a more integrated environmental knowledge base, as a prerequisite for a wider exploitation of analytics for relevant decision-making processes and management activities.

Second, the study represents the first to integrate a validated systems-thinking framework (the DPSIR framework) with environmental complaint data obtained by using text mining. Computational models for complex systems are inherently more complicated than a conceptual framework, but the DPSIR framework can serve to highlight key variables and relationships, for which parameter estimates and functions will be needed, and to identify areas where existing models may be appropriate or where new models are needed (Malmir et al., 2021). Addressing environmental challenges requires a process of problem structuring, to transform unstructured problems into ones that can be effectively addressed with sound evidence about ecological and social system structure and function (Lewison et al., 2016). This process also requires information about how citizen perceives and defines the issues. By helping to structure the analysis of complex systems, the DPSIR framework can act as a heuristic tool for complex system analysis. Theoretically, framing the environmental complaints within the DPSIR framework encourages the decision-makers to adopt a system approach and think about challenges to the problem within the larger system, even if in some cases the relationship among driver, pressure, state, impact, and response might appear obvious.

An example of a decision-making process associated with complaint handling refers to foundry cooperatives in Modena Province. As previously reported, despite regular compliance monitoring, many complaints regarding air pollution and odor issues in the present dataset revealed a relationship between these economic drivers (foundry cooperatives) and increased environmental pressures (emissions and fumes). From 2017 to 2022, citizens of Modena Province have taken several courses of action to find an effective remedy for their pollution-related complaints, including petitioning enforcement authorities and higher levels of government, appealing to the media, and engaging in collective action (demonstrations). Arpae took the frequency, duration, intensity, and hedonic tone of odor complaints into account, applying a dispersion modelling of odor emission rates to determine the number of hours of odor exposure to receptors in a year. Following the failure of other enforcement tools, in March 2022, Arpae decided to revoke permits to foundry cooperatives in Modena Province, thus definitively closing the installations (Arpae Regional Agency for Prevention, Environment and Energy Emilia-Romagna, 2022). This case provides a clear illustration of public involvement in decision-making processes.

However, the proposed framework is unable to identify all possible interactions among variables, as anthropogenic and natural systems are extremely complex, and they can be hardly represented by a set of keywords and indicators. Moreover, the categorization of keywords is a subjective process that has not yet been codified, so the keywords selected to represent environmental variables may differ from those determined by other methodologies. Furthermore, the DPSIR framework has biases toward the perspectives and concerns of several stakeholder groups, and social and economic aspects (Svarstad et al., 2008). It has been suggested to couple the DPSIR framework with analytical methods like multiple-criteria decision-making (MCDM) (Malekmohammadi and Jahanishakib, 2017), Analytic Hierarchy Process (AHP) (Neshat et al., 2014; Sun et al., 2016; Liu et al., 2018), and the structural equation model (Sun et al., 2018).

Nonetheless, the proposed approaches (methods and tools) to environmental issue monitoring and decision-making support can be used by different organizations (e.g., public, local authorities, companies, citizen associations, etc.) in different contexts, with minimal human effort, through the customization of application dictionaries and domain-specific keywords and topic clusters.

5. Conclusions and future research

The present study established a framework for the text mining of Italian environmental complaints and analyzed 1-year environmental complaints recorded by ArpaE in the Emilia-Romagna Region. The DPSIR framework was applied as a conceptual model to cluster the main keywords describing environmental complaints, using a “bottom-up” approach (from citizen to policy-makers and decision-makers). The proposed approach recognized that local communities (represented by municipalities) are critical to sustainable environmental management measures, as they determine what and when they will adjust. The present results showed that the majority of complaints were made by citizen and associated with air pollution and odor. The important drivers of environmental complaints included factories (particularly foundries and ceramic industries) and farms, which mainly affect human well-being. Inspections, controls, and management actions represent the main response activity requested by complainants.

Despite the significance of the findings, a limitation of the research is that the data cannot necessarily represent the environmental concerns of citizen who is not active in filing complaints to ArpaE. Furthermore, demographic variables were not included to control for bias in the geographical distribution of complaints, and the study analyzed data for only one year. Consequently, changes in the occurrence of complaints over time could not be addressed. For this reason, attention should be taken when generalizing the results, as the data might not be representative of all environmental complaints within the Emilia-Romagna Region or even in Italy, more broadly. Additionally, environmental complaints may be influenced by many factors (complainants’ education quality, digitalization, geographical location, age, gender, etc.).

Future work should aim at exploring how stakeholder perceptions change in particular communities, as strategies are incorporated to increase awareness and address environmental public health concerns. Over time, a greater understanding of environmental public health issues (such as those discussed in this paper) might improve communication to key segments of the public. The DPSIR framework might further allow ArpaE to refine and focus its communication strategies and information dissemination to reach out to stakeholders who may be particularly impacted by environmental management.

The development of mathematical simulation models requires integration and coordination across several disciplines including social, ecological, and health sciences. When fully parameterized, the DPSIR framework can support economic cost-benefit analysis of various responses (management actions) and the value of the impacts (ecosystem goods and services).

As the maturity of text mining tools, they are being combined with structured data analysis and reporting to better integrate into decision support processes and systems. Advanced analytics tools and well-designed platforms supporting multiple channels for the purpose of accepting requests can now rapidly identify and visualize public concerns at a global scale, and thereby address unknown, unappreciated, or underestimated drivers, pressures, and impacts.

Future implementations on ArpaE's complaint-handling process could be based on rules trained on classification and training rules on complaint keywords revealed by text mining, to realize the automation (or partial automation) of the complaint process. The goal is to train a machine learning approach that provides automated support for finding recurring patterns in the environmental complaints, thus providing a human employee with a starting point for ticket (claim) tagging, by classifying a record as a request for information, data, complaints, disservices, or environmental incidents, each of which should be handled separately and individually. We can use machine learning algorithms to automatically categorize and prioritize environmental complaints based on their content and other factors. This can help to reduce the workload of URP operators and ensure that tickets are handled more efficiently. Furthermore, machine learning algorithms might generate automated responses to environmental complaints based on the content of the ticket. This can help to reduce response times and improve the overall customer service experience.

To this end, the present research could be used as a valuable historical reference case. Future research could also explore topic-based sentiment analysis to monitor public opinion, study related discourse, and understand the impact of environmental issues on human health and well-being. Moreover, inspired by the literature on domain-specific aspect-based analysis (Distante et al., 2021; Faralli et al., 2021), sentiment analysis could be further developed by leveraging users’ emotions, as abstracted from their claims. This could automatically identify and distinguish complaint, which may express emotions such as fear or anger, or may include neutral emotions. In conclusion, the present study represents a preliminary step toward a better understanding of environmental complaints. ArpaE is interested in exploring this line of research, with the aim of enhancing citizen participation in decision-making and promoting sustainable development.

Finally, in order to make the research replicable and reproducible, the developed Orange projects and some project data are available at <https://github.com/fabimanservigi/ORANGE-ARPAE/>.

Authorship contribution statement

Fabiana Manservigi, Damiano Distante, and Stefano Faralli: conceptualization, methodology, writing-original draft, and editing. Michele Banzi, Giuseppe Bortone, Susanna Ricci, Tomaso Tonelli, and Paolo Veronesi: conceptualization, methodology, writing-review, and editing.

Declarations of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Aguwa, C., Olya, M.H., Monplaisir, L., 2017. Modeling of fuzzy-based voice of customer for business decision analytics. *Knowledge-Based Syst.* 125, 136–145.
- Al-Khoury, A.M., 2012. Customer relationship management: proposed framework from a government perspective times. *J. Manag. Strat.* 3 (4), 34–54.
- Arpaee (Regional Agency for Prevention, Environment and Energy) Emilia-Romagna, 2022. Fonderie Cooperative di Modena. <https://www.arpaee.it/il-territorio/modena/in-evidenza-a-modena/fonderie-cooperative-di-modena>. (Accessed 17 June 2023).
- Batagelj, A., Mrvar, V., 1998. Pajek-program for large network analysis. *Connections* 21 (2), 47–57.
- Bradley, P., 2015. Using the DPSIR Framework to Develop a Conceptual Model: Technical Support Document [2023-06-17]. <https://archive.epa.gov/ged/tutorial/web/html/slide0004-3.html>.
- Brewer, B., 2007. Citizen or customer? Complaints handling in the public sector. *Int. Rev. Adm. Sci.* 73 (4), 549–556.
- Bruno, M.F., Saponieri, A., Molfetta, M.G., et al., 2020. The DPSIR approach for coastal risk assessment under climate change at regional scale: the case of Apulian Coast (Italy). *J. Mar. Sci. Eng.* 8, 531. <https://doi.org/10.3390/jmse8070531>.
- Cañas, A.J., Novak, J.D., Vanhear, J., 2012. Concept Mapping and Environment as a Connection [2023-06-17]. <https://cmc.ihmc.us/cmc2012Papers/cmc2012-p26.pdf>.
- Capuano, F., de' Munari, E., Forti, S., 2018. Il controllo degli odori, nuova frontiera di ricerca. *Ecoscienza* 2, 50–51 (in Italian).
- Chen, Y.C., 2010. Citizen-centric e-government services: understanding integrated citizen service information systems. *Soc. Sci. Comput. Rev.* 28 (4), 427–442.
- Cook, G.S., Fletcher, P.J., Kelble, C.R., 2014. Towards marine ecosystem based management in South Florida: investigating the connections among ecosystem pressures, states and services in a complex coastal system. *Ecol. Indic.* 44, 26–39.
- Cormier, R., Kannen, A., Elliott, M., et al., 2013. Marine and Coastal Ecosystem-Based Risk Management Handbook. International Council for the Exploration of the Sea (ICES), Denmark, 1–60.
- Deidda Gagliardo, E., 2002. La Creazione Del Valore Nell'ente Locale. Giuffrè Publisher, Torino (in Italian), 1–456.
- Demsar, J., Turk, T., Erjavec, A., et al., 2013. Orange: data mining toolbox in Python. *J. Mach. Learn. Res.* 14 (1), 2349–2353.
- Distante, D., Faralli, S., Rittinghaus, S., et al., 2021. DomainSenticNet: an ontology and a methodology enabling domain-aware sentic computing. *Cogn. Comput.* 14, 62–77.
- EEA (European Environmental Agency), 2005a. Sustainable Use and Management of Natural Resources. EEA, Copenhagen, pp. 9–11.
- EEA, 2005b. EEA Core Set of Indicators—Guide, vols. 1–38. EEA, Luxembourg.
- EPA (Environmental Protection Agency), 1994. A Conceptual Framework to Support the Development and Use of Environmental Information. Environmental Statistics and Information Division-Making. United States Environmental Protection Agency, Washington DC, 8–17.
- European Commission, 1999. Towards Environmental Pressure Indicators for the EU, first ed. Office for Official Publications of the European Communities, Luxembourg, pp. 1–185.
- European Commission, Directorate-General for Environment, 2020. In: Environmental Compliance Assurance Vade Mecum. Publications Office of the European Union, Luxembourg, pp. 1–78.
- FAO (Food and Agriculture Organization), 2003. Data Sets, Indicators and Methods to Assess Land Degradation in Drylands. FAO, Rome, pp. 99–102.
- Faralli, S., Rittinghaus, S., Samsami, et al., 2021. Emotional intensity-based success prediction model for crowd-funded campaigns. *Inf. Process. Manag.* 58 (1), 102394. <https://doi.org/10.1016/j.ipm.2020.102394>.
- Gabrielsen, P., Bosch, P., 2003. Environmental Indicators: Typology and Use in Reporting. European Environmental Agency, Copenhagen, 1–20.
- Gebremedhin, S., Getahun, A., Anteneh, W., et al., 2018. A Drivers-Pressure-State-Impact-Responses framework to support the sustainability of fish and fisheries in Lake Tana, Ethiopia. *Sustainability* 10, 2957. <https://doi.org/10.3390/su10082957>.
- Ghodousi, M., Alesheikh, A.A., Saeidian, B., et al., 2019. Evaluating citizen satisfaction and prioritizing their needs based on citizens' complaint data. *Sustainability* 11 (17), 4595. <https://doi.org/10.3390/su11174595>.
- Harris, Z., 1954. Distributional structure. *Word* 10 (23), 146–162.
- Hartmann, S., Mainka, A., Stock, W.G., 2017. Citizen relationship management in local governments: the potential of 311 for public service delivery. In: Paulin, A., Anthopoulos, L., Reddick, C. (Eds.), *Beyond Bureaucracy. Public Administration and Information Technology*. Springer, Berlin, pp. 337–353.
- ISTAT (Italian national statistical institute), 2023. Resident Population on 1st January: Emilia-Romagna [2023-06-17]. <https://www.istat.it/en/population-and-households?>
- Joung, J., Jung, K., Ko, S., et al., 2019. Customer complaints analysis using text mining and Outcome-Driven Innovation method for market-oriented product development. *Sustainability* 11 (1), 40. <https://doi.org/10.3390/su11010040>.
- Kannabiran, G., Xavier, M.J., Anantharaaj, A., 2004. Enabling e-governance through citizen relationship management: concept, model and applications. *J. Serv. Res.* 4, 223–240.
- Katir, O., Sakalli, K., Armutlu, H., et al., 2020. Determination of customer satisfaction by text mining: case of Cappadocia Hotels. *Journal of Business Researh-Turk.* 12, 546–556.
- Khadir, A.C., Aliane, H., Guessoum, A., 2021. Ontology learning: grand tour and challenges. *Comput. Sci. Rev.* 39, 100339. <https://doi.org/10.1016/j.cosrev.2020.100339>.
- Khunanake, W., Pradatsudara, A., Pattanakiat, S., 2018. Stakeholder involvement in developing environmental indicators for the lam nam yang Part 1 watershed in the northeastern Thailand. *Applied Environmental Research* 40 (3), 28–41.
- Larson, S., Stone-Jovicich, S., 2011. Community perceptions of water quality and current institutional arrangements in the Great Barrier Reef Region of Australia. *Water Pol.* 13 (3), 411–424.
- Lewison, R.L., Rudd, M.A., Al-Hayek, W., et al., 2016. How the DPSIR framework can be used for structuring problems and facilitating empirical research in coastal systems. *Environ. Sci. Policy.* 56, 110–119.
- Liu, X., Liu, H.T., Chen, J.C., et al., 2018. Evaluating the sustainability of marine industrial parks based on the DPSIR framework. *J. Clean. Prod.* 188, 158–170.
- Loomis, D.K., Paterson, S.K., 2014. Human dimensions indicators of coastal ecosystem services: a hierarchical perspective. *Ecol. Indic.* 44, 63–68.
- Lucini, F.R., Tonetto, L.M., Fogliatto, F.S., et al., 2020. Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *J. Air Transp. Manag.* 83. <https://doi.org/10.1016/j.jairtraman.2019.101760>.
- Malekmohammadi, B., Jahanishakib, F., 2017. Vulnerability assessment of wetland landscape ecosystem services using driver-pressure-state-impact-response (DPSIR) model. *Ecol. Indic.* 82, 293–303.
- Malmir, M., Javadi, S., Moridi, A., et al., 2021. A new combined framework for sustainable development using the DPSIR approach and numerical modeling. *Geosci. Front.* 12 (4), 260–273.
- Motola, V., Colonna, N., Alfano, V., et al., 2009. Censimento Potenziale Energetico Biomasse, Metodo Indagine, Atlante Biomasse su WEB-GIS. Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Roma, pp. 1–140.
- Mozumder, M.M.H., Pyhälä, A., Wahab, M.A., et al., 2019. Understanding social-ecological challenges of a small-scale Hilsa (*Tenualosa ilisha*) Fishery in Bangladesh. *Int. J. Environ. Res. Public Health.* 16 (23), 4814. <https://doi.org/10.3390/ijerph16234814>.
- Neshat, A., Pradhan, B., Dadras, M., 2014. Groundwater vulnerability assessment using an improved DRASTIC method in GIS. *Resour. Conserv. Recycl.* 86, 74–86.
- Niemeijer, D.S., de Groot, R., 2008. A conceptual framework for selecting environmental indicator sets. *Ecol. Indic.* 8, 14–25.
- Obubu, J., Odong, R., Alamerew, T., 2022. Application of DPSIR model to identify the drivers and impacts of land use and land cover changes and climate change on land, water, and livelihoods in the L. Kyoga basin: implications for sustainable management. *Environ Syst Res* 11, 11. <https://doi.org/10.1186/s40068-022-00254-8>.
- OECD (Organisation for Economic Cooperation and Development), 1993. OECD Core Set of Indicators for Environmental Performance Reviews. A Synthesis Report by the Group on the State of the Environment. OECD Publishing, Paris, 5–11.
- OECD, 2005. Integrated Environmental Permitting Guidelines for EEECA Countries. OECD Publishing, Paris, pp. 132–138.

- OECD, 2006. Environment at a Glance: OECD Environmental Indicators. OECD Publishing, Paris, pp. 127–135.
- Patrício, J., Elliott, M., Mazik, K., et al., 2016. DPSIR—Two decades of trying to develop a unifying framework for marine environmental management? *Front. Mar. Sci.* 3, 177. <https://doi.org/10.3389/fmars.2016.00177>.
- Ponte, V., 2015. CiRM: CRM no setor público. In: Demo, G. (Ed.), *Marketing de Relacionamento & Comportamento do Consumidor*. São Paulo, Atlas, pp. 137–174 (in Portuguese).
- Rapport, D., Friend, A., 1979. *Towards a Comprehensive Framework for Environmental Statistics: A Stress-Response Approach*. Minister of Supply and Services Canada, Ottawa, 27–67.
- Richter, P., Cornford, J., 2007. Customer relationship management and citizenship: technologies and identities in public services. *Soc. Policy Soc.* 7 (2), 211–220.
- Roche, C., 2011. *Network Analysis of Semantic Web Ontologies*, vol. 13. Stanford University, California.
- Roy, S., Bose, A., Majumder, S., et al., 2022. Evaluating urban environment quality (UEQ) for Class-I Indian city: an integrated RS-GIS based exploratory spatial analysis. *Geocarto Int.* <https://doi.org/10.1080/10106049.2022.2153932>.
- Roy, S., Basak, D., Bose, A., et al., 2023. Citizens' perception towards landfill exposure and its associated health effects: a PLS-SEM based modeling approach. *Environ. Monit. Assess.* 195, 134. <https://doi.org/10.1007/s10661-022-10722-4>.
- Salem, M., Bose, A., Bashir, B., et al., 2021. Urban expansion simulation based on various driving factors using a logistic regression model: Delhi as a case study. *Sustainability* 13 (19), 10805. <https://doi.org/10.3390/su131910805>.
- Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.* 24 (5), 513–523.
- Sami, A., Irfan, A., Qureshi, M.I., et al., 2021. Sustainable public value: a step towards green public organisation for a sustainable society. *Int. J. Innov. Sustain. Dev.* 15 (2), 223–233.
- Secchi, L., 2009. Modelos organizacionais e reformas da administração pública. *Rev. Adm. Pública* 43 (2), 347–369 (in Portuguese).
- Semeoshenkova, V., Newton, A., Rojas, M., et al., 2017. A combined DPSIR and SAF approach for the adaptive management of beach erosion in Monte Hermoso and Pehuen Co (Argentina). *Ocean Coastal Manage* 143, 63–73.
- Smeets, E., Weterings, R., 1999. *Environmental Indicators: Typology and Overview*. European Environment Agency (EEA), Copenhagen, p. 19.
- Sun, S.K., Wanga, Y.B., Liu, J., et al., 2016. Sustainability assessment of regional water resources under the DPSIR framework. *J. Hydrol.* 532, 140–148.
- Sun, C.Z., Wu, Y.J., Zou, W., et al., 2018. A rural water poverty analysis in China using the DPSIR-PLS model. *Water Resour. Manag.* 32 (6), 1933–1951.
- Svarstad, H., Petersen, L.K., Rothman, D., et al., 2008. Discursive biases of the environmental research framework DPSIR. *Land Use Pol.* 25, 116–125.
- Tscherning, K., Helming, K., Krippner, B., et al., 2012. Does research applying the DPSIR framework support decision making? *Land Use Pol.* 29 (1), 102–110.
- Twizeyimana, J.D., Andersson, A., 2019. The public value of E-Government—A literature review. *Gov. Inf. Q.* 36 (2), 167–178.
- Unep (United Nations Environment Programme), 1994. *World Environment Outlook: Brainstorming Session [2023-07-17]*. [https://www.unep.org/events/Environment \(Assessment Programme. Nairobi\)](https://www.unep.org/events/Environment%20(Assessment%20Programme%20Nairobi)).
- Zhan, J., Loh, H.T., Liu, Y., 2009. Gather customer concerns from online product reviews—A text summarization approach. *Expert Syst. Appl.* 36 (2), 2107–2115.