Article

Cross-Societal Analysis of Climate Change Awareness and Its Relation to SDG 13: A Knowledge Synthesis from Text Mining

Hansu Hwang [1] **, SeJin An** [1] **, Eunchang Lee** [1,†] **, Suhyeon Han** [1] **and Cheon-hwan Lee** [1,2,*]

[1] Division of Technical Supervision, Green Technology Center (GTC), Seoul 04554, Korea; hshwang@gtck.re.kr (H.H.); sejin_an@gtck.re.kr (S.A.); ecleesia2@gmail.com (E.L.); sue@gtck.re.kr (S.H.)
[2] Program in Science and Technology Studies, Korea University, Seoul 02841, Korea
* Correspondence: chlee@gtck.re.kr; Tel.: +82-2-3393-3974
† Current address: Triplelight Co., Seoul 04520, Korea.

Abstract: The awareness and the engagement of various stakeholders play a crucial role in the successful implementation of climate policy and Sustainable Development Goals (SDGs). SDG 13, which refers to climate action, has three targets for combating climate change and its impact. Among the three targets, SDG 13.3 aims to "improve education, awareness-raising and human and institutional capacity on climate change mitigation, adaptation, impact reduction, and early warning". This target should be implemented based on the understanding of climate change awareness among various groups of societies. Furthermore, the indicator related to awareness-raising is absent in SDG 13.3. Hence, this study aims to explore the differences in climate change awareness among various social groups within a country from a text mining technique. By collecting and analyzing a large volume of text data from various sources, climate change awareness was investigated from a multilateral perspective. Two text analyses were utilized for this purpose: Latent Dirichlet Allocation (LDA) topic modeling and term co-occurrence network analysis. In order to integrate and comparatively analyze the awareness differences among diverse groups, extracted topics were compared by classifying them into four indicators derived from the detailed targets in SDG 13.3: mitigation, adaptation, impact reduction, and early warning. The results show that the Korean public exhibited a relatively high awareness of early warning compared to the other four groups, and the media dealt with climate change issues with the widest perspective. The Korean government and academia notably had a high awareness of both climate change mitigation and adaptation. In addition, corporations based in Korea were observed to have substantially focused awareness on climate change mitigation for greenhouse gas reduction. This research successfully explored the disproportion and lack of climate change awareness formed in different societies of public, social, government, industry, and academic groups. Consequently, these results could be utilized as a decision criterion for society-tailored policy formulation and promoting climate action. Our results suggest that this methodology could be utilized as a new SDG indicator and to measure the differences in awareness.

Keywords: climate change; SDGs; text mining; topic modeling; network analysis

1. Introduction

The adverse impacts of climate change have become more noticeable worldwide, the evidence of which, including rising sea levels [1], melting glaciers [2], increasing wildfires, and changing biodiversity, has been observed all over the world [3]. To respond to climate change, all the parties (195 member countries) in the United Nations Framework Convention on Climate Change (UNFCCC) committed to the Paris Agreement in December 2015 with the aim of limiting global warming to well below 2 °C above pre-industrial levels and pursuing efforts to limit warming to 1.5 °C [4,5]. Notably, the Paris Agreement provides an ambitious opportunity to consolidate the relationship between climate and development [6]. In the same year, the 2030 Agenda for sustainable development was

also adopted by the United Nations General Assembly. Under the name of Sustainable Development Goals (SDGs), 17 goals and 169 targets were established along with economic growth, social inclusion, and environmental protection [7,8]. These collaborative and actionable goals can provide new opportunities for addressing key challenges confronting humanity [9]. Among the 17 Sustainable Development Goals, SDG 13 refers to climate action: "Take urgent action to combat climate change and its impacts" [10,11]. Therefore, the Paris Agreement was the first and major worldwide step in strengthening SDG 13 [12]. Even though all 17 SDGs are closely associated with each other, goal 13 (climate action) is notably more interconnected with the other SDGs [13].

All the SDGs have their targets and indicators, designed to monitor and track each implementation with measurable data. SDG 13 (climate action) aims to "take urgent action to combat climate change and its impacts". The detailed targets of SDG 13 are: (13.1) strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries; (13.2) integrate climate change measures into national policies, strategies, and planning; and (13.3) improve education, awareness-raising, and human and institutional capacity on climate change mitigation, adaptation, impact reduction, and early warning. Like other SDGs, specific indicators were also established in detailed SDG 13 targets, and most of these indicators have been collected as measurable data [13,14].

The indicators in SDG 13.1 consist of "number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population", "number of countries that adopt and implement national disaster risk reduction strategies in line with the Sendai Framework for Disaster Risk Reduction 2015–2030", and "proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies". These three indicators have been collected for measurement of the resilience and adaptive capacity. The indicators in SDG 13.2 consist of "number of countries with nationally determined contributions, long-term strategies, national adaptation plans, strategies as reported in adaptation communications and national communications" and "total greenhouse gas emissions per year". These two indicators have also been collected for measurement of the national policy, strategies, and planning. On the contrary, the indicators in SDG 13.3 are insufficiently established compared to 13.1 and 13.2; in particular, there are no indicators for awareness-raising. SDG 13.3 has only one indicator of the "extent to which (i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment", which is biased on education and human and institutional capacity.

One target in SDG 13 is "improving education, awareness-raising, and human and institutional capacity on climate change mitigation, adaptation, impact reduction, and early warning" (SDG target 13.3) [10,11]. To achieve this goal, understanding the awareness and perception of climate change from various perspectives among diverse societies will play a crucial role in the successful implementation of climate policies and strategies. Climate action tends to be directly influenced by awareness and perception, not only at the individual levels but also at the level of scientists and governments [15,16]. Although there are some studies on identifying awareness and perception with respect to climate change at the cross-national level, several limitations, such as restrictive measurement, time-intensiveness, and geographical and cultural differences among countries, remain [17,18]. Furthermore, investigation of awareness and perception from a cross-societal perspective within a single country is also important. Thus, different levels of awareness of climate change in a specific country should be considered to facilitate a wide engagement of individuals and communities for urgent climate action.

Recently, vast amounts of digitalized data have been continuously produced in the name of big data owing to the rapid progress of information and communications technology (ICT) and digital technology [19]. Accordingly, the synthesis of useful knowledge and the discovery of valuable insights from such a large volume, variety, and velocity of data have emerged as a big challenge [20]. Since the processing and analysis of big data are not

limited to the amount and form of data, the analysis of unstructured data such as images and text is considered a great challenge, whereas traditional statistical analysis usually focuses on numeric datasets [21,22].

Generally, abundant and valuable information is underlying in the textual data [23], and many studies have been devoted to the extraction of latent semantic knowledge utilizing text mining techniques in various fields. For example, some research has focused on the investigation of public opinion through the collection and analysis of text data from social media [24], or the identification of trends and current topics by analyzing scientific literature or patents [25,26]. In addition, text analysis of news data has been utilized as an instrument for monitoring drought impacts [27]. Even though text analysis has been widely used in various disciplines, including social science, political science, computer science, information science, biological science, and environmental science [28,29], most studies have focused on specific text data from a single source. Since climate change issues are complicatedly intertwined, however, identifying awareness from a multilateral perspective should be considered a key factor for successful policymaking [17,30]. To address this challenge, this study aims to understand climate change awareness from a cross-societal perspective through a text analysis based on multi-source data from different societies.

The objective of the present research is to analyze the gap of climate change awareness among different societal groups in support of SDG 13.3 due to the lack of the measurable indicators for awareness-raising. Simultaneously, research results could provide the evidence to promote policy and action for raising climate change awareness. For these purposes, text analysis based on an unsupervised learning algorithm was utilized on multiple sources produced by different societal groups. Five different societies were selected as decisive actors responding to climate change. Social media, news, national R&D data, patents, and scientific articles were selected as the analytic objectives representing public, social, governmental, industrial, and academic awareness. Text analyses, including topic modeling and co-occurrence networks, were applied to the collected text data. Latent Dirichlet Allocation (LDA) topic modeling was used to verify the distribution of climate change awareness from diverse perspectives in each society and a term co-occurrence network analysis was employed for content analysis in detail. Consequently, the differences in climate change awareness among various societies were successfully investigated using this method. To the best of our knowledge, this is the first attempt to compile separately analyzed LDA results using a unified indicator in the field of text mining research. This approach could be helpful in planning and policy-making for raising awareness of climate change.

2. Materials and Methods

2.1. Data Collection

Five different text data were collected. Social media data were collected through a Twitter application programming interface (API) using Python. News data were collected from the BIG KINDS website (https://www.bigkinds.or.kr/, accessed on 22 March 2021), one of the largest platforms providing metadata from news articles in Korea. The search string "climate change" in Korean was used to retrieve text data containing climate change awareness information. Both sets of data, generated over one year in 2017, were collected in July 2020. The amount of collected data from each was 7632 tweets and 7634 news articles, respectively.

For scientific articles, publications written by Korean authors can be divided into two: the Korean Citation Index (KCI) category, written in the Korean language, and the Science Citation Index (SCI)/Social Sciences Citation Index (SSCI) journal categories, mostly written in English. Accordingly, the data from SCI and SSCI journals were collected from the Web of Science database (DB), and those from Korean journals were collected from the KCI DB. The search string "climate change" was also used in both platforms, and the dataset was based on 2017 publications by authors based in Korean institutions. Most KCI journals recommend, by some mandate, the submission of an additional English abstract, although all the content is written in Korean. Hence, English abstracts were additionally

collected from the KCI website by web-scrapping tools using Python, since they are not provided in the KCI DB platform. In total, 920 publications were collected from both DBs.

The Cooperative Patent Classification (CPC) system was used to collect Korean corporations' patent information. The CPC, a well-known classification system developed by the EU and US Patent Office, contains more detailed information and improved retrieval performance than the International Patent Classification (IPC) system [31,32]. A dataset based on the patents granted in 2017 by the Korean corporation was built using the CPC Y-code by web-crawler using python. The Y-codes used here are Y02 (technologies or applications for mitigation or adaptation against climate change) and Y04 (information or communication technologies having an impact on other technology areas), which are related to climate change technology [33]. In total, 7169 documents were collected from patents DB.

National R&D data were used from the National Science and Technology Information Service (NTIS), which is a specific Korean platform for managing all national R&D information. The NTIS R&D data related to climate change were collected from a national research institution, the Green Technology Center (GTC), managing climate change R&D under the Ministry of Science and ICT in Korea. In total, 9591 R&D data conducted in 2017 were collected.

2.2. Topic Modeling

Topic modeling was employed to extract latent topics from a large volume and variety of collected documents. Topic modeling is a type of unsupervised machine learning technique, which requires input and interpretation from the researcher [34]. It relies on a probabilistic model of input data, which enables the estimation of the probability of latent topics based on the frequencies and co-occurrences of words [35]. The LDA is the most widely used topic modeling algorithm. Basically, the LDA approach assumes that a single document is a mixture of several topics and each topic is characterized as a probability distribution over words [36]. At the same time, LDA requires a fixed number of topics, defined as the K value, for analysis. [37] In order to determine the number of topics (K value), perplexity, one of the measures for the goodness of fit of statistical models, was evaluated on each text dataset. [35,38]. The perplexity was calculated at a range of K value from 10 to 50. The number of topics (K) was finally determined to be 20, since the lowest perplexity value was evaluated at K = 20 for tweets, R&D, and patent data. On the other hand, the perplexity evaluated from news and scientific articles tended to increase as increasing the number of topics. However, no significant differences in perplexity were observed below the topic number of 30. Thus, the topic number was decided in accordance with the results of the other three data.

A document-term matrix (DTM), the most basic step for most text analyses, should be established before employing LDA. The DTM is composed of a set of documents in a row and terms in a column, indicating the frequency of the terms in each document. The primary procedure for this is pre-processing, which includes tokenization, removal of stop words, and stemming. Subsequently, the pre-processed data is converted to a DTM. The respective analytic text data were utilized for tweets from social media data, the full text of news articles, a purpose section of national R&D data, and a background and technical problem section of patent and abstract of scientific article, respectively. Afterward, the term frequency-inverse document frequency (TF-IDF) value was calculated to determine the weight of each term. This value indicates the importance of a given term in a document [39]. The LDA model was then applied to the TF-IDF matrix for topic extraction; the number of topics (K) was determined to be 20 for this study in order to multilaterally analyze on exclusive multi-data. The statistical software R was utilized for all these procedures. Overall analytic procedures were conducted in accordance with the Korean natural language process (NLP), except for scientific articles, since most of the documents were written in Korean [40,41].

2.3. Co-Occurrence Network Analysis

The same text data used in LDA topic modeling were employed to build a term co-occurrence network map containing climate change awareness from each of the five different societies. The weight of terms was determined using TF-IDF after the extraction of nouns from text documents. Subsequently, the correlation between terms was extracted based on the frequency of a pair of terms occurring together in a document. This process also used R software. Using this value and pairs of terms, a co-occurrence network was constructed using the Gephi software, a well-known network construction and analysis tool [42]. Approximately 125 terms were used to construct the network maps, depending on the degree of nodes. However, the co-occurrence network of tweet data consisted of 79 terms because the total text data was less than that of the other four cases. The limited amount of text information is a general characteristic and limitation of text mining from social media data [43,44]. The nodes are composed of co-occurring terms and the edges represent a correlation value to indicate how the two terms are associated with each other. The size of nodes is proportional to the degree, indicating that the larger size of the node is assigned to the larger numbers of co-occurrence with other terms. The thickness of the edges is proportional to the strength of the correlation between a pair of terms. After constructing the network, the clusters were determined by a modularity algorithm provided by Gephi software and the color of the nodes was distinguished by clusters. The position and distance of each node were modulated by the MultiGravity ForceAtlas2 layout algorithm, also provided by Gephi [45]. Each cluster was labeled manually based on the mainly used and highly relevant terms in a given cluster.

3. Results

The research framework in this study is illustrated in Figure 1. To investigate climate change awareness in five different societies, massive text documents from multiple sources as generated by each society were collected and analyzed. Five different societal groups were selected as decisive actors who affect climate change policies: public, social, government, industrial, and academic groups. The respective text data were utilized under the assumption that each text dataset encompasses awareness and perception of each society [46]. For this study, social media data (Twitter) were utilized to identify public awareness, news data were utilized for social awareness, national R&D data were utilized for governmental awareness, patents were utilized for industrial awareness, and scientific articles were utilized for academic awareness. Each dataset was collected from a specific DB platform or web scrapping, as described in detail in the Methods section. Since articles, patents, and R&D data generally contain a large volume of text information, particular sections of each document were analyzed in order to focus on the elicitation of awareness contained in each document. Accordingly, the purpose section was utilized for the analysis of national R&D, the background and technical problem sections were utilized for patent analysis, and the abstract was utilized as an analytic objective in the case of scientific articles.

After retrieving and pre-processing the analytic data, topic analysis based on the LDA model and term co-occurrence network analysis were separately performed on the respective groups' text data. This approach is undoubtedly helpful for a good understanding of the status of cross-societal awareness in a specific country, which in this paper was determined to be Korea as a case study. Finally, the analysis results were translated from Korean to English to share this worthwhile methodology with many other readers and researchers. Note that unavoidable difference in nuance between Korean and English words may exist by translation, which was conducted in consideration of interchangeable words with similar or identical meanings as well as co-occurring terms. On the other hand, scientific articles were analyzed using English abstracts regardless of Korean and English journal sources because both journal categories provide English abstracts.
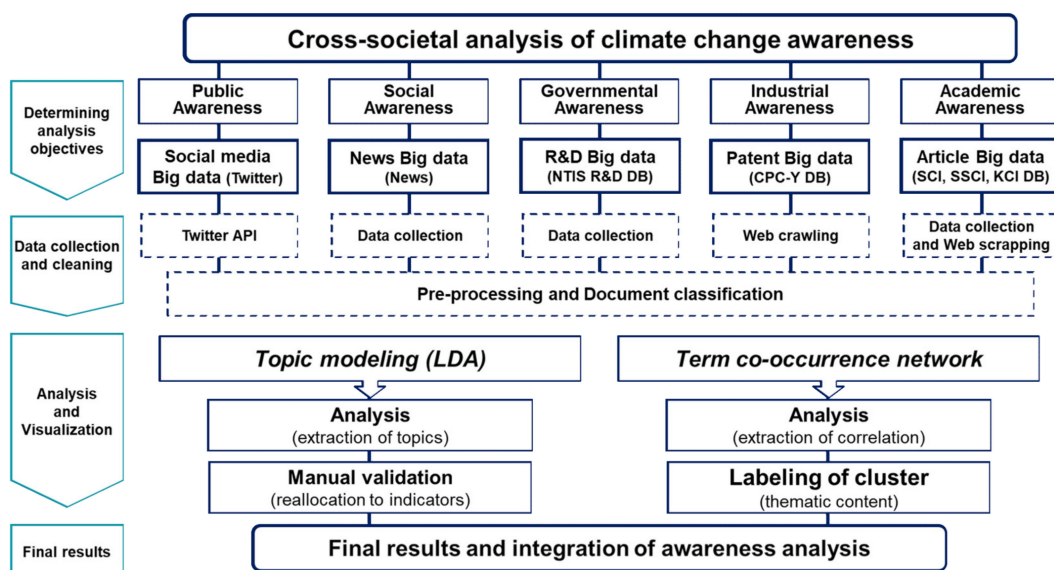
Figure 1. Schematic of the research framework of this study.

To identify the holistic awareness perspectives on climate change among different societies, topic modeling based on LDA was employed to collected text documents. As a promising unsupervised learning technique, the LDA topic model can be beneficially utilized for detecting latent topics from massive unstructured documents and classifying documents based on patterns of latent topics [47]. The LDA method was applied discretely to collected text data containing public, social, governmental, industrial, and academic awareness. This study not only aimed to extract of topics and classify documents but was also devoted to inferring how the latent topics are correlated with detailed targets of SDG 13.3 in terms of awareness of climate change. To this end, a manual validation process was additionally performed on the discerned topics in order to evaluate the relationship between each topic and each target in SDG 13.3. There are four detailed targets in SDG 13.3: "mitigation", "adaptation", "impact reduction", and "early warning" with respect to climate change. These four targets were utilized as awareness indicators to reallocate extracted topics in the manual validation process. Twenty topics were extracted from each societal document and subsequently reallocated to one of the detailed SDG 13.3 targets. The reallocation of each topic was evaluated based on relevant terms in a given topic. The awareness distribution from topic modeling with the manual validation process is illustrated in Figure 2. The detailed results of LDA are summarized in Tables A1–A5.
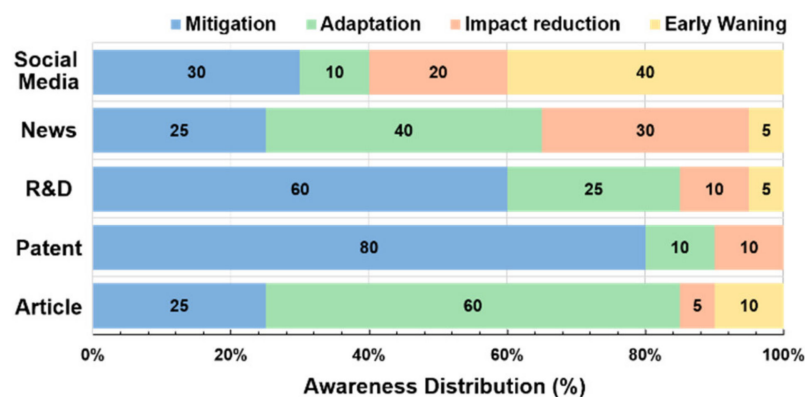

Figure 2. Climate change awareness distribution with respect to four indicators (mitigation, adaptation, impact reduction, and early warning) from five different text datasets.

As shown in Figure 2, the status of climate change awareness in each society can be comprehensively verified based on topic distribution. The climate change awareness differences were clearly observed depending on the groups of societies. Based on social media (Twitter) analysis, it was identified that the awareness of the Korean public of climate change largely focused on early warning (40%) and mitigation (30%) in 2017; less awareness of adaptation (10%) and impact reduction (20%) was also observed. News data corresponding to social awareness showed evenly distributed topics on all four awareness indicators despite a low early warning ratio (5%). Awareness of mitigation, adaptation, and impact reduction was observed to be 25%, 40%, and 30%, respectively, which implies that Korean media have comparably wider perspectives on climate change issues. Governmental and academic awareness derived from national R&D and scientific articles mostly focused on mitigation and adaptation topics with a high proportion of 85% all together. However, national R&D focused more on mitigation (60%) research, while scientific articles exhibited adaptation-related topics further with 60% ratio. Meanwhile, R&D data and academic articles showed that awareness of impact reduction and early warning lagged in both groups, inferred from the low proportion of awareness with an overall 15% ratio. Noticeably, industrial awareness represented by corporate patents was observed to be mostly concentrated on mitigation technology (80%). This indicates that the Korean industry paid attention to reducing greenhouse gas emissions for climate change mitigation. However, it is recognized that other awareness targets, such as adaptation, impact reduction, and early warning, fell behind compared to mitigation: in particular, the early warning awareness was not observed in industrial groups.

In the above section, the holistic distribution of topics from respective documents was verified to identify climate change awareness in each group through LDA topic modeling. Additionally, one more useful technique to extract valuable insights from a large volume of text data is the term co-occurrence network [48]. As a kind of content analysis, the term co-occurrence network is useful to understand the underlying content structure and relatedness in detail between co-occurring terms constituted in the document [49,50]. Visualization of the data, particularly the co-occurrence network, enables intuitive human cognition improvement [10,51]. In this regard, a co-occurrence network is advantageous for visualizing the association and patterns between items (terms) as a two-dimensional map [6,48]. Furthermore, it could provide a sub-network, called a community or cluster, which indicates aggregates of items (terms) with similarity. In other words, it implies connections between co-occurring terms within a particular theme. Term co-occurrence networks constructed using documents from each society are illustrated in Figures 3–7, and the top 5 terms as a function of degree are summarized in Table 1. The degree is defined as the number of links to a given node. High degree value indicates high co-occurring frequency with other terms, which represents the importance of a term. First, social media data exhibited five clusters having two main words "warmth" and "Earth" in the center of the network with the highest link degree, as shown in Figure 3. The five clusters composing the network were "US withdrawal from the Paris Agreement", "abnormal and extreme climate", "climate change impacts and mortality", "sea-level rise arising from global warming", and "national policy and forum on climate change response". The top three terms were "rise", "sea-level", and "case" in order. The top five terms based on degree indicate the strongest association with other terms; in other words, they frequently co-occur with other terms. Hence, these words could be considered major thematic keywords dealt with importantly by the actors in the corresponding society. The important thematic keywords dealt with in social media data are related to "warmth of the earth (indicating 'global warming')", "sea-level rise", and "cases (for abnormal climate and warming)".

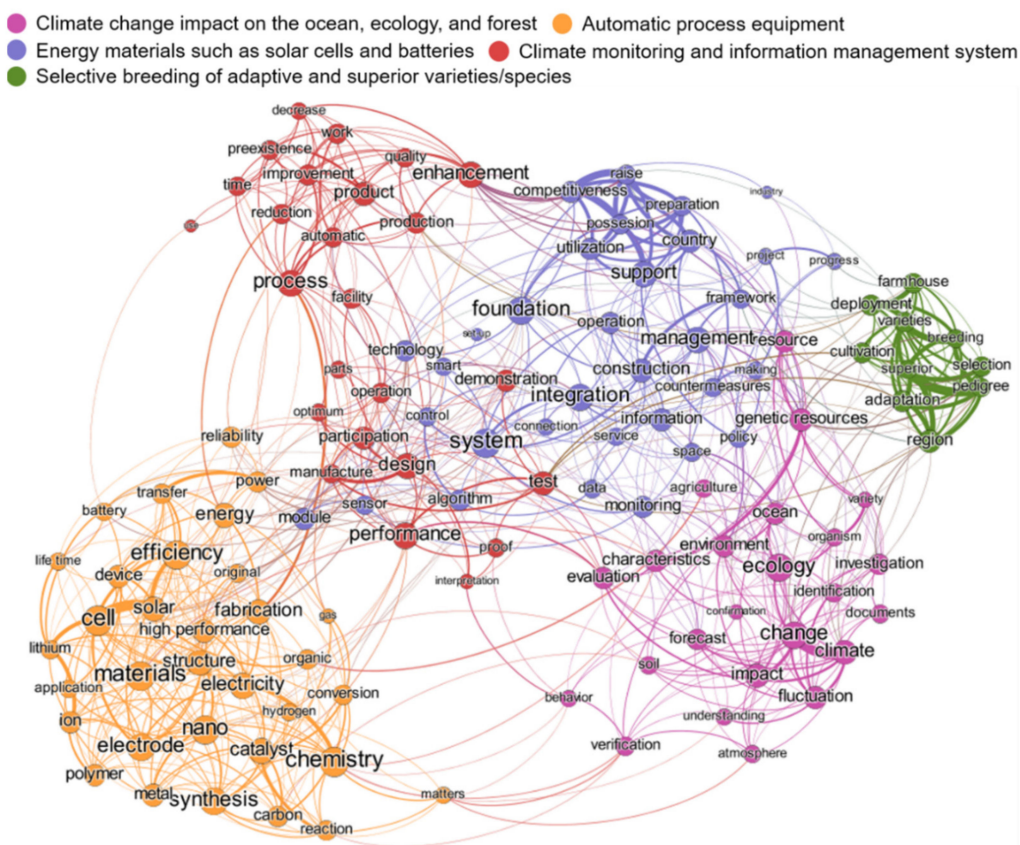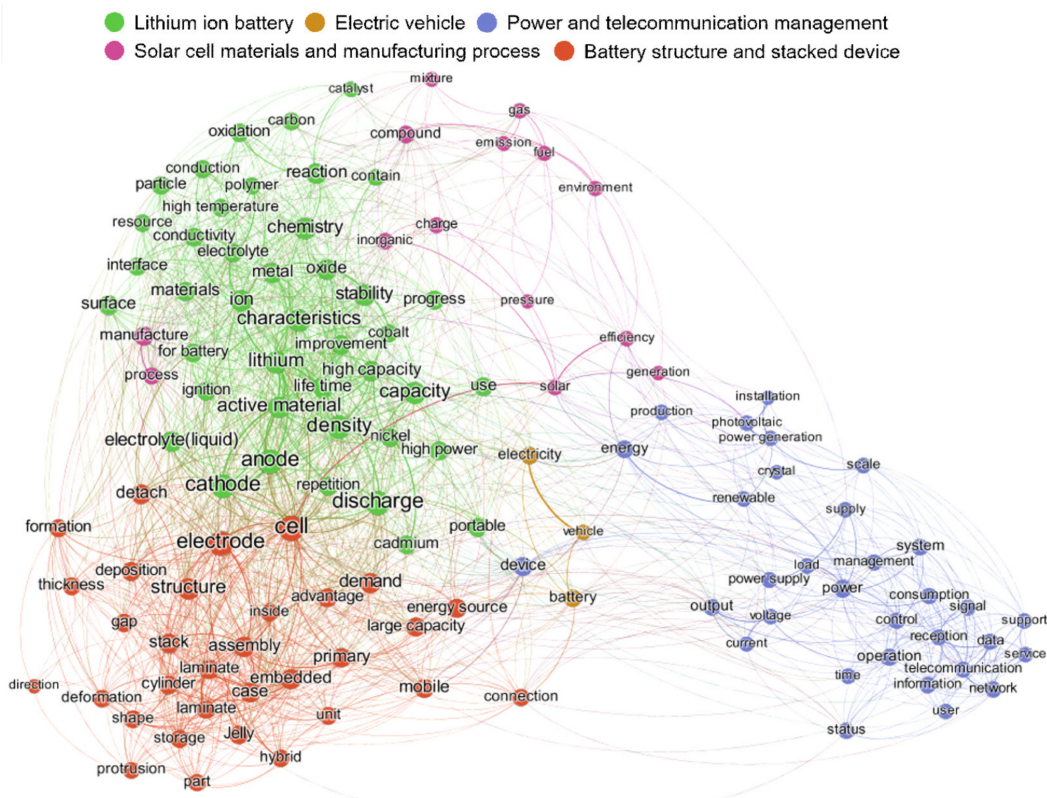**Figure 3.** Term co-occurrence network from social media data.

Figure 4. Term co-occurrence network from news data.

Figure 5. Term co-occurrence network from national R&D data.

Figure 6. Term co-occurrence network from patent data.

Figure 7. Term co-occurrence network from scientific article data.

Table 1. Top five thematic keywords depending on the degree in respective term co-occurrence network.

Groups	Rank	Terms	Cluster	Degree
Social Media	1	warmth	Abnormal and extreme climate	42
	2	Earth	Abnormal and extreme climate	39
	3	rise	Sea-level rise arising from global warming	26
	4	sea-level	Sea-level rise arising from global warming	26
	5	case	Abnormal and extreme climate	24
News	1	USA	US withdrawal from the Paris Agreement	36
	2	GHGs [1]	US withdrawal from the Paris Agreement	30
	3	China	US withdrawal from the Paris Agreement	27
	4	agreement	US withdrawal from the Paris Agreement	26
	5	Trump	US withdrawal from the Paris Agreement	26
R&D	1	efficiency	Energy materials such as solar cells and batteries	38
	2	design	Automatic process equipment	35
	3	system	Climate monitoring and information management system	34
	4	nano	Energy materials such as solar cells and batteries	34
	5	materials	Energy materials such as solar cells and batteries	34
Patent	1	cell	Battery structure and stacked device	73
	2	electrode	Battery structure and stacked device	67
	3	anode	Lithium-ion battery	66
	4	cathode	Lithium-ion battery	64
	5	discharge	Lithium-ion battery	63
Article	1	year	Number of fatality due to climate change	112
	2	system	Technology and policy research on renewable energy	112
	3	level	Number of fatality due to climate change	112
	4	development	Technology and policy research on renewable energy	111
	5	increase	Patterns and trends of temperature rise and precipitation	111

[1] GHGs: greenhouse gases.

News data representing social awareness showed the highest number of thematic clusters, seven, which implies that media have the widest perspectives on climate change awareness among the five different groups (Figure 4). This wide distribution is consistent with the LDA analysis result, as shown above. "US withdrawal from the Paris Agreement", "G20 Germany Summit", "eco-friendly renewable energy", "economy and policy research on future technology", "international organizations and forums", "countermeasures against disasters in agriculture and water resources ", and "particulate matter and air pollution" were the seven clusters observed in the co-occurrence network. The top five terms were observed as "USA", "GHG", "China", "agreement", and "Trump", in order. These five terms belonged to the same cluster, which is "US withdrawal from the Paris Agreement", implying that Korean media importantly dealt with the news related to US withdrawal from the Paris Agreement in 2017.

As presented in Figure 5, the national R&D analysis results showed five thematic clusters with a comparably balanced thematic distribution of climate change mitigation, adaptation, and impact reduction. Detailed themes of the clusters are as follows: "climate change impact on the ocean, ecology, and forest" related to impact reduction; "energy materials such as solar cells and batteries" and "automatic process equipment" related to mitigation technology; and "climate monitoring and information management system" and "selective breeding of adaptive and superior varieties/species" clusters related to climate change adaptation. The top 5 terms appearing in national R&D data were "efficiency", "design", "system", "nano", and "materials", in order, as summarized in Table 1. The terms "efficiency", "nano", and "materials" were associated with the "energy materials such as solar cells and batteries" cluster. In addition, the terms "design" and "system" were observed to be located in the center of the network.

The industrial patent analysis showed five clusters which were all associated with mitigation technology (Figure 6), as verified above in the topic analysis: "lithium-ion battery", "electric vehicle", "power and telecommunication management", "solar cell materials and manufacturing process", and "battery structure and stacked device". The terms "cell", "electrode", "anode", "cathode", and "discharge" were the top five thematic keywords in patent data (Table 1). These five keywords were associated with battery (or solar cell) technologies for mitigating greenhouse gas. Thus, the results reveal that industrial awareness of climate change in Korea is mostly focused on mitigation technology.

Lastly, the term co-occurrence network of the scientific articles was constructed using VOSviewer software, as illustrated in Figure 7. Since scientific articles solely provide English textual abstracts differently from other documents, VOSviewer, which is specialized for text mining in English, was utilized for the co-occurrence network analysis [22]. However, the method of constructing the network based on the term association between extracted nouns is basically same mechanism [22,52]. The term co-occurrence network built from articles also showed five clusters, but two exhibited minor proportions. There were three major clusters, "technology and policy research on renewable energy", "patterns and trends of temperature rise and precipitation", and "number of fatalities due to climate change", and two thematic clusters, "impact on natural disaster damage" and "soil carbon and plant varieties", were observed as a minor proportion. The terms "year", "system", "level", "development", and "increase" were highly important thematic keywords in the academic society. These important keywords are deeply associated with climate change adaptation themes.

4. Discussion

Knowledge synthesis based on machine learning is useful to discover underlying patterns or insights from a variety and large volume of big data, which provides evidence for data-driven decision-making [53,54]. The computer-assisted learning technique enables the improvement of the human capacity to handle unstructured text data rapidly, massively, and automatically [21,51,53]. Text analytic methods have already been utilized to track and identify sustainability indicators [34,55]. Likewise, the present method could be an

effective tool for sustainably monitoring and managing climate change awareness from diverse perspectives in support of SDG 13 and its targets. Furthermore, this methodology could also be applied to a cross-national level analysis with relevant data, even though this study focused on identifying the differences among cross-societal awareness of climate change at a particular period of time and in a specific country.

Generally, text mining techniques have been applied to text data from a single source or aggregated documents with the same context despite different sources. On the contrary, by conducting text analysis on documents collected from different sources, comprehensive knowledge can be synthesized to infer multifaceted aspects involving diverse stakeholders. Thus, we endeavored to develop this methodology further to improve availability beyond detecting topics and classifying documents. We suggested a validation process for reallocation after topic extraction in an effort to integrate mutually exclusive data from different sources.

The general unsupervised learning method is limited to document classification, which requires additional interpretation [56,57]. Thus, an additional process is required to extract the relationship between topics and specific indicators and to integrate separately analyzed data. For this purpose, awareness indicators derived from detailed targets in SDG 13.3 were utilized. The target of SDG 13.3 is "improving education, awareness-raising, and human and institutional capacity on climate change mitigation, adaptation, impact reduction, and early warning". Accordingly, four detailed targets, which are (1) mitigation, (2) adaptation, (3) impact reduction, and (4) early warning, were utilized as criteria for the derivation of awareness indicators. The reallocation process was performed after the extraction of latent topics using these indicators in the manual validation process. Based on this relationship, the integration of exclusive data and multilateral distribution of awareness was successfully ascertained. In addition, the relation of climate change awareness to SDG 13 can be elicited simultaneously. This novel approach, which utilizes a reallocation procedure using unified awareness indicators, will provide informative and valuable insights in the measurement of climate change awareness. This methodology can be utilized to obtain an opportunity for the achievement of sustainable development goals with high reproducibility and scalability. In addition, a broader or more specific point of view could be clarified if the relevant text data can be collected.

Typically, awareness identification research has been widely conducted through questionnaires, telephone surveys, or interviews [17,58]. These empirical methods have several limitations, such as time and cost intensiveness, issues in sample size and response rate, challenges in real-time collection, missing samples, and potential bias. Nevertheless, there are many advantages, such as collecting information suitable for the purpose from well-designed questionnaires as well as abundant individual-level background information of respondent groups [55,59,60]. The text analytic method based on computer-assisted learning has recently emerged to complement these limitations of empirical surveys [38,61]. Text mining techniques have also been utilized to elicit valuable insights from the data surveyed by empirical methods [30]. In particular, those are highly favorable for the direct detection of users' opinions from text data owing to the advantages of machine learning from data collection to analysis steps. Likewise, this study successfully captured the discriminative characteristics of climate change awareness among various groups. Therefore, the methodology used in this study can provide scientific evidence for identifying awareness from diverse stakeholder groups, which could complement survey-based research.

As a result of the LDA with the reallocation process, climate change awareness can be successfully inferred from five different societies in Korea: public, social, governmental, industrial, and academic awareness. Based on the collection and analysis of a large volume of text data from various sources, the identification of climate change awareness from diverse perspectives and their relationships with the SDG 13.3 targets was successfully investigated. The LDA topic model was used to explore the relationship between awareness indicators and each topic, as well as the distribution of awareness in terms of mitigation, adaptation, impact reduction, and early warning with respect to climate change. The term

co-occurrence network analysis was used as a content analysis, complementing the LDA results in detail. All groups showed relatively high climate change mitigation awareness in the levels with a more than 25% distribution rate. While early warning awareness lagged in the four other groups, the Korean public paid it more attention. The media showed the broadest awareness perspectives on climate change, as verified by LDA and network analysis simultaneously. The awareness of Korean industry was found, by both text analyses, to concentrate on mitigation technology in an effort to reduce greenhouse gas. The government and researchers mostly focused on both mitigation and adaptation awareness of climate change.

The distribution of all topics from the LDA results was illustrated as a network visualization based on the association between topics and awareness indicators, as shown in Figure 8. The four awareness indicators are represented as a hexagonal node and the topics, extracted from each text dataset, are represented as a circular shaped node. The colors of the nodes are distinguished by respective data sources and awareness indicators. Figure 8 is directly correlated with the summation of Tables A1–A5. News topics representing social awareness are located in the center of the network because of even distribution of the four awareness indicators. Patent topics (industrial awareness) are mostly focused on mitigation, and scientific articles are focused on adaptation; R&D topics are primarily distributed to both mitigation and adaptation. In addition, impact reduction and early warning awareness were observed to be at a low level in the overall groups of societies. Using this network visualization, all the relationships between each topic and each awareness indicator can be intuitively recognized with respect to all societies.
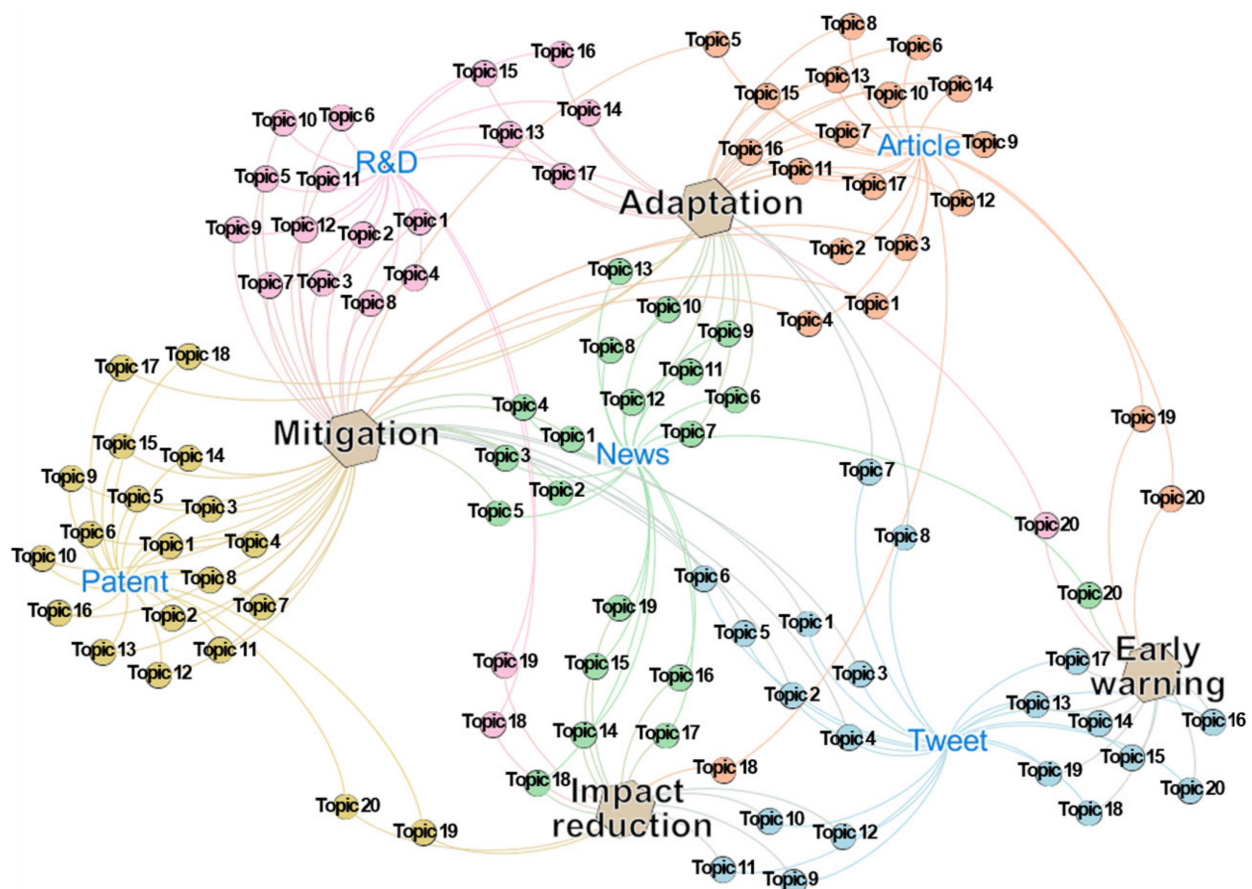

Figure 8. Network visualization of the awareness distribution based on the association between topic and indicator. The circular node represents each topic and the hexagonal node represents awareness indicators.

Sustainability and climate change issues require mutual understanding among stakeholders from various societal groups and backgrounds. In this respect, this study is well suited to provide useful knowledge to be used in policy-making and management contexts that support the implementation of SDG 13 (climate action). Data-driven knowledge management based on multiple sources presented in this study could be extended with relevant datasets to any other policy- and decision-making strategy involving diverse stakeholders. Since this approach is based on text big data analysis, it is appropriate for inferring valuable insights from various and large volume of data. Furthermore, it can be applied to real-time monitoring and analysis with tracking the increase in the velocity of data. Accordingly, this method can also be utilized for other SDG targets and as evidence for the formulation of national policy. In addition, the presented methodology could be used for other decision-making processes such as marketing strategies and sustainable management in corporate management as well as the environmental-social-governance (ESG) agenda, which requires various stakeholders' awareness. It is hoped that this methodology will be utilized widely and developed further for valuable knowledge management.

5. Conclusions

Raising-awareness of climate change is considered one of the key factors in the sustainable development goal and climate policy [62]. Climate action is most likely dependent on climate change awareness, not only from individuals but from diverse social groups [15,16]. Furthermore, there are no detailed indicators related to awareness-raising in SDG 13.3, even though the goals for climate change awareness-raising have been included. Since policy-making and strategies to address climate change issues and promoting climate action are highly dependent on various stakeholders, understanding the awareness from diverse perspectives should be considered essential. In this respect, text analyses based on multi-source data were utilized as a promising methodology to identify climate change awareness from the diverse perspectives of different groups in this study.

Five different groups were selected as decisive actors responding to climate change, and the text data produced by each of them were collected and analyzed. The differences in climate change awareness have been clearly verified through the LDA topic modeling and co-occurrence network analysis on collected text data. In order to integrate and analyze the awareness differences among these groups, the respective text data were comparatively analyzed by classifying them into four indicators, mitigation, adaptation, impact reduction, and early warning, based on the detailed targets in SDG 13.3. Through this research, the status and distribution of climate change awareness with respect to the disproportion and deficiency can be grasped from five different groups: public, social, governmental, industrial, and academic societies. Accordingly, it can be used as a criterion for decision- and policy-making tailored for each society to raise awareness.

Our research provides two main contributions to the literature. First, the multilateral view of climate change awareness among diverse cross-societal groups in a single country was successfully inferred. We proposed the methodology for comparative analysis by analyzing text data from multiple sources rather than a single source and integrating them into unified indicators. This methodology is beneficial to identify the awareness distribution among different groups based on an unsupervised learning technique. Compared to the survey method, the awareness differences of cross-societal groups could be readily identified with high-reproducibility and resource-saving owing to the computer-assisted procedures [34,55]. However, there are some limitations such as insufficient background information with regard to each stakeholder. In addition, prerequisites of collecting and handling relevant text data which can represent their groups will not be applicable to all countries at present. Nevertheless, this method is favorable for prompt responses and reducing the time and resources needed to investigate awareness. Thus, it is well suited to fill the gap in the limitations of the survey method [17,63].

Second, this methodology and these results could be utilized for the indicator of SDG 13.3 with regard to awareness-raising. However, the methodology should be further

developed as a quantitative indicator for this purpose. Nevertheless, the present method is highly appropriate as a qualitative indicator to understand the status of climate change awareness on mitigation, adaptation, impact reduction, and early warning. Thus, this methodology will encourage policymakers and researchers to develop new indicators for identifying, tracking, and monitoring SDGs [8,14], and could at the same time help collect reliable information about the status of climate change awareness. The results investigated here could be utilized for knowledge management and national policy formulation to raise awareness of climate change.

Author Contributions: Conceptualization, H.H. and C.-h.L.; data collection, S.A., E.L., S.H. and C.-h.L.; topic modeling, S.A. and C.-h.L.; network analysis, H.H. and E.L.; writing, H.H.; editing, H.H., S.H. and C.-h.L.; supervision, C.-h.L.; project administration, S.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Appendix A

Table A1. Extracted topics and reallocation results to awareness indicators from social media data.

Topics	Indicators	Relevant Keywords
1	Mitigation	temperature, fatality, increase, condition, death, increase, warming, decrease, Kilimanjaro, prediction, difficulty, dry
2	Mitigation	experience, education, summer, environment, winter, week, life, population, France, worst, time, stability
3	Mitigation	response, country, environment, Republic of Korea, global village, policy, researcher, evaluation, project, roughly, target, director
4	Mitigation	rise, during, environment, mankind, response, Republic of Korea, perception, science, technology, temperature, fine dust, level
5	Mitigation	response, energy, policy, environment, possibility, development, hold, city, greenhouse gases, regeneration, forum, solution
6	Mitigation	high, disease, malnutrition, abnormality, region, Earthquake, outbreak, collision, Asia, threat, destruction, asteroid
7	Adaptation	general assembly, free, trade, article, anxiety, effect, environment, improvement, control, ranking, dollar, eradication
8	Adaptation	fatality, decline, event, increase, temperature, heat, cold, the Ministry of Environment, parties, extent, disease, general assembly
9	Impact reduction	problem, China, think, human, science, damage, world, country, people, knowledge, story, environment
10	Impact reduction	rise, sea-level, ice, impact, melt, increase, Greenland, temperature, decrease, glacier, warming, ocean
11	Impact reduction	Antarctica, ice, sea-level, glacier, hot, region, precipitation, continent, model, country, trend, thick
12	Impact reduction	present, high, minister, temperature, average, temperature, sea-level, representative, regulatory, minister, Antarctica, finance
13	Early warning	ice, Antarctica, average, sea-level, ocean, height, rise, North Pole, float, temperature, carbon dioxide, abnormality
14	Early warning	news, drought, storm, report, people, evidence, flood, crop failure, population, benefit, response, women
15	Early warning	forecast, collision, abnormality, polar bear, Africa, death, disaster, possibility, occurrence, long term, asteroid, photograph

Table A1. *Cont*.

Topics	Indicators	Relevant Keywords
16	Early warning	Trump, USA, withdrawal, Agreement, president, Paris Agreement, government, declaration, serious, possibility, drought, season
17	Early warning	forecast, sea-level, rise, during, warming, movie, truth, sink, discomfort, scene, reality, media
18	Early warning	world, future, warming, occurrence, glacier, enormous, rise, high, coast, sea-level, tsunami, melting
19	Early warning	people, death, glacier, fatality, abnormality, weather, growth, reduction, case, severe cold, winter, summer
20	Early warning	fatality, decrease, increase, cold, shortage, development, people, herd, water sampling, UK, heat, increase

Table A2. Extracted topics and reallocation results to awareness indicators from news data.

Topics	Indicators	Relevant Keywords
1	Mitigation	summit, USA, Germany, China, trade, meeting, Merkel, Europe, international, Trump, economy, free
2	Mitigation	power plant, greenhouse gases, energy, power generation, carbon, photovoltaic, emission, nuclear power, fuel fossil, eco-friendly, coal, carbon dioxide
3	Mitigation	sustainable, management, society, finance, corporation, contribution, stakeholder, development, eco-friendly, committee, group, institution
4	Mitigation	industry, revolution, data, science, economy, corporation, future, resources, job, artificial intelligence, competitiveness, innovation
5	Mitigation	USA, convention, Paris, Trump, Agreement, Obama, greenhouse gases, China, administration, the White House, France, emission
6	Adaptation	Gwangju, Jeju, Jeonnam, international, environment, center, university, first-come-first-served, competition, culture, forum, expert
7	Adaptation	Chungnam, rural area, public corporation, management, area, reservoir, drought, water resources, headquarter, countermeasure, agricultural water, project
8	Adaptation	ecology, national, park, marine, organism, resources, fisheries, diversity, arboretum, region, extinction, Jeonnam
9	Adaptation	committee, expert, government, policy, the Ministry of Environment, project, president of committee, Moon Jae-in, assembly, citizen, candidate, organization
10	Adaptation	safety, region, project, facility, Ulsan, management, planning, local governments, disaster, Gyeongbuk, city, accident
11	Adaptation	international, cooperation, Incheon, Agreement, Asia, economy, center, project, response, green, task, organization
12	Adaptation	Seoul, secretary general, society, international, USA, UN, market, world, Ban Ki-moon, New York, Seoul, City
13	Adaptation	forest, the Forest Service, resources, tree, carbon, water source, wood, economy, construction, management, forest fires, region
14	Impact reduction	environment, education, experience, student, council, life, school, energy, possibility, Chungbuk, activity, practice
15	Impact reduction	weather, the Meteorological Administration, country, warming, Earth, temperature, average, precipitation, province, center, science, Seoul
16	Impact reduction	management, vulnerability, region, damage, countermeasures, summertime, resident, health, prevention, occurrence, disaster, welfare
17	Impact reduction	environment, pollution, atmosphere, fine dust, China, matters, Seoul, policy, the Minister of Environment, environment service, vehicle, emission
18	Impact reduction	agriculture, cultivation, income, farmhouse, agricultural product, region, rural area, food, production, insurance, education, subtropical regions
19	Impact reduction	city, citizen, greenhouse gases, eco-friendly, energy, housing, space, vehicle, stakeholder, green, park, facility
20	Early warning	USA, scientist, Earth, warming, science, U.K., world, people, Antarctica, global village, laboratory, Europe

Table A3. Extracted topics and reallocation results to awareness indicators from R&D data.

Topics	Indicators	Relevant Keywords
1	Mitigation	process, production, improvement, automatic, product, improvement, operation, processing, quality, defect, time, occurrence
2	Mitigation	solar, battery, device, material, structure, process, nano, thin film, characteristics, organic, electrode, energy
3	Mitigation	systems, energy, power, power generation, photovoltaic, electricity, battery, control, module, charging, grid, wind power
4	Mitigation	product, market, improvement, domestic, function, demand, lighting, effect, design, industry, oversea, export
5	Mitigation	above, below, temperature, measurement, time, criteria, level, maintenance, drying, pressure, control, air
6	Mitigation	electrode, cell, material, ion, nano, lithium, characteristics, chemistry, electricity, structure, metal, catalyst
7	Mitigation	material, process, coating, composite, characteristics, surface, product, eco-friendly, fiber, enhancement, strength, excellence
8	Mitigation	design, test, production, evaluation, prototype, system, module, structure, parts, proof, reliability, improvement
9	Mitigation	production, hydrogen, catalyst, reaction, fuel, gas, bio, conversion, energy, separation, system, carbon dioxide
10	Mitigation	system, demonstration, operation, energy, facility, integration, design, evaluation, site, method, utilization, model
11	Mitigation	institution, participation, evaluation, design, university, production, process, characteristics, derivation, condition, operation, facility
12	Mitigation	sensor, data, control, vehicle, driving, communication, autonomy, environment, real-time, algorithm, platform, smart
13	Adaptation	varieties, cultivation, region, crop, genetic source, pedigree, incubation, selection, characteristics, production, superior, deployment
14	Adaptation	microorganisms, function, production, plant, food, organism, activation, genomes, algae, incubation, verification, animal
15	Adaptation	resources, support, international, utilization, industry, national, management, standard, system, training, education, manpower
16	Adaptation	diagnosis, virus, disease, vaccine, evaluation, infection, system, livestock, antibody, detection, antigen, outbreak
17	Adaptation	model, interpretation, characteristics, experiment, technique, forecast, performance, offshore, behavior, phenomenon, evaluation, occurrence
18	Impact reduction	evaluation, safety, criteria, accident, nuclear power plant, interpretation, verification, regulation, management, domestic, nuclear fuel, Earthquake
19	Impact reduction	matters, contamination, treatment, environment, removal, fine dust, water quality, emission, atmosphere, filter, chemical, reduction
20	Early warning	change, climate, data, environment, impact, forecast, region, evaluation, ecology, occurrence, marine, weather

Table A4. Extracted topics and reallocation results to awareness indicators from patent data.

Topics	Indicators	Relevant Keywords
1	Mitigation	converter, drone, transformer, flight, resonance, invalidity, harmonic wave, unmanned, video, command, condenser, rectifier
2	Mitigation	fuel cell, duct, junction, gasket, HVAC, catalyst, skin, burner, jet, water, fine dust, residence
3	Mitigation	capacitor, super, luminous, touch, ultra, phenol, porous, mask, host, graphene, sulfide, ink
4	Mitigation	lighting, monitor, humidifier, street lamp, power saving, ultrasound, LED, standby power, humidification, illumination, vibrator, noise
5	Mitigation	catalyst, exhaust, material, pyrolysis, adsorbent, seawater, saltwater, steam, waste heat, olefin, evaporator, slip
6	Mitigation	sun, wafer, dye, substrate, sensitized, rear, copper foil, photovoltaic, boat, pigmentation, glass, satellite

Table A4. *Cont.*

Topics	Indicators	Relevant Keywords
7	Mitigation	forming, glass, mold, pole plate, window, thermoelectricity, curved surface, wave power, scrap, lens, positive pole, negative pole
8	Mitigation	terminal, packet, node, seedling, base station, downlink, wake-up, object, actuator, direction, vanadium, resources
9	Mitigation	photovoltaic, panel, tracking, water-floating, buoyancy, roof, building, string, solar, insulation, mooring, support
10	Mitigation	turbine, ship, blade, rotor, propeller, wing, hull, liquefaction, bearing, boat, navigation, aircraft
11	Mitigation	graphene, algae, bacteria, ferrite, sulfide, infection, antibiotic, feed, shielding, body parts, locking, lipids
12	Mitigation	valve, cylinder, section, exhaust, bypass, turbocharger, regeneration, rpm, metering, brake, head, rail
13	Mitigation	hybrid, nanometer, chemical formula, trade, container, cartridge, flexible, terminal, calculation, object, hookup, charger
14	Mitigation	battery, charger, electric vehicle, socket, master, water cooled, agricultural, parking, booth, loading, relay, car body
15	Mitigation	substrate, station, slag, cutting, color, relay, scribing, super-electric, panel, protective film, cellulose, smelting
16	Mitigation	fermentation, exchanger, fertilizer, waste, food, geothermal, stirring, compost, underground, groundwater, organic, compartment
17	Adaptation	glass, rubber, separator, refrigerant, cutting, filament, glass plate, car body, evaporator, butadiene, door
18	Adaptation	cultivation, plant, crop, hydroponic, greenhouse, nutrient solution, vegetable, house, farming, water quality, culture medium, cultivator
19	Impact reduction	grid, bus, reactor, tower, nuclear fuel, distribution, core, coolant, uranium, fuel rod, washing machine, host
20	Impact reduction	cell, sequence, virus, antibody, immunity, protein, antigen, vector, tumor, infection, nucleic acid, vaccine

Table A5. Extracted topics and reallocation results to awareness indicators from article data.

Topics	Indicators	Relevant Keywords
1	Mitigation	CCS [1], housing, underground, German, smart, thing, hydropower, insurance, insulation, transmission, disclosure, tunnel
2	Mitigation	shale, governance, aid, cell, diffusion, strain, bacteria, cyanobacteria, blooms, outbreak, cyanobacterial, airborne
3	Mitigation	vehicle, electric, ecofriendly, liquid, catalyst, injection, synthesis, vehicles, delay, pyrolysis, decomposition, polymer
4	Mitigation	student, education, school, attitude, saving, curriculum, patient, literacy, unprecedented, teacher, Tokyo, attitudes
5	Mitigation	tax, stock, New Zealand, credit, patent, litter, agenda, bill, GATT [2], ODA [3], afforestation, discourse
6	Adaptation	holocene, Jeju, pollen, deposit, paleoclimate, fluvial, subsurface, volcanic, alluvial, carbonate, diatom, shelf
7	Adaptation	phytoplankton, gene, protein, integration, Antarctic, japonicus, mitochondrial, amino, copepods, toxicity, acid, biochemical
8	Adaptation	drought, rice, irrigation, corn, paddy, yields, fertilizer, GPP [4], japonica, canopy, drought, compost
9	Adaptation	flood, nuclear, resilience, earthquake, drainage, pump, inundation, refrigerant, debris, minute, storm, acceptance
10	Adaptation	pine, storm, surge, radial, topographic, palm, bioethanol, mortality, mercury, oak, biofuel, cherry
11	Adaptation	groundwater, image, algal, aquifer, salinity, membrane, aquifers, algae, estuary, freshwater, desalination, separation
12	Adaptation	microbial, communities, abundance, salt, tidal, bay, egg, abiotic, taxonomic, rivers, beetle, biotic

Table A5. *Cont.*

Topics	Indicators	Relevant Keywords
13	Adaptation	cloud, dust, GHGs [5], fine, aerosol, convective, liability, cloud, bottom up, OECD [6], saving, scarcity
14	Adaptation	vulnerability, Busan, Daegu, outdoor, grassland, family, subalpine, maritime, moral, endemic, Incheon, VESTAP [7]
15	Adaptation	tourism, consumer, satisfaction, perception, tourist, media, logistics, news, behavioral, intention, occupant, comfort
16	Adaptation	disease, particulate matter, mortality, death, ozone, ambient, particulate, bulb, episode, cardiovascular, cultivar, FAO [8]
17	Adaptation	concrete, snow, snowfall, fishing, block, corrosion, ventilation, removal, illegal, welfare, carbonation, chamber
18	Impact reduction	dam, landslide, hazard, beach, nonlinear, shoreline, sludge, hydrology, hydroclimatic, sewage, warning, incineration
19	Early warning	ice, Arctic, polar, Antarctic, Joseon, sea-ice, Thailand, yellow, infiltration, permafrost, retreat, nonstationary
20	Early warning	anomalous, El Niño, EAWM [9], westerly, anticyclone, TCS [10], attribution, eddy, catch, air-sea, gyre, hydroclimate

[1] CCS (Carbon dioxide Capture and Storage); [2] GATT (the General Agreement on Tariffs and Trade); [3] ODA (Official Development Assistance); [4] GPP (Gross Primary Production); [5] GHGs (Greenhouse Gases); [6] OECD (Organization for Economic Cooperation and Development); [7] VESTAP (Vulnerability Assessment Tool to build Climate Change Adaptation Plan); [8] FAO (Food and Agriculture Organization); [9] EAWM (the East Asian winter monsoon); [10] TCS (Tropical Cyclones).

References

1. Cazenave, A.; Dieng, H.-B.; Meyssignac, B.; von Schuckmann, K.; Decharme, B.; Berthier, E. The rate of sea-level rise. *Nat. Clim. Chang.* **2014**, *4*, 358–361. [CrossRef]
2. Lamarche-Gagnon, G.; Wadham, J.L.; Sherwood Lollar, B.; Arndt, S.; Fietzek, P.; Beaton, A.D.; Tedstone, A.J.; Telling, J.; Bagshaw, E.A.; Hawkings, J.R.; et al. Greenland melt drives continuous export of methane from the ice-sheet bed. *Nature* **2019**, *565*, 73–77. [CrossRef]
3. Nicholls, S. Impacts of environmental disturbances on housing prices: A review of the hedonic pricing literature. *J. Environ. Manag.* **2019**, *246*, 1–10. [CrossRef] [PubMed]
4. Unfccc Adoption of The Paris Agreement. Rep. No. FCCC/CP/2015/L.9/Rev.1. 2015. Available online: https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf (accessed on 22 March 2021).
5. Rogelj, J.; den Elzen, M.; Höhne, N.; Fransen, T.; Fekete, H.; Winkler, H.; Schaeffer, R.; Sha, F.; Riahi, K.; Meinshausen, M. Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature* **2016**, *534*, 631–639. [CrossRef]
6. Sebestyén, V.; Domokos, E.; Abonyi, J. Focal points for sustainable development strategies—Text mining-based comparative analysis of voluntary national reviews. *J. Environ. Manag.* **2020**, *263*, 110414. [CrossRef]
7. UN General Assembly. *Transforming Our World: The 2030 Agenda for Sustainable Development*; A/RES/70/1; United Nations: New York, NY, USA, 2015. Available online: https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E (accessed on 22 March 2021).
8. Yonehara, A.; Saito, O.; Hayashi, K.; Nagao, M.; Yanagisawa, R.; Matsuyama, K. The role of evaluation in achieving the SDGs. *Sustain. Sci.* **2017**, *12*, 969–973. [CrossRef]
9. Rosati, F.; Faria, L.G.D. Addressing the SDGs in sustainability reports: The relationship with institutional factors. *J. Clean. Prod.* **2019**, *215*, 1312–1326. [CrossRef]
10. Lee, H.; Son, J.; Joo, D.; Ha, J.; Yun, S.; Lim, C.-H.; Lee, W.-K. Sustainable Water Security Based on the SDG Framework: A Case Study of the 2019 Metro Manila Water Crisis. *Sustainability* **2020**, *12*, 6860. [CrossRef]
11. Doni, F.; Gasperini, A.; Soares, J.T. What is the SDG 13? In *SDG13—Climate Action: Combating Climate Change and Its Impacts*; Emerald Publishing Limited: Bingley, UK, 2020; pp. 21–30.
12. Gomez-Echeverri, L. Climate and development: Enhancing impact through stronger linkages in the implementation of the Paris Agreement and the Sustainable Development Goals (SDGs). *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2018**, *376*, 20160444. [CrossRef]
13. Coscieme, L.; Mortensen, L.F.; Anderson, S.; Ward, J.; Donohue, I.; Sutton, P.C. Going beyond Gross Domestic Product as an indicator to bring coherence to the Sustainable Development Goals. *J. Clean. Prod.* **2020**, *248*, 119232. [CrossRef]
14. Fraisl, D.; Campbell, J.; See, L.; Wehn, U.; Wardlaw, J.; Gold, M.; Moorthy, I.; Arias, R.; Piera, J.; Oliver, J.L.; et al. Mapping citizen science contributions to the UN sustainable development goals. *Sustain. Sci.* **2020**, *15*, 1735–1751. [CrossRef]
15. Clayton, S.; Devine-Wright, P.; Stern, P.C.; Whitmarsh, L.; Carrico, A.; Steg, L.; Swim, J.; Bonnes, M. Psychological research and global climate change. *Nat. Clim. Chang.* **2015**, *5*, 640–646. [CrossRef]

16. Doherty, K.L.; Webler, T.N. Social norms and efficacy beliefs drive the Alarmed segment's public-sphere climate actions. *Nat. Clim. Chang.* **2016**, *6*, 879–884. [CrossRef]

17. Lee, T.M.; Markowitz, E.M.; Howe, P.D.; Ko, C.-Y.; Leiserowitz, A.A. Predictors of public climate change awareness and risk perception around the world. *Nat. Clim. Chang.* **2015**, *5*, 1014–1020. [CrossRef]

18. Knight, K.W. Public awareness and perception of climate change: A quantitative cross-national study. *Environ. Sociol.* **2016**, *2*, 101–113. [CrossRef]

19. Haara, A.; Pykäläinen, J.; Tolvanen, A.; Kurttila, M. Use of interactive data visualization in multi-objective forest planning. *J. Environ. Manag.* **2018**, *210*, 71–86. [CrossRef]

20. Aykroyd, R.G.; Leiva, V.; Ruggeri, F. Recent developments of control charts, identification of big data sources and future trends of current research. *Technol. Forecast. Soc. Chang.* **2019**, *144*, 221–232. [CrossRef]

21. Khan, Z.; Vorley, T. Big data text analytics: An enabler of knowledge management. *J. Knowl. Manag.* **2017**, *21*, 18–34. [CrossRef]

22. Li, J.; Antonenko, P.D.; Wang, J. Trends and issues in multimedia learning research in 1996–2016: A bibliometric analysis. *Educ. Res. Rev.* **2019**, *28*, 100282. [CrossRef]

23. Hemmatian, F.; Sohrabi, M.K. A survey on classification techniques for opinion mining and sentiment analysis. *Artif. Intell. Rev.* **2019**, *52*, 1495–1545. [CrossRef]

24. Li, D.; Zhang, Y.; Li, C. Mining Public Opinion on Transportation Systems Based on Social Media Data. *Sustainability* **2019**, *11*, 4016. [CrossRef]

25. Wang, B.; Liu, S.; Ding, K.; Liu, Z.; Xu, J. Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology. *Scientometrics* **2014**, *101*, 685–704. [CrossRef]

26. Ho, L.; Alonso, A.; Eurie Forio, M.A.; Vanclooster, M.; Goethals, P.L.M. Water research in support of the Sustainable Development Goal 6: A case study in Belgium. *J. Clean. Prod.* **2020**, *277*, 124082. [CrossRef]

27. Madruga de Brito, M.; Kuhlicke, C.; Marx, A. Near-real-time drought impact assessment: A text mining approach on the 2018/19 drought in Germany. *Environ. Res. Lett.* **2020**, *15*, 1040a9. [CrossRef]

28. Huo, F.; Xu, L.; Li, Y.; Famiglietti, J.S.; Li, Z.; Kajikawa, Y.; Chen, F. Using big data analytics to synthesize research domains and identify emerging fields in urban climatology. *Wiley Interdiscip. Rev. Clim. Chang.* **2021**, *12*, e688. [CrossRef]

29. Zhang, L.; Wang, S.; Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1253. [CrossRef]

30. Tvinnereim, E.; Fløttum, K.; Gjerstad, Ø.; Johannesson, M.P.; Nordø, Å.D. Citizens' preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions. *Glob. Environ. Chang.* **2017**, *46*, 34–41. [CrossRef]

31. Chae, S.; Gim, J. A Study on Trend Analysis of Applicants Based on Patent Classification Systems. *Information* **2019**, *10*, 364. [CrossRef]

32. Lee, H.J.; Oh, H. A Study on the Deduction and Diffusion of Promising Artificial Intelligence Technology for Sustainable Industrial Development. *Sustainability* **2020**, *12*, 5609. [CrossRef]

33. Uspto Cooperative Patent Classification. Available online: https://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y.html (accessed on 22 March 2021).

34. Park, K.; Kremer, G.E.O. Text mining-based categorization and user perspective analysis of environmental sustainability indicators for manufacturing and service systems. *Ecol. Indic.* **2017**, *72*, 803–820. [CrossRef]

35. De Clercq, D.; Diop, N.-F.; Jain, D.; Tan, B.; Wen, Z. Multi-label classification and interactive NLP-based visualization of electric vehicle patent data. *World Pat. Inf.* **2019**, *58*, 101903. [CrossRef]

36. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation David. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022. [CrossRef]

37. Boussalis, C.; Coan, T.G. Text-mining the signals of climate change doubt. *Glob. Environ. Chang.* **2016**, *36*, 89–100. [CrossRef]

38. Nam, S.; Ha, C.; Lee, H. Redesigning In-Flight Service with Service Blueprint Based on Text Analysis. *Sustainability* **2018**, *10*, 4492. [CrossRef]

39. Zhang, W.; Yoshida, T.; Tang, X. A comparative study of TF*IDF, LSI and multi-words for text classification. *Expert Syst. Appl.* **2011**, *38*, 2758–2765. [CrossRef]

40. Joung, J.; Jung, K.; Ko, S.; Kim, K. Customer Complaints Analysis Using Text Mining and Outcome-Driven Innovation Method for Market-Oriented Product Development. *Sustainability* **2019**, *11*, 40. [CrossRef]

41. Do Park, S.; Lee, J.Y.; Wang, B. The main factors affecting cultural exchange between Korea and China: A semantic network analysis based on the cultural governance perspective. *Int. J. Intercult. Relat.* **2019**, *71*, 72–83. [CrossRef]

42. Abrahams, B.; Sitas, N.; Esler, K.J. Exploring the dynamics of research collaborations by mapping social networks in invasion science. *J. Environ. Manag.* **2019**, *229*, 27–37. [CrossRef]

43. Dahal, B.; Kumar, S.A.P.; Li, Z. Topic modeling and sentiment analysis of global climate change tweets. *Soc. Netw. Anal. Min.* **2019**, *9*, 1–20. [CrossRef]

44. Pickering, C.M.; Norman, P. Assessing discourses about controversial environmental management issues on social media: Tweeting about wild horses in a national park. *J. Environ. Manag.* **2020**, *275*, 111244. [CrossRef]

45. Jacomy, M.; Venturini, T.; Heymann, S.; Bastian, M. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLoS ONE* **2014**, *9*, e98679. [CrossRef] [PubMed]

46. Lam, J.C.K.; Cheung, L.Y.L.; Wang, S.; Li, V.O.K. Stakeholder concerns of air pollution in Hong Kong and policy implications: A big-data computational text analysis approach. *Environ. Sci. Policy* **2019**, *101*, 374–382. [CrossRef]

47. Callaghan, M.W.; Minx, J.C.; Forster, P.M. A topography of climate change research. *Nat. Clim. Chang.* **2020**, *10*, 118–123. [CrossRef]

48. Zandi, S.; Nemati, B.; Jahanianfard, D.; Davarazar, M.; Sheikhnejad, Y.; Mostafaie, A.; Kamali, M.; Aminabhavi, T.M. Industrial biowastes treatment using membrane bioreactors (MBRs)—A scientometric study. *J. Environ. Manag.* **2019**, *247*, 462–473. [CrossRef] [PubMed]

49. Indulska, M.; Hovorka, D.S.; Recker, J. Quantitative approaches to content analysis: Identifying conceptual drift across publication outlets. *Eur. J. Inf. Syst.* **2012**, *21*, 49–69. [CrossRef]

50. Stockwell, P.; Colomb, R.M.; Smith, A.E.; Wiles, J. Use of an automatic content analysis tool: A technique for seeing both local and global scope. *Int. J. Hum. Comput. Stud.* **2009**, *67*, 424–436. [CrossRef]

51. Nunez-Mir, G.C.; Iannone, B.V.; Pijanowski, B.C.; Kong, N.; Fei, S. Automated content analysis: Addressing the big literature challenge in ecology and evolution. *Methods Ecol. Evol.* **2016**, *7*, 1262–1272. [CrossRef]

52. Van Eck, N.J.; Waltman, L. Text Mining and Visualization using VOSviewer. *arXiv* **2011**, arXiv:1109.2058.

53. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144. [CrossRef]

54. Nilsson, M.; Chisholm, E.; Griggs, D.; Howden-Chapman, P.; McCollum, D.; Messerli, P.; Neumann, B.; Stevance, A.-S.; Visbeck, M.; Stafford-Smith, M. Mapping interactions between the sustainable development goals: Lessons learned and ways forward. *Sustain. Sci.* **2018**, *13*, 1489–1503. [CrossRef]

55. Rivera, S.J.; Minsker, B.S.; Work, D.B.; Roth, D. A text mining framework for advancing sustainability indicators. *Environ. Model. Softw.* **2014**, *62*, 128–138. [CrossRef]

56. Suominen, A.; Toivanen, H. Map of science with topic modeling: Comparison of unsupervised learning and human-assigned subject classification. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 2464–2476. [CrossRef]

57. Hagen, L. Content analysis of e-petitions with topic modeling: How to train and evaluate LDA models? *Inf. Process. Manag.* **2018**, *54*, 1292–1307. [CrossRef]

58. Wang, S.W.; Lee, W.K.; Brooks, J.; Dorji, C. Awareness of the Convention on Biological Diversity and provisions regarding access and benefit sharing among multiple stakeholder groups in Bhutan. *Oryx* **2020**, *54*, 735–742. [CrossRef]

59. Dalla Costa, E.; Tranquillo, V.; Dai, F.; Minero, M.; Battini, M.; Mattiello, S.; Barbieri, S.; Ferrante, V.; Ferrari, L.; Zanella, A.; et al. Text Mining Analysis to Evaluate Stakeholders' Perception Regarding Welfare of Equines, Small Ruminants, and Turkeys. *Animals* **2019**, *9*, 225. [CrossRef] [PubMed]

60. Lesnikowski, A.; Belfer, E.; Rodman, E.; Smith, J.; Biesbroek, R.; Wilkerson, J.D.; Ford, J.D.; Berrang-Ford, L. Frontiers in data analytics for adaptation research: Topic modeling. *Wiley Interdiscip. Rev. Clim. Chang.* **2019**, *10*, e576. [CrossRef]

61. Puerta, P.; Laguna, L.; Vidal, L.; Ares, G.; Fiszman, S.; Tárrega, A. Co-occurrence networks of Twitter content after manual or automatic processing. A case-study on "gluten-free". *Food Qual. Prefer.* **2020**, *86*, 103993. [CrossRef]

62. Luís, S.; Vauclair, C.-M.; Lima, M.L. Raising awareness of climate change causes? Cross-national evidence for the normalization of societal risk perception of climate change. *Environ. Sci. Policy* **2018**, *80*, 74–81. [CrossRef]

63. Chapman, A.; Shigetomi, Y. Developing national frameworks for inclusive sustainable development incorporating lifestyle factor importance. *J. Clean. Prod.* **2018**, *200*, 39–47. [CrossRef]