

## Article

# Identifying Emerging Issues in the Seafood Industry Based on a Text Mining Approach

Kiuk Han <sup>1</sup> , Jaesun Yeom <sup>2</sup>  and Keunsuk Chung <sup>2,\*</sup> 

<sup>1</sup> Fisheries Policy Implementation, Korea Maritime Institute, Haeyang-ro 301-gil 26, Busan 49111, Republic of Korea; hankiuk@kmi.re.kr

<sup>2</sup> School of Business Admin, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulsan 44919, Republic of Korea; jaesunyeom22@unist.ac.kr

\* Correspondence: chungk@unist.ac.kr

**Abstract:** Identification of emerging issues has garnered growing interest as a way to establish proactive policy formulation. However, in fisheries research, analyzing such issues has largely depended on the literature or researchers' judgment. We use keyword analysis, targeting news application programming interfaces (News APIs) (72,981 news sources and blogs), to investigate issues in the global seafood industry from January 2019 to March 2022. Among a variety of topics identified by year and country, in general, seafood market function, health, and tariffs were the main issues in 2019, while COVID-19-related issues were primarily mentioned between 2020 and 2021. After 2022, the role of the market regained attention, and various new issues rose to the surface. To identify emerging issues, we jointly employ dynamic time warping (DTW) and growth models, which derive several keywords, including coercion, cuisines, food safety, ketones, plastic ingestions, seafood alcohol, urbanization, wastewater treatment, and the World Trade Organization (WTO). High interest in food safety, environmental change, trade conflict, and seafood value improvement reveal the need for proper policy responses.

**Keywords:** global issues; emerging issues; seafood; horizon scanning; text mining



**Citation:** Han, K.; Yeom, J.; Chung, K. Identifying Emerging Issues in the Seafood Industry Based on a Text Mining Approach. *Appl. Sci.* **2024**, *14*, 1820. <https://doi.org/10.3390/app14051820>

Academic Editors: Giovanni Luca Russo and Antonio Luca Langellotti

Received: 15 December 2023

Revised: 20 February 2024

Accepted: 20 February 2024

Published: 22 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Rapid changes and rising uncertainty in the global landscape have made monitoring of emerging issues increasingly popular. To reflect this trend, growing bodies of researchers, governments, and industries around the world are actively conducting horizon-scanning studies in the dimensions of business and public policy (see [1–5]). Moreover, as the pace of such changes accelerates, formerly minor issues can be transformed into major trends in a condensed timeframe. Thus, there is a growing need to identify such minor issues and monitor changing trends systematically.

In the seafood industry, trade accounts for a higher proportion (37%) compared to other primary industries (9.8% for meat). Moreover, global issues have a significant impact on the supply chains of individual countries [6]. Thus, in order to devise proactive policy measures and strategies for the seafood industry in a timely manner, mere monitoring of current issues may prove inadequate. It is necessary to identify emerging issues that may garner significant attention in the future.

Although horizon scanning is widely conducted on various subjects in many fields, it is rare to find literature in fisheries that delivers comprehensive and systematic analysis, except for a few studies focusing on specific topics and regions ([7] on marine microplastic debris; and [8] on Europe). However, those studies either show limited capability in continuous monitoring of general trends or identify a small number of candidates, failing to acknowledge other possible issues (for instance, [9] only proposes COVID-19 as an emerging issue may exist). Therefore, we can see that there is a requirement for an approach that can continuously monitor various fields.

To overcome these methodological difficulties, we aim to derive issues with high external validity based on various data sources by year. In [10], the author provides some clues for determining the *emergingness* of issues. According to him, the slopes of issue curves take off slowly, then rise sharply, and finally taper off. This S-curve concept helps us identify issues with similar growth patterns by fitting the curve. The S-curve is useful in helping us understand that disruption and reordering are cyclical [11]. It is also being applied to trend projections and technology forecasting (see [12–14]).

Various approaches are used to analyze trending issues. For example, efforts are being made to perform keyword analysis on journals, reports, and internet articles (e.g., [15–17]). In predicting keyword trends, ordinary least squares regression (OLS) [18] or machine learning [19,20] are used. The authors of [21] propose using the Web as a source of information, i.e., Google search, while checking if the keywords of interest follow an S-curve. The process of deriving new problems using computer algorithms has expected advantages such as enhanced transparency and independence from individual bias or specific intentions [22,23].

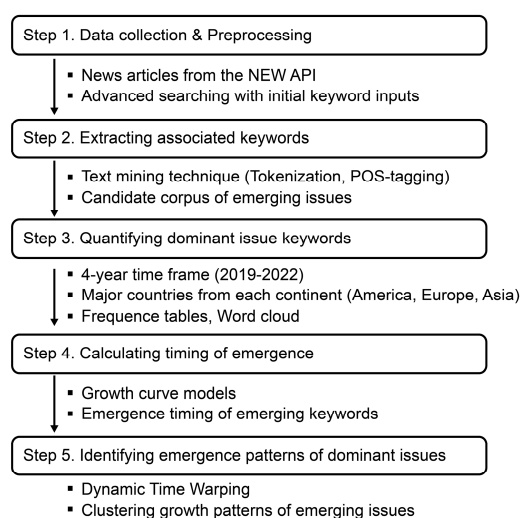
Reflecting these research trends, this study attempts to reveal major issues in the seafood sector and expands the intellectual realm of issue analysis. To this end, articles related to the seafood industry were extracted via news application programming interfaces (News APIs), encompassing more than 80,000 news and blog sources, and the issues that received attention were investigated by year. As a follow-up, considering the S-shaped growth pattern of the issues, we classify and introduce them as emerging issues that will receive attention in the future.

## 2. Materials and Methods

### 2.1. Research Framework

To identify emerging issues in seafood consumption, we trace frequently mentioned keywords from global news articles. The premise of this study is that global news articles contain emerging issues of seafood consumption correlated with social, economic, and political events. Furthermore, news articles are considered structured documents that can be used as refined sources of opinion mining and topic modeling. Owing to such characteristics of news articles, they contain keywords that are more likely to become crucial sources of emerging trends.

Figure 1 presents the process of our analysis in five steps: (1) data collection and preprocessing; (2) extracting associated keywords; (3) quantifying dominant issue keywords; (4) identifying emerging patterns of dominant issues; and (5) deciding on candidates for emerging issues. We provide detailed explanations of each sub-step in the following subsections.



**Figure 1.** Research framework.

## 2.2. Text Mining (Keyword Analysis)

### 2.2.1. Data Collection and Preprocessing

As noted earlier, the News APIs used contain current and historical news articles from more than 80,000 published sources worldwide. It also offers advanced search parameters that confine the scope of news content. First, we construct initial keyword inputs, which contain globally consumed major seafood items as well as the names of major seafood-exporting/importing countries (see Table 1). These keyword inputs filter the titles and contents of news articles relevant to the seafood domain, yielding a collection of 72,981 news articles from January 2019 to March 2022. After the initial search, we further filter out less related data by eliminating articles containing unrelated terms in their titles.

**Table 1.** Data collected from the News APIs by year.

Country		2019	2020	2021	2022
East Asia	China	2795	11,255	3986	697
	Japan	1986	4108	1605	469
	South Korea	600	2815	605	199
Southeast Asia	Indonesia	477	737	396	88
	Thailand	798	3132	687	100
	Vietnam	800	1508	651	153
Europe	France	1600	2851	1393	363
	Italy	1399	2440	1227	298
	Spain	1176	1077	854	200
Americas	Canada	1699	2691	1495	400
	USA	2794	4940	2648	789
Total		16,124	37,554	15,547	3756

### 2.2.2. Extracting Associated Keywords

From the constructed news article source, we extract associated keywords, i.e., a candidate corpus of emerging issues. The extraction process is facilitated by applying text mining techniques, namely tokenization, part-of-speech tagging (POS tagging), and N-grams, which collectively identify keyword candidates from the news content. These techniques are implemented using the Python (version 3.10) NLTK library. First, the application of the text mining technique starts with sentence tokenization, the goal of which is to slice chunk-level text into minimum units. Thus, each paragraph in a news article is sliced up into sentences, and each sentence is converted into a set of words. Second, POS tagging aims to identify specific grammatical word forms (nouns, verbs, adjectives, adverbs, or other forms) within the sentence. As people usually understand a news article's content by looking at noun terminologies, we identified noun terms using the POS tagging technique. Third, the noun terminologies could be composite noun terms made up of two or more words (e.g., food market, protein ingredients, etc.). To extract these composite noun terms (bigrams and trigrams), we utilize the N-gram technique. Finally, we establish a set of noun keywords comprising candidate corpora of emerging issues.

### 2.2.3. Quantifying Dominant Issue Keywords

In this step, we construct keyword frequency tables by year and country based on associated keywords from the previous step. Since we aim to identify recent emerging issues, we quantify yearly keywords in the past four years, assuming a four-year timeframe is sufficiently long to capture emerging issues and their changes. Dominant topics are summarized in frequency tables (the frequency tables are not present in the article but will be available upon request to authors) that demonstrate the transition of focus over time.

We visualize the information in the frequency tables with word clouds, illustrating the dominant keywords that have emerged recently.

### 2.3. Emerging Issues Analysis

#### 2.3.1. Identifying Emergence Patterns of Dominant Issues

According to [10,24], a particular issue transitions from an emerging one to a problem/opportunity, eventually fading away, following an S-curve that represents the issue's emergence, growth, maturity, and decline. However, capturing and addressing individual issues one by one presents practical challenges to researchers and policymakers. We classify common growth patterns among major issues by adopting the dynamic time warping (DTW) technique, which is known for its strength in measuring the similarity between two time series sequences with possible differences either in length or timing. Unlike Euclidean distance, which compares signals one-to-one along an identical time axis, DTW aligns signal sequences based on their order of occurrence and compares the amplitudes of the matched signals, thereby providing a better similarity score.

Curve-fitted keywords showing similar growth patterns are clustered by the DTW algorithm to serve two purposes. One is to verify the presence of the S-curve growth patterns in the frequency of keywords, and the other is to classify many keywords at once. Two time series' consistency can be determined using DTW, even if there is warping between them [25]. It also minimizes the distance between the two time series and calculates the cumulative distance to the minimum [26].

When comparing two time series,  $X = (x_1, x_2, \dots, x_m)$  and  $Y = (y_1, y_2, \dots, y_n)$ , with lengths of  $m$  and  $n$ , respectively, DTW creates an  $m \times n$  matrix and then calculates the distance between  $x_i$  and  $y_j$  by  $(x_i - y_j)^2$ . See [25] for detailed explanations about the process.

Between k-medoids and k-means algorithms, we opt for the former, since the latter is vulnerable to the existence of outliers [27,28]. In the context of our study, outliers could be predominant keywords related to coronavirus issues. If the algorithm of our choice is vulnerable to the presence of such dominant keywords, potentially important issues that are less relevant to the pandemic could be substantially overlooked. Since our study focuses on extracting "emerging" issues for proactive policy formulation, we try to minimize the possibility of such biases inherent in the k-means algorithm. Through DTW, we can identify the appropriate number of significant themes of emerging issues in seafood. Specific emergence patterns are expected to reveal the rapidly changing seafood issues and trends. To determine the number of clusters, we employ internal cluster validity assessment through the examination of multiple index metrics. For some indices—Silhouette index, score function, Canlinski–Harabasz index, and Dun index—higher values indicate better performance of clustering method, while for others—Davies–Bouldin index, modified Davies–Bouldin index, and COP index—lower values indicate better performance [12].

#### 2.3.2. Decision on Emerging Issues Candidates

DTW has a significant advantage in identifying candidates for major emerging issues from a wide range of keywords. However, it is also true that the method has limitations in prioritizing individual keywords from multiple candidates of emerging issues. To obtain clues on the emergence time of each candidate, we build on the characteristics of growth curve models, which have strength in modeling natural and social events that involve change over time [12]. In this study, we adopt a logistic growth curve model because it effectively detects the S-curve growth pattern of a specific event, which could be the growth trend of each emerging keyword candidate. Through this process, we select keywords that deserve the highest priority among emerging issue candidates derived by DTW.

We calculate the growth curve of each dominant keyword based on its monthly mention counts. Furthermore, in consideration of the possible high volatility inherent in the monthly counts, we compute another growth curve utilizing accumulated mention counts corresponding to each dominant issue keyword. Through this process, we select

keywords that deserve the highest priority among emerging issue candidates derived by DTW. Equation (1) expresses the equation for the logistic curve.

$$f(x) = \alpha + \frac{\beta - \alpha}{(1 + \exp(\gamma(x - \delta)))^\varepsilon}, \quad (1)$$

where  $\alpha$  is the lower limit,  $\beta$  is the upper limit,  $\delta$  is the inflection point of the curve, and  $\gamma$  is the slope at the inflection point ( $x = \delta$ ).

### 3. Results and Discussion

#### 3.1. Word Cloud

##### 3.1.1. The World

Word clouds, derived from analyzing the frequency of mentions, can visualize various issues in the global seafood industry. Not confined to the issues of seafood's efficacy, various keywords such as species of fish, tariffs, cooking styles, and disease issues such as COVID-19 have emerged. More importantly, differences in major issues have been observed from year to year. Despite their intuitive appeal and conciseness, however, word clouds may not provide quantitative details. Thus, we provide the result of keyword analysis in Appendix A in the form of frequency tables summarizing the top ten mention counts for selected countries (frequency tables for the full sample are available from the authors upon request).

In 2019, *health* and *tariffs* (we italicize the keywords to distinguish them from the ordinary text of this article; even for the same word, the term is in italics if it is used as a particular keyword, while it is not in italics if it is used for general description) received the most mention counts. Particularly, *tariffs* drew high attention because of the US–China trade dispute, which led to the reciprocal imposition of tariffs on seafood. Also, keywords associated with the efficacy of seafood (*protein* and *diet*) as well as the ones for commonly consumed aquatic foods (*shrimp* and *salmon*) were prominent. Additionally, certain types of seafood dishes (*soup* and *sushi*) and environmental issues (*sustainability*, *plastic*, and *food safety*) also garnered high attention.

In 2020, the spread of COVID-19 was rampant. Thus, the global issues were dominated by *coronavirus*. Although the term *coronavirus* was mentioned most often, various keywords were derived from it. For example, *Huanan Seafood Market* in Wuhan, the designated epicenter of the disease, as well as some past epidemics, such as *SARS* and *Ebola*, were mentioned frequently. *Lockdowns*, *scanners*, and *precautions* against coronavirus were also discussed.

*Coronavirus* and *health* remained at the center of attention throughout 2021. Also, COVID-19-related keywords such as *SARS*, *virology*, and *Huanan Market* continued to draw high attention throughout the year. Moreover, it is noteworthy that several novel issues started emerging this year. For example, shipping delays from lockdowns and logistics blockades brought about keywords related to *logistics* and *distribution channels*. In 2021, environmental issues (*plastic consumption*) due to the pandemic and *food safety* gained renewed attention.

In 2022, we saw a shift in the rankings of leading keywords. *Coronavirus* lost its top spot and fell to fifth place, while major keywords that were prevalent before the pandemic (*health*, *food market*, and *protein*) returned to the top. In short, public attention to the function of the seafood market was on the rise again. Figure 2 exhibits the word clouds for the entire sample countries by year. In the following subsections, we provide similar word clouds for sub-sample countries by continent.

##### 3.1.2. Individual Countries

We conducted country-by-country comparisons for selected countries to see whether there were any country-specific issue trends. From our full sample incorporating eleven countries—namely, China, South Korea, and Japan from East Asia; Indonesia, Thailand, and Vietnam from Southeast Asia; France, Italy, and Spain from Europe; and Canada and the US from North America—we report the results from only four countries, China,



Thailand, France, and the US (we report only four country results to keep the discussion concise; word clouds for these countries are presented in Figure 3, the full report on all eleven countries is available from the authors upon request).

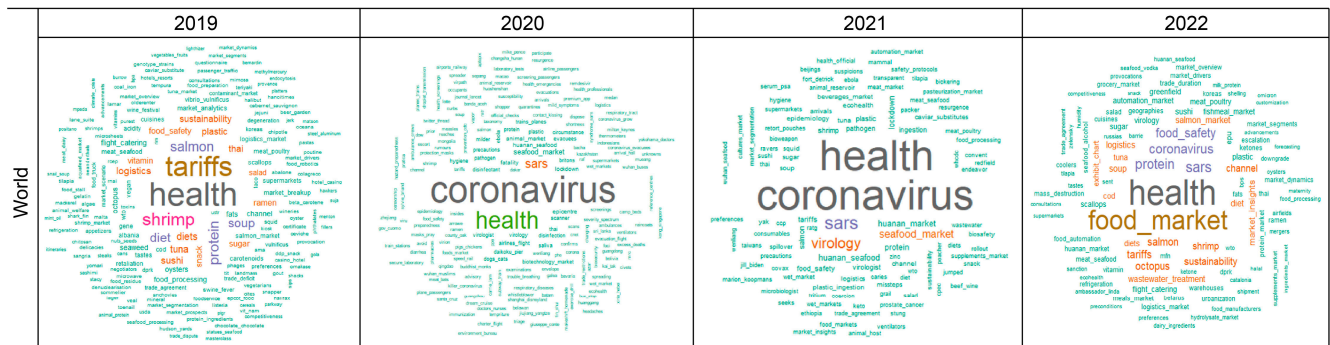


Figure 2. Word cloud for the top 200 words in the entire sample.

In China, 2019 was a year when the US–China trade dispute intensified. Thus, the year’s popular issue was *tariffs*, followed by *health* and *protein*. In 2020, *coronavirus* was by far the most mentioned keyword, followed by *health* and then *SARS*. In 2021, these three keywords continued to dominate. However, the popularity of *coronavirus* had slightly decreased, with *health* being mentioned as much as *coronavirus*, while mentions of *SARS* remained at a similar level as in 2020. Meanwhile, in the second half of 2021, disruptions in the supply chains brought more attention to *logistics* as an emerging issue. In 2022, with the changing perception of aquatic food, interests in *food safety* and *sustainability* were added to the existing interest in *health* functions. Moreover, continuing supply chain disruptions, currently due to the Russia–Ukraine war, kept *logistics* on the list.

As for Thailand’s keywords, in 2019, *soup* related to *Thai* food culture was mentioned the most, followed by *snacks*, *health*, and *sustainability*. For 2020 and 2021, as in other countries, *coronavirus* and *health*-related keywords received the most attention. Interests in *disinfection* and *precautions* during 2020, as well as in *shrimp*, *tuna*, *food safety*, and *sustainability* during 2021, were significantly high. In 2022, *health* and *protein* were mentioned the most, and interest in the *fishmeal market* was high, which might have reflected expectations of the recovery of the seafood processing industry as the pandemic spread subsided. Additionally, an accident on Thailand’s eastern coast increased the mentions of *oil spill*.

In France, during 2019, along with *tariffs* and *health*, keywords related to greenhouse gas (*per capita emissions*) received high attention as discussions of net zero by 2050 developed. In 2020, the issues of *coronavirus* and *health* dominated, but keywords regarding food processing (*food market*, *food processing*, and *food automation*) also received frequent mentions. Throughout 2021, pandemic issues were still dominant, alongside growing interest in the *health supplements market*. The year 2022 was marked by the public’s attention to seafood as a *ketogenic* food in relation to *health* as well as interest in *food safety* and *wastewater treatment*.

In the US, in 2019, keywords such as *health*, *diet*, *fats*, *vitamin*, *vegetarianism*, *food safety*, and the US–China trade-war-related *tariffs* received high attention. In 2020 and 2021, like in other countries, the focus was mostly on *coronavirus*, *health*, and *SARS*. In 2021, along with the changing perception of resource management and environmental pollution, the mention counts of *sustainability* and *plastic* rose rapidly. Also, keywords on seafood variety (*salmon*, *squid*, and *tuna*) regained some attention. In 2022, *tariffs* made it back to the list. Unlike in 2019, this time this was due to the US sanctions against Russia after its invasion of Ukraine. Additionally, several war- and sanctions-related keywords such as *Zelensky*, *Belarus*, *Russia*, *MFN* (most favored nation), *sanctions*, and *seafood alcohol* appeared along with conventional seafood-related keywords like *health*, *food market*, *protein*, *salmon*, and *flight catering*.

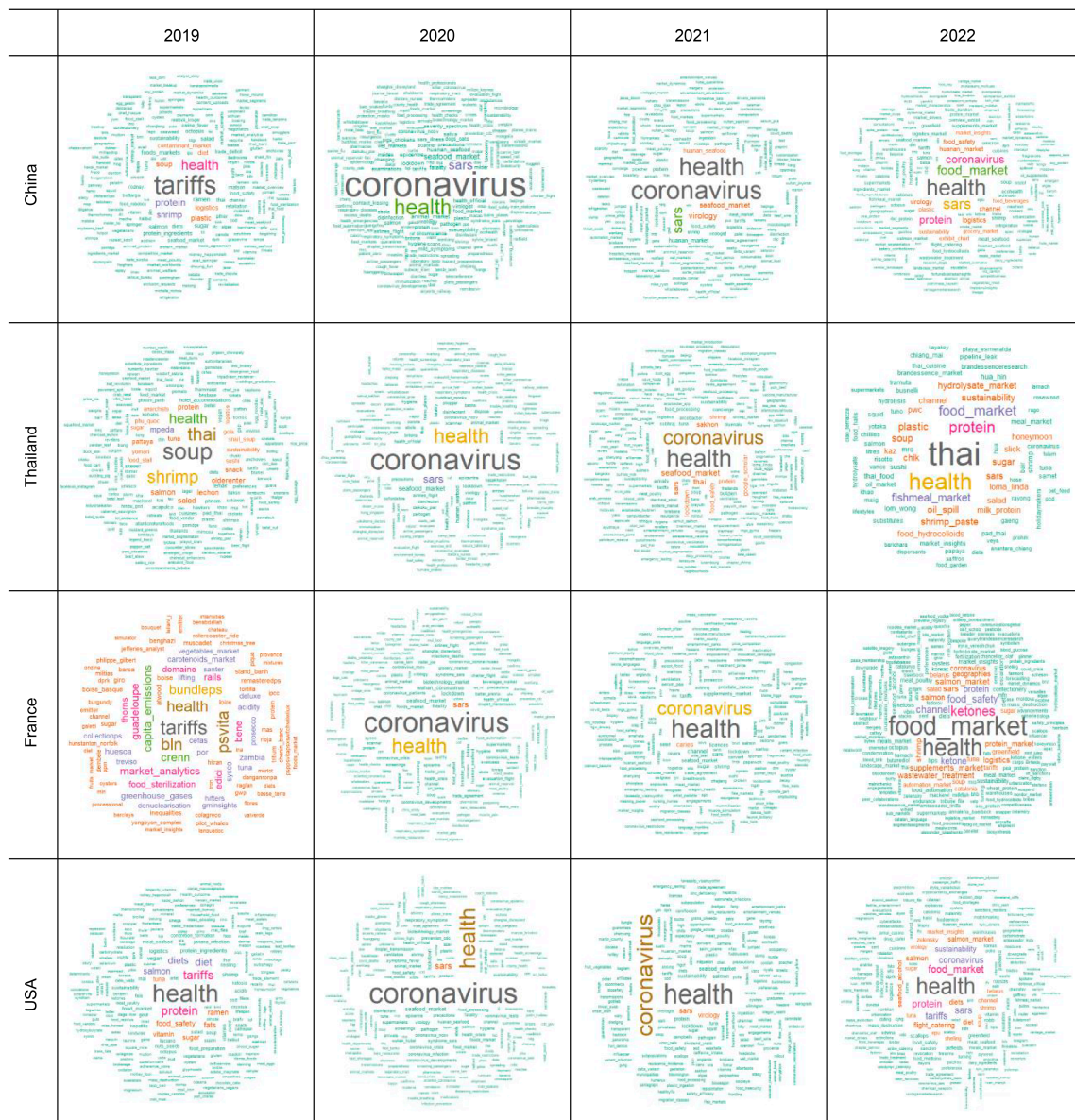


Figure 3. Word cloud for the top 200 words in the entire sample by country (East Asia).

### 3.2. Dynamic Time Warping (DTW)

As described in the previous section, we employ the DTW method to classify emerging issues. With DTW, we categorize keywords into similar growth trends by comparing the changes in their mention counts over time. One advantage of this model is that using statistical techniques minimizes the reliance on researchers’ subjective biases in classifying numerous keywords. We normalize the data to focus on the S-shaped growth pattern of issues, instead of being influenced by extreme mention counts of particular issues (e.g., coronavirus). To determine the number of clusters, we use several types of internal cluster validity indices as our guidelines. Table 2 summarizes the values of these indices based on monthly data, in which higher values of Sil (Silhouette index, [28]), SF (score function, [29]), CH (Calinski–Harabasz index, [30]), and D (Dunn index, [30]) indicate better clustering method performance. In contrast, lower values of DB (Davies–Bouldin index, [30]), DB star (modified Davies–Bouldin index, [31]), and COP index (context-independent optimality and partiality properties index, [30]) indicate better performance. We conduct the same analyses by converting the monthly data into cumulative ones and the results are reported

in Table 3. For the monthly data, we opt for four clusters, while three clusters are chosen for the cumulative data.

**Table 2.** Internal evaluation results (monthly).

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Sil	0.20	0.10	0.11	0.12	0.11	0.10	0.10
SF	0.00	0.00	0.00	0.00	0.0	0.00	0.00
CH	54.12	40.35	34.09	30.09	20.66	20.78	19.31
D	0.17	0.12	0.08	0.07	0.07	0.03	0.08
DB	1.59	2.19	2.09	1.41	2.12	1.49	1.58
DBstar	1.59	2.78	2.82	1.79	2.77	1.80	2.17
COP	0.50	0.46	0.44	0.42	0.41	0.40	0.38

**Table 3.** Internal evaluation results (cumulative).

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Sil	0.34	0.45	0.47	0.40	0.26	0.27	0.23
SF	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CH	96.67	84.85	76.38	64.62	48.17	35.34	39.75
D	0.03	0.04	0.08	0.09	0.00	0.06	0.05
DB	1.02	0.82	0.71	1.44	1.30	1.53	1.33
DBstar	1.02	0.98	0.91	1.80	1.93	2.12	2.13
COP	0.29	0.17	0.12	0.12	0.11	0.12	0.10

Figure 4 shows the normalized frequency of keywords from the sample (the sample targeted the top 50 keywords with the highest frequency each year; among these, those that did not overlap by year are added and used), clustered by monthly data. The bold dashed line in each panel represents the medoid time series in the cluster. The first cluster (Periodical Issues) consists of keywords that have been seen periodically, especially before *coronavirus* issues. The second cluster (Corona Issues) includes keywords with increasing frequencies once the *coronavirus* issues occurred. The third cluster (Post-Corona Issues) includes the keywords with rapidly increasing frequencies after the coronavirus issues. The fourth cluster (Emerging Issue Candidates) is the group of keywords that have drawn attention recently, which can be seen as a candidate group for emerging issues.

Figure 5 displays the normalized frequency of keywords from the cumulative data. The first cluster (Coronavirus Issues) refers to a set of keywords that have significantly increased in frequency due to the COVID-19 outbreak but have not increased in frequency recently. The second cluster (Issues of Constant Interest) corresponds to the set of keywords with a steady frequency throughout the sample period between 2019 and 2022. The third cluster (Issues of Recent Interest) refers to a set of keywords that did not draw much attention in the beginning but received frequent post-pandemic mentions.

Table 4 lists the major keywords for each cluster from the monthly data. Cluster 1 comprises the keywords that receive regular attention, including certain types of food items, health, diet, and trade issues. In Cluster 2, with the outbreak of the pandemic, the main keywords are coronavirus and related issues such as its epicenter, similar diseases, hygiene, and prevention. In Cluster 3, coronavirus and related issues such as lockdown and logistics problems still predominate, but some other keywords related to diet and trade issues also appeared. Cluster 4 focuses on keywords that are currently receiving the most attention, which are potential candidates to become emerging issues. Similarly, Table 5 classifies the keywords in each cluster from the cumulative data. Except for the keywords related to the single coronavirus issue (Cluster 1), the keywords in Clusters 2 and 3 can be identified as emerging issue candidates. To sum up, keywords that are not limited to the pandemic issues but continue to receive attention and interest even after the pandemic can



be emerging issue candidates. Therefore, Cluster 4 on a monthly basis and Clusters 2 and 3 on a monthly cumulative basis can be targeted.

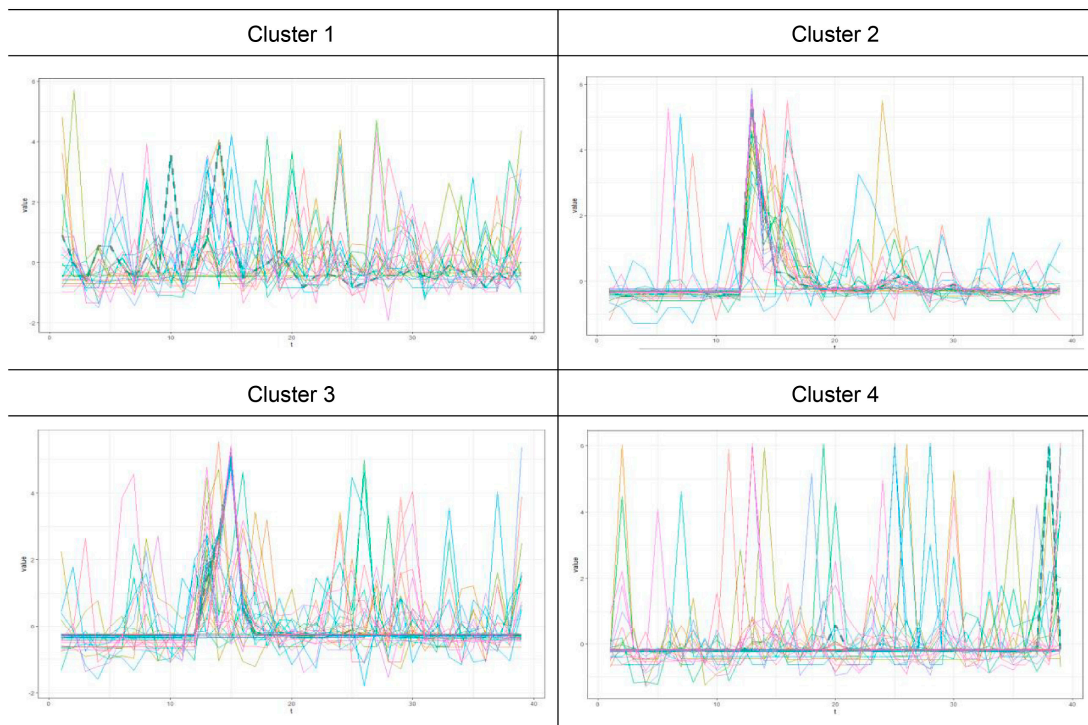


Figure 4. Cluster members by dynamic time warping (monthly).

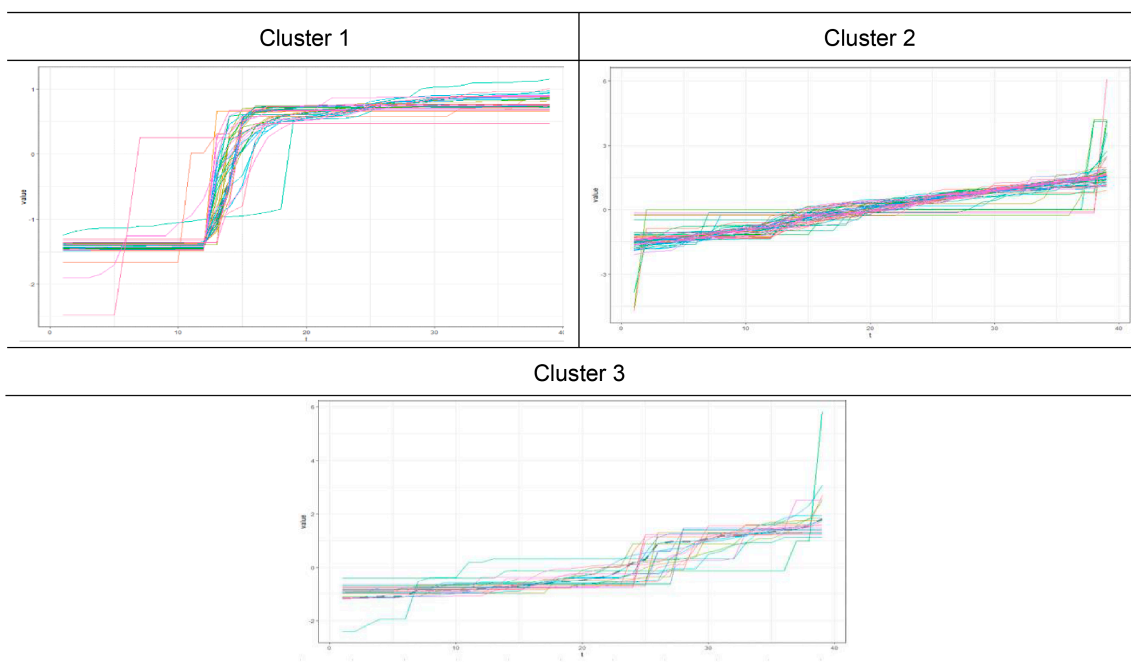


Figure 5. Cluster members by dynamic time warping (cumulative).

**Table 4.** List of keywords by dynamic time warping (monthly).

Group Name	Keywords			
Cluster 1 (Periodical Issues)	Cod	Diets	Eco health	Fats
	Market analytics	Meat market	Meat seafood	Octopus
	Oysters	Ramen	Retaliation	Seaweed
	Snack	Supermarkets	Sustainability	Swine fever
	Tariffs	Tastes	Tuna	Vegan
	Warehouses	Wastewater		
Cluster 2 (Corona Issues)	Acidity	Airlines flight	Animal market	Arrivals
	Beverages_market	Corona	Coronavirus	Disinfectant
	Ebola	Eco health	Epicenter	Epidemiology
	Evacuees	Gene	Health	Huanan Market
	Huanan seafood	Hygiene	Kazakhstan	Meat poultry
	Pathogen	Precautions	Prevention CDC	SARS
	Scanner	Seafood market	Soup	Thai
	Wenliang	Wet markets		
Cluster 3 (Post-Corona Issues)	Automation market	Britons	Ccp	Changing
	Channel	Circumstance	Contact kissing	Diet
	Disinfection	Dogs cats	Escalation	Examinations
	Fatality	Ingestion	Insides	Lockdown
	Logistics	Market segmentation	Milder	Missteps
	Pigs chickens	Plastic	Protein	Protein market
	Salad	Saliva	Salmon market	Scans
	Severity spectrum	Spillover	Sugar	Syndrome SARS
	Taxonomy	Trade agreement	<i>Vibrio Vulnificus</i>	Virologist
	Virology	Vitamin	Whcdc	
Cluster 4 (Emerging Issues Candidates)	Caries	Caviar substitutes	Coercion	Consumables
	Cuisines	EPU	Fishmeal market	Flight catering
	Food processing	Food Safety	Grail	Greenfield
	Ketones	Logistics market	Market insights	Meals market
	Packer	Plastic ingestion	Poacher	Rotort_pouches
	Salmon	Seafood alcohol	Squid	Sushi
	Trade duration	Urbanization	Wastewater treatment	WTO

### 3.3. Growth Model

Once DTW classifies the emerging issue candidates, we reexamine the frequency of individual keyword mentions to identify the most likely emerging issues. To capture the S-curve growth pattern, we use a logistic growth model. In examining the frequency of keyword mentions, we exclude keywords related to advertisements and market research reports (the reason we exclude market research reports' titles and advertisements is that keywords are highlighted by some company's publicity). *Coercion*, *cuisines*, *food safety*, *ketones*, *plastic ingestions*, *seafood alcohol* (Russia–Ukraine-related import ban), *urbanization*, *wastewater treatment*, and *World Trade Organization (WTO)* could be selected based on Figure 6.

**Table 5.** List of keywords by dynamic time warping (cumulative).

Group Name	Keywords			
Cluster 1 (Corona Issues)	Airlines flight	Animal market	Britons	Changing
	Circumstance	Contact kissing	Corona	Coronavirus
	Disinfectant	Disinfection	Dogs cats	Ebola
	Epicenter	Evacuees	Examinations	Fatality
	Health	Huanan seafood	Hygiene	Insides
	Kazakhstan	Lockdown	Milder	Pigs chickens
	Prevention CDC	Saliva	SARS	Scanner
	Scans	Seafood market	Severity spectrum	Swine fever
	Taxonomy	<i>Vibrio Vulnificus</i>	Wenliang	Wet markets
Cluster 2 (Issues of Constant Interest)	Acidity	Arrivals	Ccp	Channel
	Cuisines	Diet	Diets	Eco health
	Epidemiology	EPU	Escalation	Fats
	Fishmeal market	Flight catering	Food safety	Food processing
	Gene	Huanan market	Ketones	Logistics
	Logistics market	Market analytics	Meals market	Meat seafood
	Octopus	Oysters	Pathogen	Plastic
	Protein	Ramen	Retaliation	Salad
	Salmon	Salmon market	Seafood alcohol	Seaweed
	Snack	Soup	Squid	Sugar
	Supermarkets	Sushi	Sustainability	Tariffs
	Tastes	Thai	Trade agreement	Trade duration
	Tuna	Urbanization	Vegan	Virologist
	Virology	Vitamin	Warehouses	Wastewater treatment
	WTO			
	Cluster 3 (Issues of Recent Interest)	Automation market	Beverages_market	Caries
Coercion		Consumables	Food market	Geographies
Grail		Greenfield	Ingestion	Market breakup
Market insights		Market segmentation	Meat market	Meat poultry
Missteps		Packer	Plastic ingestion	Poacher
Protein market		Retort_pouches	Spillover	Wastewater
Whcdc				

*Urbanization* is classified as an emerging issue. Formerly, the term was used in the context of enhanced consumer convenience (e.g., interest in the processed seafood industry and development of e-commerce). However, more recently, *urbanization* has been used as one of the causes of environmental problems including the food crisis. For example, the increased use of plastics due to urbanization causes environmental pollution, which in turn negatively affects the seafood production system. Accordingly, the public interest in industries with relatively fewer environmental problems (e.g., *alternative meat*) has increased.

*Food safety* is an issue of the highest priority in food consumption. While safe and healthy eating habits have always been a subject of high interest, the recent focus on “safe foods” is shifting toward its new role as an export barrier for seafood, justified by quarantines for food hygiene and safety. These food safety issues are highly likely to

become an effective non-tariff barrier for fishery product exports because tariff barriers are to be eliminated gradually by large free trade agreements.



**Figure 6.** Keyword trends using the logistic growth model in Cluster 4.

With the growing volume of international trade, *WTO* became increasingly significant. In particular, along with the increase in the volume of global commerce, interests in trade or dispute topics also have increased as part of the *WTO*-related issues. Specific examples mentioning the *WTO* with regard to the trade of seafood include the US–China trade war in 2018, the *WTO*'s decision to uphold Korea's ban on seafood imports from Fukushima, Japan in 2019, and the Brexit-related EU–UK trade and cooperation agreement in 2020. Another



case is increased mention counts of *WTO* when the Australian government requested the *WTO* to establish a dispute resolution subcommittee regarding China's imposition of additional tariffs on Australian barley in 2021. The mention counts of *WTO* have been continuously on the rise and significantly increased in 2022 as well.

Due to increasing global awareness of healthy diet and nutrition, there is a surge in the popularity of related issues. One such keyword is *Ketones*. With the popular trend focusing on low carb, high fat, and high protein diets, including 2022's so-called bulletproof diet, mention counts of the *ketogenic diet*, a representative of low carb diet with a focus on seafood, has continued to rise. Since health and well-being are positioned as essential values for food consumption led by younger generations, a strategy that meets this demand for "value consumption" is expected to be essential to revitalize the related industries. In addition, with the growing interest in seafood as a source of a healthy diet, recipes for consuming seafood are steadily gaining attention, suggesting that we need to focus on how we consume seafood.

*Wastewater* has been a topic of periodic discussions, but interest in *wastewater treatment* is also increasing. This keyword was sporadically mentioned during 2019–2020 by Greenpeace, an international environmental organization, as they raised awareness of the increasing risk of contamination from Fukushima, Japan. After the International Atomic Energy Agency investigation team visited the site in February 2022 to verify the safety of the Fukushima Daiichi nuclear power plant's offshore discharge plan, articles about protests by the area residents after the visit were mentioned frequently. Throughout our sample period, the majority of mentions about *contaminated water treatment* were related to Fukushima, Japan. Aside from the Fukushima wastewater discharge, the issue of water pollution at a seafood processing facility was garnering attention due to the growing awareness of environmental deterioration.

*Seafood and alcohol* came to attention after the US ban on Russian seafood and alcohol (pollock, salmon, crab, vodka, etc.) in March 2022 as a part of sanctions to punish Russia's invasion of Ukraine, which could generate significant effects on the world's seafood supply chain.

The keyword *coercion* also receives frequent mentions when international disputes occur. The term's mention counts rose significantly between 2020 and 2021 with China's sanctions on Australian fruits and aquatic products, and also after Russia's invasion of Ukraine in 2022. These cases also indicate some signs that non-tariff barriers between countries are increasing.

Keywords pertaining to *plastics ingestion* started gaining attention with the increased awareness and interest in climate change and environmental pollution after the Paris Climate Agreement in 2015. Media attention to a 2019 report [32] brought this issue to the forefront of environmental discussions. After 2021, *food safety* and *sustainability* emerged as major issues as part of the discussions regarding this keyword. As the awareness of environmental problems increased, keywords related to *plastic consumption* also gained increased attention.

#### 4. Conclusions

We have used keyword analysis, jointly with DTW and the logistic growth model to investigate the major issues and emerging issues in the seafood industry during January 2019–March 2022. To identify emerging issues, we first collected news articles and extracted the candidate issues via text mining and DTW. Among the extracted candidates, we derived validated issues by fitting the logistic growth model. During the process, we were able to reveal some pressing issues.

This study contributes to the literature on horizon scanning research and its practical applications in seafood industries. From the academic perspective, this study offers a novel and objective way of identifying emerging issues in fisheries research, while existing studies have relied on media attention or researchers' judgments for issue identification. The methodology we employed is highly applicable to other fields of research with a similar

interest. Our results show the major issues that were brought into the spotlight each year and highlight the imminent and significant issues in policy formulation for the future. From our analysis, we found that policymakers need to respond to the public's high interests in food safety, the changing environment, possible future trade disputes, and the value of marine products as nutritional components.

Our findings also offer several practical insights for policymaking. First, there is a need for a sustainable seafood supply system that fulfills consumer demand and enhances the value of seafood. For instance, over time, food safety standards will advance and become more refined. Thus, establishing a traceable seafood supply chain and safety management system will be necessary. Moreover, considering the growing interest in seafood consumption due to its nutritional value, it is necessary to provide user-friendly recipes so that consumers can increase their familiarity with seafood. Second, policymakers need to devise plans to cope with environmental changes. Climate change may negatively affect the productivity of capture fisheries and related industries. In tackling climate change, the fundamental approach would be carbon emission control. However, parallel to this fundamental response, policymakers can move society one step further toward securing its supply of fishery products by fostering the development of aquaculture industries. Supporting investment in food technology could be an alternative. Third, a comprehensive review of recent trade conflicts is necessary. Getting into bilateral trade conflicts may leave the involved countries with aggravated consequences, especially when their reliance on each other is high. Facing the heightened uncertainty in the global playing fields, one cannot neglect the importance of diversifying trade partners and building up a systematic process of risk management.

The findings of this study have high external validity due to the extensive use of a large database covering more than 80,000 sources of media coverage. One caveat is that, to use a sufficiently large international database, the main language was limited to English. However, detailed records of a particular market or country are likely to be better preserved in a local language database. It is necessary to develop a model that allows balanced coverage of data sources that minimize potential linguistic, economic, or political biases. Additionally, in deriving emerging issues, our focus on quantitative methods such as frequency analysis was intended to gain independence from researchers' subjective biases. However, future research may benefit from incorporating quantitative and qualitative—e.g., expert opinions—evaluations in the process. Also, delving into alternative text analysis methods other than frequency analysis would offer further insights into identifying emerging issues. Finally, this study focuses on identifying issues in a broad category of the seafood sector. In terms of practical policy formulation, applying the model to specialized subfields of the seafood sector would be more useful. For example, keywords related to processed seafood or particular subspecies of seafood such as tuna and salmon could be identified and categorized as emerging issues for policy development and refinement.

**Author Contributions:** Conceptualization, K.H.; Methodology, K.H. and J.Y.; Formal analysis, K.H., J.Y. and K.C.; Investigation, K.H. and J.Y.; Writing—original draft, K.H.; Writing—review and editing, J.Y. and K.C.; Visualization, K.H., J.Y. and K.C.; Project administration, K.H. and K.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Korea Maritime Institute Fund (Ocean and Fisheries Future Risk: 2nd Year Performance). Keunsuk Chung acknowledges the support from UNIST Education and Research Innovation Fund for this research.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the authors. The data are not publicly available due to privacy.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

Calinski–Harabasz index (CH); Comprehensive and Progressive Partnership for the Trans-Pacific Partnership (CPTPP); Davies–Bouldin index (DB); Dunn index (D); dynamic time warping (DTW); Food and Agriculture Organization of the United Nations (FAO); modified Davies–Bouldin index (DBstar); score function (SF); Silhouette index (Sil); World Trade Organization (WTO).

## Appendix A

The following tables provide the results of keyword analysis for selected four countries in the form of frequency tables.

**Table A1.** The results of keyword analysis (China).

NO	2019		2020		2021		2022	
	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.
1	tariffs	1002	coronavirus	69,524	coronavirus	4994	health	1052
2	health	498	health	40,238	health	4746	sars	684
3	protein	312	sars	19,020	sars	2526	food_market	554
4	shrimp	252	seafood_market	7586	virology	1156	coronavirus	490
5	soup	192	fatality	3370	seafood_market	922	protein	460
6	plastic	174	huanan_seafood	2620	huanan_seafood	644	logistics	220
7	logistics	156	animal_market	2600	huanan_market	584	huanan_market	214
8	diet	142	lockdown	2508	protein	468	virology	206
9	contaminant_market	134	epicenter	2456	food_safety	388	food_safety	196
10	foods_markets	122	virology	2196	lockdown	360	channel	182

Unit: Number.

**Table A2.** The results of keyword analysis (Thailand).

NO	2019		2020		2021		2022	
	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.
1	soup	346	coronavirus	35,134	health	776	thai	178
2	thai	264	health	23,184	coronavirus	624	health	120
3	shrimp	250	sars	9532	seafood_market	194	protein	74
4	health	194	seafood_market	3962	thai	174	food_market	64
5	mpeda	94	huanan_seafood	2090	sars	150	fishmeal_market	46
6	salmon	86	epicenter	1380	sakhon	134	plastic	40
7	salad	84	lockdown	1248	lockdown	128	soup	38
8	lechon	80	scanner	1188	food_safety	108	oil_spill	36
9	olderenter	80	thai	1172	protein	102	sustainability	36
10	protein	74	airlines_flight	786	shrimp	102	shrimp_paste	36

Unit: Number.

**Table A3.** The results of keyword analysis (France).

NO	2019		2020		2021		2022	
	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.
1	tariffs	56	coronavirus	12,098	foods_market	422	food_market	434

Table A3. Cont.

NO	2019		2020		2021		2022	
	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.
2	bln	48	health	8150	breader_premixes	392	health	384
3	psvita	46	sars	2458	britons	360	ketones	200
4	health	44	seafood_market	948	food_processing	356	food_safety	140
5	bundleps	36	lockdown	784	sars	150	channel	130
6	crenn	32	wuhan_coronavirus	762	sakhon	134	ketone	116
7	capita_emissions	30	foods_market	422	lockdown	128	protein	112
8	guadeloupe	28	breader_premixes	392	food_safety	108	salmon_market	104
9	market_analytics	28	britons	360	protein	102	supplements_market	101
10	edici	24	food_processing	356	shrimp	102	sars	98

Unit: Number.

Table A4. The results of keyword analysis (USA).

NO	2019		2020		2021		2022	
	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.	Keywords	Freq.
1	health	1044	coronavirus	19,696	health	2684	health	798
2	tariffs	474	health	14,776	coronavirus	2214	food_market	314
3	protein	458	sars	4098	sars	472	protein	306
4	diet	374	seafood_market	1890	virology	366	tariffs	264
5	diets	330	lockdown	1138	salmon	240	sars	264
6	salmon	286	protein	746	seafood_market	228	coronavirus	260
7	food_safety	220	tariffs	508	sustainability	192	sustainability	200
8	fats	218	epicenter	442	protein	192	salmon	182
9	ramen	216	precautions	428	lockdown	182	flight_catering	170
10	vitamin	178	virology	422	zinc	166	diets	168

Unit: Number.

## References

- Gersl, A.; Hermanek, J. Indicators of financial system stability: Towards an aggregate financial stability indicator? *Prague Econ. Pap.* **2008**, *2008*, 127–142. [\[CrossRef\]](#)
- Hines, A.; Baldwin, B.P.; Bengston, D.N.; Crabtree, J.; Christensen, K.; Frankowski, N.; Schlehuber, L.; Westphal, L.M.; Young, L. Monitoring emerging issues: A proposed approach and initial test. *World Futures Rev.* **2021**, *13*, 195–213. [\[CrossRef\]](#)
- Marsden, G.; Kelly, C.; Snell, C. Selecting indicators for strategic performance management. *Transp. Res. Rec.* **2006**, *1956*, 21–29. [\[CrossRef\]](#)
- Spangenberg, J.H. Scenarios and Indicators for Sustainable Development: Towards a Critical Assessment of Achievements and Challenges. *Sustainability* **2019**, *11*, 942. [\[CrossRef\]](#)
- World Development Indicators*; The World Bank: Washington, DC, USA, 2023.
- Natale, F.; Borrello, A.; Motova, A. Analysis of the determinants of international seafood trade using a gravity model. *Mar. Policy* **2015**, *60*, 98–106. [\[CrossRef\]](#)
- Barboza, L.G.A.; Vethaak, A.D.; Lavorante, B.R.B.O.; Lundebye, A.-K.; Guilhermino, L. Marine microplastic debris: An emerging issue for food security, food safety and human health. *Mar. Pollut. Bull.* **2018**, *133*, 336–348. [\[CrossRef\]](#) [\[PubMed\]](#)
- Miraglia, M.; Marvin, H.J.P.; Kleter, G.A.; Battilani, P.; Brera, C.; Coni, E.; Cubadda, F.; Croci, L.; De Santis, B.; Dekkers, S.; et al. Climate change and food safety: An emerging issue with special focus on Europe. *Food Chem. Toxicol.* **2009**, *47*, 1009–1021. [\[CrossRef\]](#) [\[PubMed\]](#)
- Food and Agriculture Organization of the United Nations (FAO), Part 4 Emerging issues and outlook. In *The State of World Fisheries and Aquaculture*; FAO: Rome, Italy, 2022; pp. 195–223.
- Molitor, G.T. How to anticipate public-policy changes. *SAM Adv. Manag. J.* **1977**, *42*, 4–13.



11. Rhemann, M. Understanding disruption through Molitor's models. *World Futures Rev.* **2018**, *10*, 34–37. [[CrossRef](#)]
12. Han, K.; Leem, K.; Choi, Y.R.; Chung, K. What drives a country's fish consumption? Market growth phase and the causal relations among fish consumption, production and income growth. *Fish. Res.* **2022**, *254*, 106435.
13. Adamuthe, A.C.; Thampi, G.T. Technology forecasting: A case study of computational technologies. *Technol. Forecast. Soc. Change* **2019**, *143*, 181–189. [[CrossRef](#)]
14. Kucharavy, D.; De Guio, R. Application of S-shaped curves. *Procedia Eng.* **2011**, *9*, 1877–7058. [[CrossRef](#)]
15. Krigsholm, P.; Riekkinen, K. Applying Text Mining for Identifying Future Signals of Land Administration. *Land* **2019**, *8*, 181. [[CrossRef](#)]
16. Bai, X.; Zhang, X.; Li, K.X.; Zhou, Y.; Yuen, K.F. Research topics and trends in the maritime transport: A structural topic model. *Transp. Policy* **2021**, *102*, 11–24. [[CrossRef](#)]
17. Hase, V.; Mahl, D.; Schäfer, M.S.; Keller, T.R. Climate change in news media across the globe: An automated analysis of issue attention and themes in climate change coverage in 10 countries (2006–2018). *Glob. Environ. Change* **2021**, *70*, 102353. [[CrossRef](#)]
18. Velvizhi, V.; Billewar, S.R.; Londhe, G.; Kshirsagar, P.; Kumar, N. Big data for time series and trend analysis of poly waste management in India. *Mater. Today Proc.* **2021**, *37*, 2607–2611. [[CrossRef](#)]
19. Kurian, D.; Sattari, F.; Lefsrud, L.; Ma, Y. Using machine learning and keyword analysis to analyze incidents and reduce risk in oil sands operations. *Saf. Sci.* **2020**, *130*, 104873. [[CrossRef](#)]
20. Sharma, D.; Kumar, B.; Chand, S. Trend analysis in machine learning research using text mining. In Proceedings of the 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 12–13 October 2018; pp. 136–141.
21. Carbonell, J.; Sánchez-Esguevillas, A.; Carro, B. Assessing emerging issues. The external and internal approach. *Futures* **2015**, *73*, 12–21. [[CrossRef](#)]
22. Wang, Q. A bibliometric model for identifying emerging research topics. *J. Assoc. Info. Sci. Technol* **2018**, *69*, 290–304. [[CrossRef](#)]
23. Wever, M.; Shah, M.; O'Leary, N. Designing early warning systems for detecting systemic risk: A case study and discussion. *Futures* **2022**, *136*, 102882. [[CrossRef](#)]
24. Dator, J. Emerging Issues Analysis: Because of Graham Molitor. *World Futures Rev.* **2018**, *10*, 5–10. [[CrossRef](#)]
25. Keogh, E.J.; Pazzani, M.J. Scaling up dynamic time warping to massive datasets. In *Principles of Data Mining and Knowledge Discovery*; Żytkow, J.M., Rauch, J., Eds.; Springer: Berlin/Heidelberg, Germany, 1999; pp. 1–11.
26. Müller, M. Dynamic Time Warping. In *Information Retrieval for Music and Motion*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 69–84.
27. Arora, P.; Deepali; Varshney, S. Analysis of K-Means and K-Medoids Algorithm For Big Data. *Procedia Comput. Sci.* **2016**, *78*, 507–512. [[CrossRef](#)]
28. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
29. Saitta, S.; Raphael, B.; Smith, I.F. *A Bounded Index for Cluster Validity*; Springer: Berlin/Heidelberg, Germany, 2007.
30. Arbelaitz, O.; Gurrutxaga, I.; Muguerza, J.; Pérez, J.M.; Perona, I. An extensive comparative study of cluster validity indices. *Pattern Recognit.* **2013**, *46*, 243–256. [[CrossRef](#)]
31. Kim, M.; Ramakrishna, R.S. New indices for cluster validity assessment. *Pattern Recognit. Lett.* **2005**, *26*, 2353–2363. [[CrossRef](#)]
32. Advisors, D.; de Wit, W.; Bigaud, N. *No Plastic in Nature: Assessing Plastic Ingestion from Nature to People*; World Wildlife Fund for Nature: Gland, Switzerland, 2019.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.