

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Mutanov, Galimkair; Omirbekova, Zhanar; Shaikh, Aijaz A.; Issayeva, Zhansaya

**Title:** Sustainability-Driven Green Innovation : Revolutionising Aerospace Decision-Making with an Intelligent Decision Support System

**Year:** 2024

**Version:** Published version

**Copyright:** © 2023 by the authors

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Mutanov, G., Omirbekova, Z., Shaikh, A. A., & Issayeva, Z. (2024). Sustainability-Driven Green Innovation : Revolutionising Aerospace Decision-Making with an Intelligent Decision Support System. *Sustainability*, 16(1), Article 41. <https://doi.org/10.3390/su16010041>

## Article

# Sustainability-Driven Green Innovation: Revolutionising Aerospace Decision-Making with an Intelligent Decision Support System

Galimkair Mutanov <sup>1</sup>, Zhanar Omirbekova <sup>1,2</sup> , Aijaz A. Shaikh <sup>3,\*</sup>  and Zhansaya Issayeva <sup>4</sup>

<sup>1</sup> Institute of Information and Computational Technologies, P.O. Box 050010 Almaty, Kazakhstan; galimkairmutanov@gmail.com (G.M.); zh.omirbekova@ipic.kz (Z.O.)

<sup>2</sup> Department of Computer Science, Al-Farabi Kazakh National University, P.O. Box 050040 Almaty, Kazakhstan

<sup>3</sup> Department of Marketing, Jyväskylä University School of Business and Economics, University of Jyväskylä, P.O. Box 35, FI-40014 Jyväskylä, Finland

<sup>4</sup> Faculty of Oriental Studies, Al-Farabi Kazakh National University, P.O. Box 050040 Almaty, Kazakhstan; zhansaya.issayeva@kaznu.edu.kz

\* Correspondence: aijaz.a.shaikh@ju.fi

**Abstract:** Green innovation refers to developing and implementing new technologies, practices, products, and processes that promote sustainability and reduce environmental impacts. This article postulates the conceptualisation and implementation of an intelligent decision support system (IDSS) tailored to the aerospace technology sector. The data were collected from open sources such as social media and analyzed using the natural language processing tool. The envisaged IDSS is a comprehensive and seamlessly integrated platform designed to undergird decision-making, problem-solving, and research initiatives within the aerospace industry. Catering to the sector's engineers, technicians, and managerial cadres, it aims to unravel complex datasets, proffer incisive analyses, and furnish prudent advice and recommendations. Its multifaceted capabilities range from data search and optimisation to modelling and forecasting. With an emphasis on harmonious integration with extant aerospace systems, it strives to provide engineers and technicians with enriched data insights. Moreover, its design ethos is centred on user-friendliness, underscored by an intuitive graphical interface that expedites seamless access and utilisation. Ultimately, the envisioned IDSS will augment the aerospace industry's analytical prowess and will serve as a potent instrument for effective decision-making.

**Keywords:** sustainable green innovations; aerospace; intelligent decision support system



**Citation:** Mutanov, G.; Omirbekova, Z.; Shaikh, A.A.; Issayeva, Z. Sustainability-Driven Green Innovation: Revolutionising Aerospace Decision-Making with an Intelligent Decision Support System. *Sustainability* **2024**, *16*, 41. <https://doi.org/10.3390/su16010041>

Academic Editors: Darjan Karabasevic and Francesco Tajani

Received: 8 September 2023

Revised: 22 November 2023

Accepted: 8 December 2023

Published: 20 December 2023



**Copyright:** © 2023 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

Green innovation refers to developing and implementing new technologies, practices and products that promote sustainability and reduce environmental impacts. It encompasses many fields, from renewable energy and clean transportation to waste reduction and eco-friendly manufacturing. Green innovation is essential for addressing the urgent challenges of climate change and ensuring a sustainable future for our planet. By leveraging the latest advances in artificial intelligence (AI), machine learning (ML) and other cutting-edge technologies, we can accelerate green innovation and create a better world for future generations [1,2].

As the aerospace industry continues to evolve and grow, especially after the launch of SpaceX by Elon Musk in 2002, there is an increasing need for intelligent decision-making systems that can help streamline processes and improve overall efficiency. Moreover, recent advancements seen in the shape of the proliferation of artificial intelligence applications, tools, and systems, the need for developing a new intelligent decision support system (IDSS) was felt more recently. This is where an improved version of IDSS comes in. By harnessing

the power of AI and ML, this system can revolutionise how aerospace professionals make decisions, providing real-time insights and recommendations to help identify problems and opportunities before they become critical. Whether one is working in aircraft design, logistics, or any other area in the aerospace field, the IDSS can help one stay ahead of the curve and make better, more informed decisions every step of the way.

The aerospace sector, characterised by a dynamic landscape, has embraced decision support systems (DSSs) as instruments for discerning and tracking emergent technological paradigms poised to shape its trajectory. These systems are pivotal in steering strategic resource allocation decisions by corporations and government entities. The present discourse on such systems is focussed on delineating their conceptual underpinnings, particularly in the context of foresight research within the aerospace industry. Our research considered burgeoning trends and technologies within space to assess their imminent and transformative potential. Following this evaluation, a structured framework was devised to monitor and manage the trajectories of these nascent paradigms. Central to this endeavour was the imperative to furnish an array of stakeholders with cogent insights requisite for informed strategic determinations.

Revolutionising aerospace decision-making with an IDSS is crucial for ensuring the safety and efficiency of air travel. There is a significant research gap in this area, particularly in Kazakhstan, where the development and implementation of such systems are still in their early stages. By implementing advanced technologies such as AI and ML, we can improve the decision-making process and enhance the overall performance of aerospace systems. This is essential for meeting the growing demand for air travel while reducing environmental impacts and promoting sustainability.

Therefore, the present study was conducted to develop an information system utilising AI technology that would support decision-making in the aerospace industry. Specifically, the system focusses on enhancing safety and efficiency in air travel. The following research questions were considered, which are essential when developing an IDSS utilising AI technology to support decision-making in the aerospace industry:

1. What are the primary safety concerns in the aerospace industry, and how can AI technology be used to address them?
2. How can AI-powered systems be used to optimise air traffic management and enhance air travel efficiency?
3. How can ML algorithms be used to improve the performance and reliability of aerospace systems?
4. What ethical considerations must be considered when developing an AI-powered IDSS for the aerospace industry?

The present study has made significant contributions. By addressing the research questions outlined above, it has provided insights into how AI technology can address safety concerns in the aerospace industry, optimise air traffic management, and minimise environmental impacts. Additionally, it has provided valuable information on how ML algorithms can improve the performance and reliability of aerospace systems and the ethical considerations that need to be considered when developing an AI-powered IDSS for the aerospace industry. These findings can benefit the industry and the academic community by allowing them to better understand how AI technology can be effectively utilised in the aerospace industry, leading to safer and more efficient aerospace technologies.

The rest of the paper is organised as follows. The Section 1 provides an overview of the research questions and objectives of the present study. The Section 2 examines previous studies on AI technology in the aerospace industry, including its potential benefits and challenges. Section 3 describes the research design, data collection, analysis methods, and ethical considerations. The Results section presents the study's findings, including insights into how AI technology can address safety concerns in the aerospace industry, optimise air traffic management, and minimise environmental impacts. Lastly, the Section 6 summarises the main findings, discusses their implications for the industry and academic community, and provides directions for future research.

## 2. Literature Review

### 2.1. Green Innovation

As the global community becomes increasingly concerned about the impact of human activities on the environment, various subsectors of the economy and industry, including the aerospace subsector, are recognising the importance of implementing sustainable practices and reducing their carbon footprints. Green innovations, which encompass a range of solutions, from renewable energy sources to eco-friendly packaging options, have emerged as an important topic in the business world [3].

Research has shown that companies that implement green innovations not only benefit the environment but can also see a positive impact on their bottom line. Consumers are becoming more environmentally conscious and are often willing to pay a premium for sustainable and environmentally friendly products. Moreover, sustainable practices can lead to long-term savings by reducing energy and waste management costs [4]. For example, the authors of Ref. [5], while examining sustainability in the aerospace industry, suggested a novel decision support method that uses both qualitative and quantitative analysis.

The extractive approach is deemed easier as it achieves good grammar and accuracy levels by copying large chunks of text from a source document; however, the complex abilities essential for high-quality generalization, including paraphrasing, generalization, and incorporating real-world knowledge, are feasible only within an abstract framework, and despite the increased challenges, notable advancements have been made, particularly with recent developments in deep learning. In this context, the development of a novel decision support method that integrates both qualitative and quantitative analysis raises the prospect of achieving long-term savings in the aerospace industry, providing a more comprehensive and efficient approach to decision-making.

Companies that prioritise sustainability are also viewed more favourably by consumers, leading to increased brand loyalty and a positive reputation. Moreover, implementing green innovations is essential for meeting global sustainability goals and ensuring a habitable planet for future generations [6].

### 2.2. Intelligent Decision Support Systems

IDSSs have recently gained significant popularity in the aerospace industry. An IDSS is a software system that utilises AI and other cutting-edge technologies to assist decision-makers in solving complex problems and making informed decisions. IDSSs collect and analyse data from different sources to create meaningful insights for analysts. AI-based DSSs then make recommendations and communicate them to users in an understandable way [7]. At present, most organisations can easily obtain enough relevant data, but analysing these data and determining what to do with the insights derived from them have become the most difficult and time-consuming tasks. One of the problems that any organisation may face is knowing how to obtain the right document classification due to a shortage of data [8].

IDSSs can be used in various aerospace fields, such as aircraft design and air traffic control. At its nucleus, an IDSS in the aerospace domain epitomises a cognitive computing mechanism leveraging AI to augment the decision-making capacities of experts, researchers, and engineers. By seamlessly amalgamating data, models, and analytical tools, this system can establish a holistic arsenal for dissecting intricate quandaries. The system's utility manifests in diverse aerospace domains, encompassing strategic planning, launch logistics, satellite operations, and mission command, and further extends to development strategy blueprinting, resource allocation, and scheduling. Its capabilities also extend into modelling scenarios to evaluate the performance of novel technologies and to gauge the repercussions of proposed modifications within the space ecosystem. The accelerating trajectory of space technologies is inextricably linked to the need for effective information retrieval mechanisms.

Several researchers have explored the application of IDSS in aerospace. For example, Ref. [9] investigated the benefits of using an IDSS in aircraft design. They highlighted how

an IDSS can help designers create more efficient and safe aircraft designs by exploring the potential of an IDSS in improving air traffic control operations and reducing the risk of accidents. They noted that an IDSS can assist air traffic controllers in making more accurate and timely decisions. These studies demonstrated the versatility and potential of IDSSs in the aerospace industry.

### 2.3. Artificial Intelligence and Ethical Issues

The prevalence of AI in our daily lives has grown significantly in recent years [10], perhaps due to the Fourth Industrial Revolution. Defined as the ability of machines to carry out tasks through intelligent, human-like behaviour [11], AI has undoubtedly revolutionised industries, offering unprecedented capabilities in data analysis, automation, and decision-making. From healthcare to finance, AI-powered solutions have become an integral part of our lives. However, with the increased popularity and usefulness of AI, a wide range of ethical questions have arisen, such as how AI can be taught to make moral decisions, how its decision-making processes can be made adequately visible to humans, and who should be held accountable for the conclusions it arrived at [10]. AI's exponential expansion (e.g., ChatGPT) and pervasive influence make it even more important and pressing to resolve these moral dilemmas [11], but the current empirical research landscape and the state of ethical discourse are too conceptually disorganised and fragmented to reach a conclusion. This will perhaps be more evident in the coming days, when generative AI and similar tools are expected to become more common, thus providing a more stable opinion on the ethical issues surrounding AI.

## 3. Methods

Today, one of the best ways to track technological trends is to monitor open data in the global information environment. In other words, all the necessary information is contained in scientific publications, patents, analytical reports, news media, and social networks of experts in the aerospace industry. Due to the decentralised structure of science and technology, the global process of innovative development is reflected in openly published documents. The analysis will help identify new technological trends and track the progress of existing ones [6].

### 3.1. Datasets

Our developed IDSS uses various data collection methods to provide comprehensive and reliable information for aerospace decision-making. It utilises web crawlers to search for and extract information related to scientific research and technological innovations in the aerospace industry. It also employs a research rubric to classify and organise data based on their thematic affiliations, making it easier to analyse and use the data. The system is integrated with academic library application programming interfaces (APIs), which provide access to authoritative scientific articles, publications and research. Additionally, it uses patent databases and integrates with the U.S. Patent and Trademark Office to track new patents, inventions, and technical developments related to the aerospace industry. By equipping the developed IDSS with these functionalities, we strive to provide the most up-to-date and comprehensive information to support decision-making in the aerospace sector [9].

To develop an intelligent decision-making system, we collapsed data and divided them into several categories, as follows:

- Online databases: The NASA Technical Report Server, the NASA Astrophysical Data System, the European Space Agency's space science portal, and other similar online databases are repositories that grant access to an expansive collection of documents spanning the space technology spectrum.
- Search engines: Google, Bing, and other conventional search engines offer avenues for scouring space technology-related information that may provide pertinent insights.

- Journals and publications: Specialised journals and publications, such as the Journal of Space Technology, International Journal of Space Science and Technology, and Space Technology International Forum, furnish granular insights into the dynamic flux characterising the ever-evolving landscape of space technology.
- Social media platforms: Twitter, Reddit, Quora, and other social media platforms have emerged as alternative forums for probing and extracting information pertinent to the space technology narrative.
- Specialised databases: The International Space Station Experiment Database, the Space Research Network, and other databases meticulously tailored to the space domain offer an array of datasets germane to the realm of space technologies.

### 3.2. Data Collection

Methods for collecting relevant information and creating a pilot registry of searchable sources of information are essential steps in the research process. This approach allows one to ensure the efficiency and accuracy of data collection and to assess the potential value of the collected data for a specific research task.

One type of data collection method is the target-oriented data collection method, which involves collecting relevant information and creating a registry of sources that allows researchers to focus on specific aspects of the study. This approach helps minimise information noise and focusses efforts on key sources. When using data collection methods, data quality improvement is applied, which allows for optimising the selection of sources based on their relevance and contributes to the collection of better information, which increases the reliability of the research results.

An example of a method for creating a pilot registry is listing keywords and terms related to the research topic and using them to search for information in databases, on-line archives, and other resources. Collecting expert opinions or conclusions is also very important. Domain experts must be consulted to identify the most relevant and authoritative sources.

It is also important to create a small pilot registry from several sources selected based on their relevance. This will enable the evaluation of data quality and the formulation of a larger plan for collecting information. The main approach is a systematic literature review—that is, analysing and summarising the results of previous studies to identify the most significant sources and assess their relevance [12].

#### 3.2.1. Metadata and Rating Analysis

Metadata (e.g., citations, ratings, reviews) can be used to identify popular and reputable sources. Collecting relevant information and creating a registry help focus research, improve data quality, and optimise resource usage.

Regarding existing monitoring problems and the relevance of collecting relevant industry information, the aerospace industry faces challenges in real-time monitoring due to its dynamic nature, and the relevance of collecting industry-specific information lies in addressing technological advancements, supply chain complexities, and regulatory changes to enhance decision support systems for optimizing processes, ensuring compliance, and maintaining competitiveness.

The basis of all analytical systems is the data that are increasingly available in the public domain on the global internet, and the effectiveness of the developed AI-based expert DSS for the space industry is primarily determined by the quality of the processed data. However, the exponential growth of internet resources, the rapid development of social networks, and the transition of almost all media outlets to the internet lead to the repeated duplication of information and information noise, which significantly complicates the search for relevant information both for users and search robots.



### 3.2.2. Search Engine Data

The Google search engine, which accounts for more than 62% of the global search market, generates millions or billions of responses for space-related queries in a split second (see Table 1).

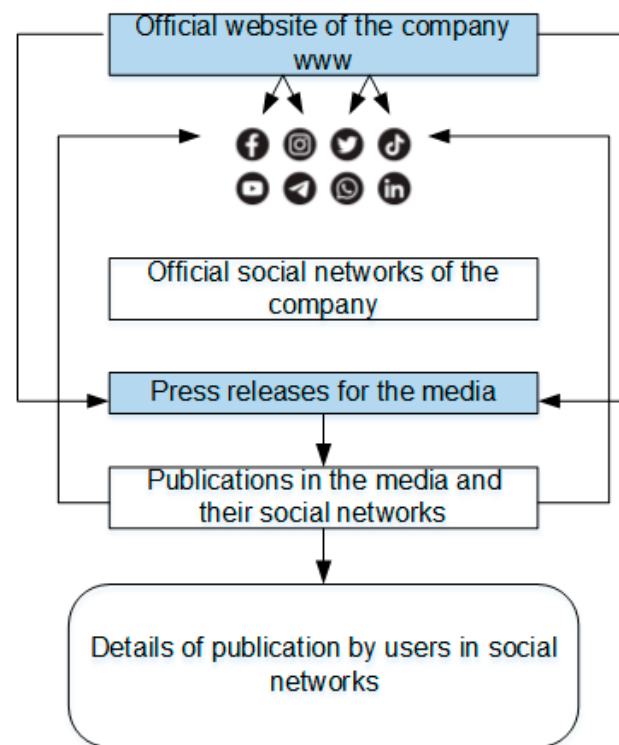
**Table 1.** Output of responses to queries in the Google search engine on space topics.

Search Queries	Number of Responses	Search Time
Space technologies in the world	6.31 million	0.5 s
Space technologies in the world	4.31 billion	0.65 s
Satellite manufacturers	11 million	0.48 s
Satellite manufacturers	149 million	0.58 s
Spaceports in the world	1.47 million	0.56 s
Cosmodromes in the world	6.68 million	0.59 s
Production of space rockets	355 thousand	0.45 s
Production of space rockets	8.54 million	0.64 s
Space rocket manufacturers	127 thousand	0.45 s
Space rocket manufacturers	18.3 million	0.67 s
Space technology innovations	1.56 million	0.45 s
Space technology innovations	186 million	0.74 s
Space technology market	2.14 billion	0.81 s
Space technology market	829 thousand	0.53 s
Space technology market technology	10.55 million	0.59 s

Note: Own research as of 5 November 2021.

The number of search results or responses to space-related queries as of 12 October 2021 differs significantly from that which shows in Table 1 as of 5 November 2021. The number of responses to the query regarding the space technologies in the world on 12 October 2021 was 1.26 billion, while on 5 November 2021, the search engine issued 4.31 billion responses, about 3.3 times more, which indicates search problems. In contrast, for the space technology market technology query, the results were optimised in 2021, amounting to only 10.55 million responses, compared to 452 million responses in April 2018. Google is constantly improving its algorithms and is one of the world leaders in AI development investment. However, the problem of providing relevant information in response to queries has not yet been solved. The search speed for queries and keywords does not provide advantages as the results contain much ‘garbage’.

With the growth of information on the internet on the one hand and increasing transparency requirements for the activities of private companies and government agencies on the other, companies spend significant funds to make themselves known to the world. Almost every company has an official website and social media accounts where news is duplicated (the first round of duplication). Furthermore, all significant market players present official press releases for the media, which also have websites and social media accounts. Thus, readers (internet users) are connected and repeatedly duplicate the publications of companies and mass media on their own social media accounts (see Figure 1). As a result, there are usually a dozen to a thousand identical publications for a single event. The number of publications in business media depends on the importance of the news, person, or event (for the region, country, or world) covered. The more significant the person or event is, the greater the duplication of the publications on it. Quantity does not translate into quality. The farther away from the original source a publication is, the more distorted the information. According to experts, everyone’s media activity does not increase but reduces the availability of information.



**Figure 1.** Simplified scheme of formation of industry-specific information flows on the internet.

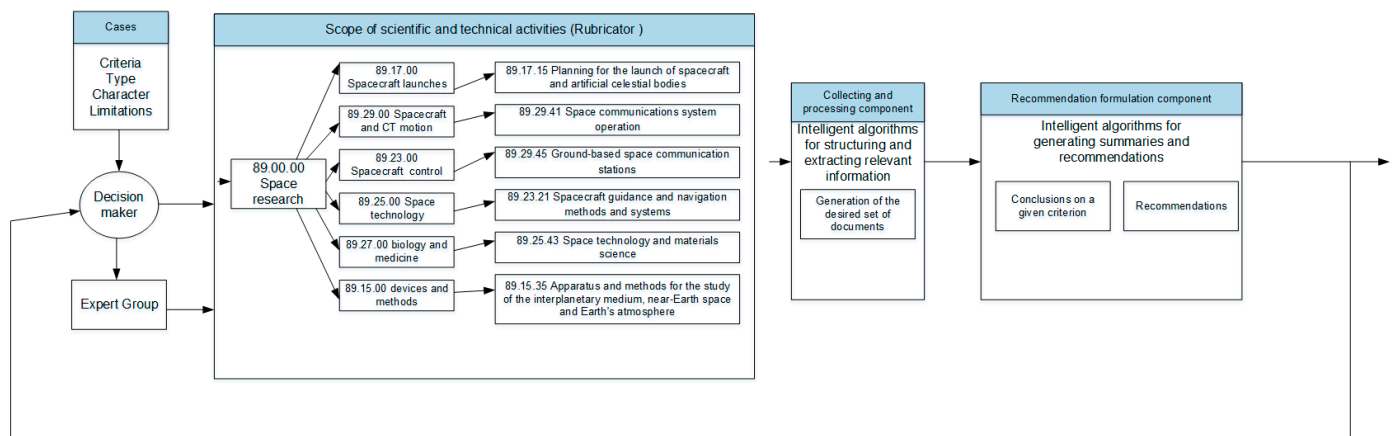
The global issue of information noise has been a subject of discourse for some time. Initially, there was optimism regarding the use of automation and AI to address this problem. However, an examination of the international literature revealed that even well-funded western AI developers are grappling with challenges akin to escalated information noise and subpar input data quality. As highlighted in the quote below [9], this phenomenon distorts the outcomes generated by AI systems.

... Data serves as the lifeblood for artificial intelligence models; it cannot merely serve as a means to an end in the modelling process. . . What is fed in determines what is produced—a longstanding principle of the modelling paradigm. . . In the era of contemporary AI and the concurrent deluge of data that machine learning models must contend with, deciphering flawed outcomes has become a more intricate task.

Our investigation revealed that the quantity of responses obtained is contingent upon multiple factors, including the specific wording of the query, the language used, the geographic region, and the temporal aspect of the request. The observed variations are conspicuous even within a single diurnal cycle, let alone over a span of a week or a month.

Semantic analysis is a method that has proven efficacious in orchestrating the influx of incoming data. To operationalise this approach, it is imperative to first delineate the pivotal concepts and nomenclature germane to the subject domain. This was conducted by interfacing classifiers and internationally recognised rubricators that have garnered consensus within the scientific and technical communities. The prescribed classifiers are used to scrutinise fresh publications within the aerospace sector, with the objective of elucidating salient concepts and terminologies and discerning scholarly trajectories to encapsulate the contemporary landscape of the industry. This not only befits the economy of time and effort but also furnishes a means of vigilantly monitoring consequential intelligence. Figure 2 provides a comprehensive schematic representation for elucidation.





**Figure 2.** Visual depiction of the flowchart delineating the process of semantic analysis predicated upon a universally acknowledged rubricator.

### 3.2.3. Data Collected from Public Sources

The data collected by the developed IDSS from open sources from 28 October to 5 November 2021 were news articles and posts on social media platforms (Vk.com, Instagram, Facebook, and Twitter) with publication dates from 1 January 2016 to 5 November 2021. There were 4538 articles in all. Table 2 shows the word and character statistics from the collected data.

**Table 2.** Data collected from public sources from 1 January 2016 to 5 November 2021.

Word Resource	Word	Symbols per Word	Symbols per Word
Facebook	80.00	634.04	8.50
Instagram	139.77	1126.81	8.03
Twitter	16.03	118.04	7.41
VK.com	26.67	219.79	8.69
News portals	230.10	2013.24	9.01

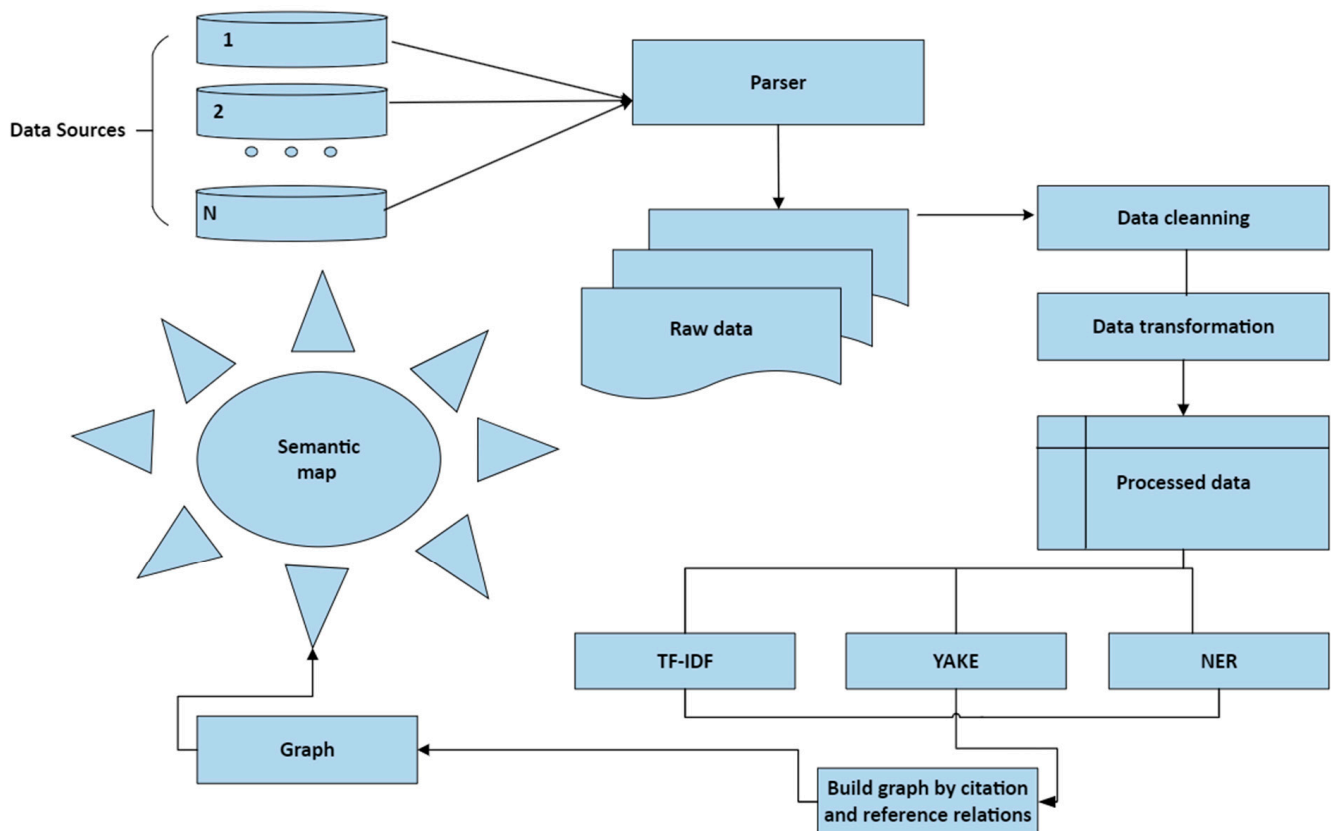
Articles were selected according to the list of keywords (space science, space industry, rocket, and satellite). Statistics on symbols and words in publications show that, on average, publications on news portals consist of 230 words and 2000 symbols. The second place was taken by Instagram posts, with 140 words and 126 characters per post. On Facebook, Twitter, and Vk, the words/symbols were 80/634, 16/118, and 27/220, respectively (see Table 2).

### 3.3. Preprocessing

Integral to the maturation of a robust decision-making apparatus, an array of numerical experiments was meticulously conducted, focussed on elucidating trends, clustering patterns, and semantic citation maps within the milieu of the space industry. The lynchpin of enquiry revolves around the developmental trajectory of space technologies, as elucidated within the scientific and technical literary stratum. This endeavour is imbued with the exploration of algorithmic paradigms for articulating trends, clustering patterns, and semantic citation maps, underpinned by graph-theoretical constructs and a coherent formal architecture inherent to scientific documents.

The efficacy of the algorithms was rigorously examined, manifesting affirmative outcomes in delineating discrete subdomains within the overarching subject area. This discernment of nuanced subtopics bears significant utility in the realm of scientific and technological monitoring of burgeoning space industries, alongside facilitating information retrieval, abstracting, and other language processing tasks [13–18].

Of note is the proposed algorithm for the instantiation of semantic citation maps, premised on graph-theoretical paradigms and the meticulous elucidation of formal scientific document structures. Figure 3 shows a schematic depiction of this algorithmic instantiation. A pivotal facet of this algorithm is its astute treatment of inter-document relationships, manifesting as citations, and its ability to unravel keywords that holistically encapsulate the semantic essence of each document. This algorithm was meticulously verified and validated, leveraging an expansive dataset culled from diverse online sources dedicated to the monitoring of publication activity [19–22]. Notably, the corpus comprised 5762 articles. Before data cleansing and subsequent transformation into a machine-readable format, an article–article adjacency matrix was realised, informed by reference linkages derived from the List of Sources Used section. The algorithm seamlessly incorporates the Term Frequency-Inverse Document Frequency (TF-IDF), Yet Another Keyword Extractor (YAKE), and Named-Entity Recognition (NER) techniques to ascertain common keywords, thereby facilitating the modelling of thematic linkages between scientific and technical publications [23–29].



**Figure 3.** Conceptual diagram of the algorithm for constructing a semantic citation map.

As a result of the aforementioned algorithmic transformations, the dataset metamorphoses into a structured schema featuring 26,066 rows, where key\_1 and key\_2 denote outgoing and incoming article IDs, respectively, underscored by their respective keywords and titles. Additional dimensions include abstracts, further cementing the comprehensiveness of the structured schema.

This article’s new GreedySummariser method outperforms traditional abstracting methods on test data on all synthetic metrics. It is based on calculating the importance of sentences through a common statistical measure, TF-IDF, which is computed as follows:

$$\begin{aligned} \text{tf}(t, d) &= \frac{t}{d} \\ \text{idf}(t) &= \log\left(\frac{N}{df+1}\right) \\ \text{tf-idf}(t, d) &= \text{tf}(t, d) \times \text{idf}(t) \end{aligned} \quad (1)$$

where  $t$  is the number of terms in the document,  $d$  is the paper’s length,  $N$  is the number of copies, and  $df$  is the number of documents containing a term.

The algorithm’s sequence is as follows:

Step 1. The full text of the article ( $T$ ) is divided into sentences ( $S$ ).

Step 2. For each sentence from  $S$ , the TF-IDF statistical metric is estimated for the whole document collection, and a matrix ( $Mx$ ) is generated based on this estimation.

Step 3. A loop with a stop condition is run until  $Mx$  is empty or the desired number of sentences is selected.

Step 4.  $Mx$  along axis (row) 0 is summarised, and the maximum value index, which is the index of the sentence containing the maximum TF-IDF value words from  $T$ , is obtained. The index is saved on the index list ( $IL$ ).

Step 5: The matrix is updated by removing the columns corresponding to non-zero values for the sentence with the maximum sum of words of TF-IDF value from  $T$ . If the stop condition is triggered, step 6 is proceeded to. Otherwise, step 4 is repeated.

Step 6.  $IL$  is updated with the best combination of sentences.

Step 7. The indices in  $IL$  are arranged in ascending order to restore the article’s original sentence sequence order.

Step 8. The summary is assembled by taking the sentences from  $T$  by the indices in the sorted  $IL$ .

The extractive approach is deemed easier as it achieves good grammar and accuracy levels by copying large chunks of text from a source document; however, the complex abilities essential for high-quality generalization, including paraphrasing, generalization, and incorporating real-world knowledge, are feasible only within an abstract framework, and despite the increased challenges, notable advancements have been made, particularly with recent developments in deep learning. Deep learning methods are the most sophisticated methods for the mentioned purpose, but classical ML algorithms are still used at present for many tasks, such as text generalisation. Such algorithms are faster and more computationally efficient, but deep learning methods produce more human-like summaries and try to provide more details in less space. Thus, based on the results of the numerical experiments conducted, GreedySummariser is the most efficient method for abstracting space text.

#### 4. Findings

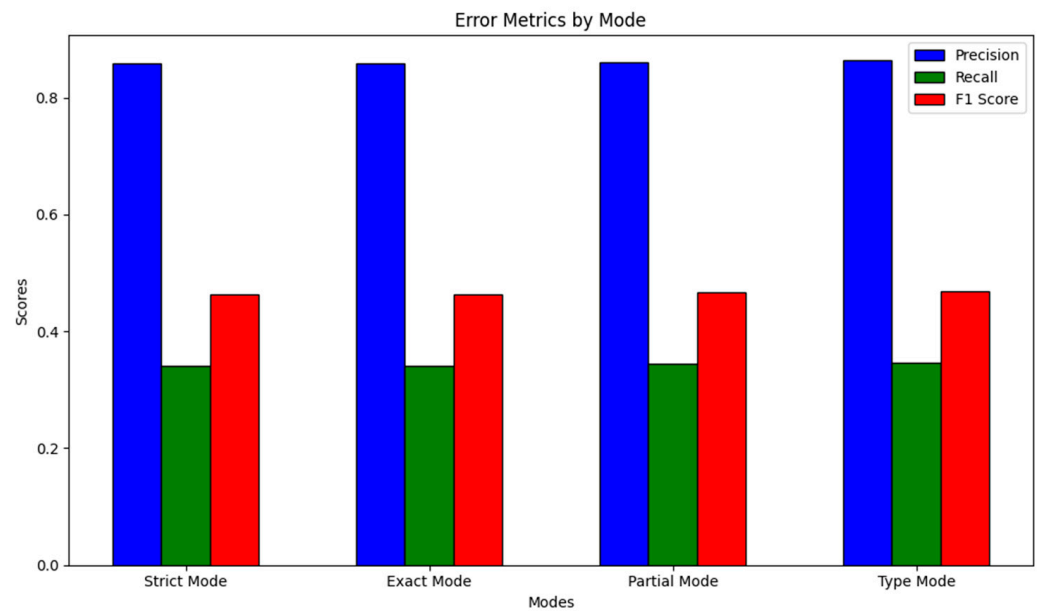
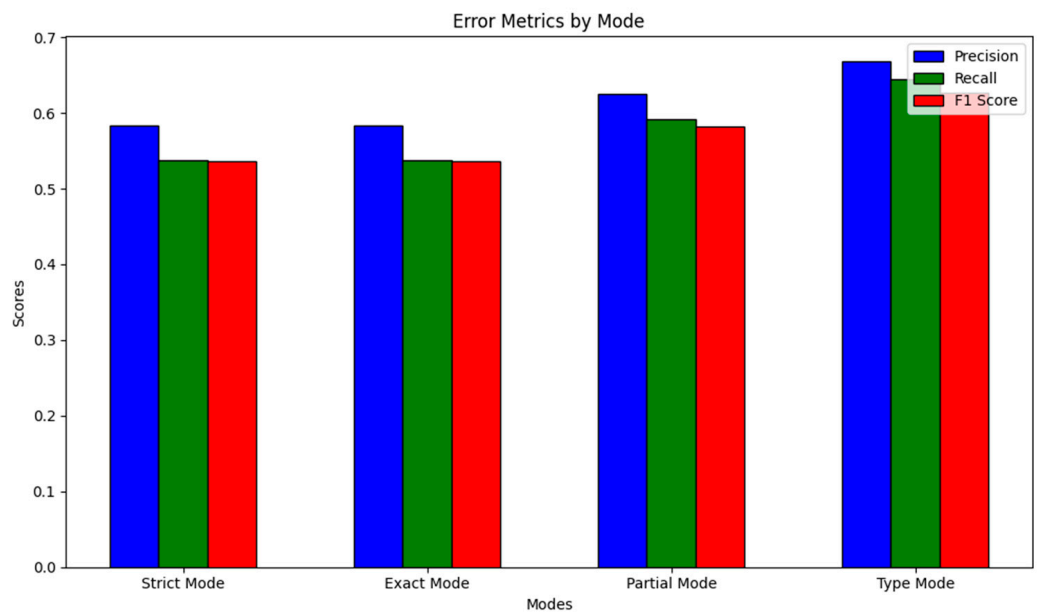
Tables 3–5 cogently present the outcomes of the rigorous algorithmic examinations in encapsulating the crux of the advancements and insights wrought by the numerical experimentation conducted in the present study. For enhanced comprehension, Figures 4–6 delineate the error metric across distinct modes.

**Table 3.** Results of selecting and classifying keywords by the ‘rocket’ class.

	Precision	Recall	F1 Score
Strict mode	0.8581	0.3413	0.4641
Exact mode	0.8581	0.3413	0.4641
Partial mode	0.8607	0.3437	0.4667
Type mode	0.8634	0.3461	0.4692

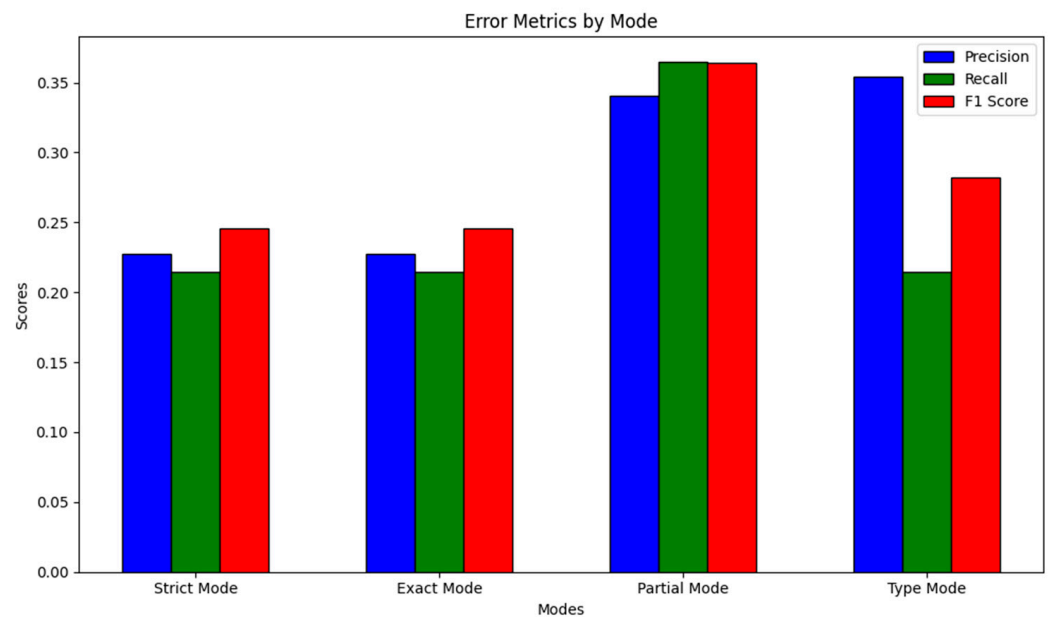
**Table 4.** Results of selecting and classifying keywords by the 'satellites' class.

	Precision	Recall	F1 Score
Strict mode	0.5830	0.5379	0.5369
Exact mode	0.5830	0.5379	0.5369
Partial mode	0.6254	0.5912	0.5816
Type mode	0.6678	0.6445	0.6263

**Figure 4.** Results of classifying keywords by the 'rocket' class.**Figure 5.** Results of classifying keywords by the 'satellites' class.

**Table 5.** Results of selecting and classifying keywords by the ‘technologies’ class.

	Precision	Recall	F1 Score
Strict mode	0.2272	0.2148	0.2457
Exact mode	0.2272	0.2148	0.2457
Partial mode	0.3406	0.3646	0.3641
Type mode	0.3540	0.2145	0.2825

**Figure 6.** Results of classifying keywords by the ‘technologies’ class.

Precision and recall have close values in the strict and exact modes, which may indicate homogeneity in the model, but the accuracy and completeness of the model in these modes are low.

Precision is higher and closer to recall in the partial mode, which may indicate a more balanced relationship between accuracy and completeness.

In the type mode, precision is higher than recall, and the difference between the two is significant. This may indicate a more focussed model approach to accuracy at the expense of completeness.

Clustering is a method of unsupervised ML that combines data points with similar characteristics. It is also a useful tool for finding trends in large datasets and for recognising complex relationships between data points based on keywords.

After analysing the results of keyword selection and classification for different classes, such as ‘rocket’, ‘satellites’ and ‘technologies’, we used multiple modes: the strict, exact, partial, and type modes. We evaluated the performance of the classification models using the precision, recall, and F1 score metrics.

The classification models demonstrated high precision, recall, and F1 scores across all modes for the ‘rocket’ class, with F1 scores ranging from 0.4641 to 0.4692. The ‘satellites’ class also exhibited a reasonable performance, particularly in the partial and type modes, with the F1 scores reaching 0.5816 and 0.6263, respectively. However, the ‘technologies’ class struggled to achieve high precision, recall, or F1 scores, with the highest F1 score of 0.3641 observed in the partial mode. It is noteworthy that clustering, an unsupervised ML method, has proven helpful in identifying trends and complex relationships within large datasets based on keywords.

Further research and model tuning may be necessary to enhance the performance of classification models, especially for classes with lower F1 scores. Nevertheless, the findings

of the present study provide valuable insights into the use of clustering techniques for keyword-based data analysis in the aerospace industry.

The aforementioned methods are compared in Table 6 as summarisation results on the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) metric.

**Table 6.** Comparative results of existing text summarisation methods on scientific and media data (the higher the value, the better).

Methods	BigPatent		ArXiv		PubMed		KZ Data	
	R-1	R-2	R-1	R-2	R-1	R-2	R-1	R-2
TextRank	35.99	11.14	-	-	-	-	31.9	6.7
LexRank	35.57	10.47	33.85	10.73	39.19	13.89	31.5	6.5
LSA	-	-	29.91	7.42	33.89	9.93	-	-
SumBasic	27.44	7.08	29.47	6.95	37.15	11.36	-	-

One of the most effective metrics for evaluating summarisation is ROUGE [30]. The basic idea behind this metric is to count the number of word and/or phrase matches (also known as  $n$ -grams) between a generated graded resume and an excellent human-generated gold standard. While there are many ways to measure similarity between reference and candidate resumes, the ROUGE metric remains the global standard in general text summarisation tasks.

There are variations in the ROUGE metric in the literature. The most common are ROUGE-N, ROUGE-L, ROUGE-W, and ROUGE-S, which are included in the publicly available Natural Language Toolkit (NLTK) natural language processing package. Formally, ROUGE-N is an  $n$ -gram recall between a candidate summary and a set of reference summaries. The authors have reported the stability and robustness of ROUGE on various sample sizes [30,31]. Nevertheless, achieving a high correlation with human judgment when summarising multiple documents, as ROUGE has achieved in single-document summarisation tasks, is still an open research topic.

Another approach to calculating the score is using traditional ML metrics, such as cosine distance or RMS error, over vectorised word representations.

This article discusses the concept and implementation of an IDSS for the aerospace technology sector. The IDSS utilises AI algorithms to extract, process, and analyse data from diverse sources to obtain insights critical for informed decision-making. It has various functionalities, ranging from data search and optimisation to modelling and forecasting, and is designed to undergird decision-making, problem-solving, and research initiatives within the aerospace industry, catering to engineers, technicians, and managerial cadres. It aims to optimise performance, curtail costs, and ensure safety and efficiency in aircraft design, flight operations, maintenance, security, and logistics. This article highlights the importance of effective information retrieval mechanisms, including online databases, search engines, and specialised journals and publications, to support the accelerating trajectory of space technologies.

## 5. Discussion

### 5.1. Theoretical Implications

Based on the results of our analysis, it is clear that classification models can be highly effective in identifying trends and patterns within large datasets based on keywords. The high performance observed in the ‘rocket’ and ‘satellites’ classes suggests that these models can be beneficial in the aerospace industry for identifying and tracking developments in these areas.

The concept of green innovation is significant in today’s world [32], where we face a pressing need to address climate change and reduce environmental impacts. Our analysis suggests that ML and AI can play vital roles in accelerating green innovation by enabling



researchers to identify emerging technologies and practices that promote sustainability and reduce environmental impacts.

Our findings suggest that classification models and clustering techniques can be particularly useful in identifying trends and patterns within large datasets related to green innovation. By analysing such datasets, researchers can gain valuable insights into market trends, competitor activities, and emerging technologies, which can inform investment, research and development, and strategic planning.

Furthermore, our findings suggest that clustering can be a valuable tool for unsupervised ML in the aforementioned context, enabling researchers to identify complex relationships and trends within datasets that may not immediately be apparent.

### *5.2. Managerial Implications*

For managers in the aerospace industry, the use of classification models and clustering techniques can provide valuable insights into market trends, competitor activity, and emerging technologies. By analysing large datasets based on keywords, managers can make informed decisions about investment, research and development, and strategic planning.

For managers working in the field of green innovation, the use of ML and AI can provide valuable insights into market trends, emerging technologies, and best practices. By leveraging these technologies, managers can make informed decisions about investment, research and development, and strategic planning, which can help promote sustainability and reduce environmental impacts. However, it is essential to note that the performance of ML and AI models can vary depending on the class being analysed. As such, it may be necessary to conduct further research and model tuning to enhance the accuracy and reliability of these techniques.

### *5.3. Limitations and Future Research Directions*

One limitation of our study is that it focussed solely on keyword-based data analysis. Future research may explore the use of other techniques, such as natural language processing and sentiment analysis, to gain a more nuanced understanding of trends and patterns within large datasets.

Further research may be necessary to evaluate the performance of ML and AI models across a wider range of datasets related to green innovation, which can help identify areas where these techniques are particularly effective and areas where they may require further development. By continuing to explore the potential of these technologies for promoting sustainability and reducing environmental impacts, we can create a better world for future generations.

Additionally, further research may be necessary to evaluate the performance of classification models and clustering techniques across a wider range of classes and datasets. This can help identify areas where these techniques are particularly effective and areas where they may require further development.

### *5.4. Ethical Implications of Artificial Intelligence Research*

In light of the growing concern about ethical considerations in AI technology, a dedicated section that delves into the ethical implications of the research topic was added to our study. AI has undoubtedly revolutionised industries, offering unprecedented capabilities in data analysis, automation, and decision-making. From healthcare to finance, AI-powered solutions have become an integral part of our lives. However, when it comes to space technologies, the implications of AI have become even more profound. The ability to process vast amounts of data and make split-second decisions is critical in space exploration, making AI a valuable tool for space agencies. The issue of bias in AI algorithms is at the forefront of ethical concerns about AI. Unintentional bias can perpetuate discrimination and social injustice. Discussing fairness-aware ML techniques and debiasing strategies is essential to addressing this issue and its consequences.

## 6. Conclusions

The article discussed the use of classification models and clustering techniques in the aerospace industry in relation to green innovation. The research findings highlight how these models can effectively identify patterns and trends in datasets using keywords. It suggests that these techniques can help track developments and promote practices thus accelerating innovation within the aerospace industry. These methods are also seen as tools for decision-making in areas such as investment, research, and strategic planning. However, the article acknowledges some limitations. It recognizes that focussing solely on keyword-based analysis may not provide an understanding of the data. Therefore, future research should explore techniques like natural language processing and sentiment analysis to gain a nuanced perspective. The article also emphasizes considerations in AI, particularly addressing biases in algorithms and stressing the importance of fairness ML techniques and strategies to eliminate bias. In conclusion, while this study highlights the applications of classification models and clustering techniques, it encourages research to improve their accuracy and dependability. Additionally, it underscores the need for considerations to mitigate biases within AI systems, promoting an inclusive integration of artificial intelligence across various industries.

**Author Contributions:** Conceptualisation, G.M. and Z.O.; data collection, A.A.S. and Z.I.; methods, G.M. and Z.I.; data analysis, A.A.S. and Z.O.; writing—original draft preparation, G.M. and Z.O.; writing—review and editing, all the authors; supervision, G.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Aerospace Committee of the Ministry of Digital Development, Innovations and Aerospace Industry of the Republic of Kazakhstan, grant number: BR11265420.

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

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are contained within the article.

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

## References

1. Zhao, P.; Gao, Y.; Sun, X. How does artificial intelligence affect green economic growth?—Evidence from China. *Sci. Total Environ.* **2022**, *834*, 155306. [CrossRef] [PubMed]
2. Zhou, H.; Yip, W.S.; Ren, J.; To, S. Topic discovery innovations for sustainable ultra-precision machining by social network analysis and machine learning approach. *Adv. Eng. Inform.* **2022**, *53*, 101715. [CrossRef]
3. The Benefits of Implementing Green Innovations in Business. Sustainable Business Toolkit. Available online: <https://ied.eu/blog/green-innovation-in-business> (accessed on 16 August 2021).
4. Why Is Sustainability in Business Important and How to Achieve It. Available online: <https://emeritus.org/blog/sustainability-why-is-sustainability-important/> (accessed on 9 December 2022).
5. Khurana, A.; Bhatnagar, V. Investigating Entropy for Extractive Document Summarization. *Expert Syst. Appl.* **2022**, *187*, 115820. [CrossRef]
6. Sharma, R.; Shaikh, A.A.; Bekoe, S. Ramasubramanian Information, communications and media technologies for sustainability: Constructing data-driven policy narratives. *Sustainability* **2021**, *13*, 2903. [CrossRef]
7. The Business Case for Sustainability. Harvard Business Review. Available online: <https://online.hbs.edu/blog/post/business-case-for-sustainability> (accessed on 16 August 2021).
8. Sánchez-Marrè, M. *Intelligent Decision Support Systems*; Springer International Publishing: Cham, Switzerland, 2022; pp. 77–116.
9. Bonvillian, W.B. DARPA and its ARPA-E and IARPA clones: A unique innovation organization model. *Ind. Corp. Chang.* **2018**, *27*, 897–914. [CrossRef]
10. Qasem, M.H.; Aljaidi, M.; Samara, G.; Alazaidah, R.; Alsarhan, A.; Alshammari, M. An Intelligent Decision Support System Based on Multi Agent Systems for Business Classification Problem. *Sustainability* **2023**, *15*, 10977. [CrossRef]
11. Hallstedt, S.I.; Bertoni, M.; Isaksson, O. Assessing sustainability and value of manufacturing processes: A case in the aerospace industry. *J. Clean. Prod.* **2015**, *108*, 169–182. [CrossRef]
12. Sanchez-Gomez, J.M.; Vega-Rodríguez, M.A.; Pérez, C.J. A multi-objective memetic algorithm for query-oriented text summarization: Medicine texts as a case study. *Expert Syst. Appl.* **2022**, *198*, 116769. [CrossRef]

13. Du, X.; Lu, Z.; Wu, D. An intelligent recognition model for dynamic air traffic decision-making. *Knowl.-Based Syst.* **2020**, *199*, 105274. [CrossRef]
14. Data Scientists: Bring the Narrative to the Forefront: [Electronic Resource]. 2021. Available online: <https://techcrunch.com/2021/04/16/data-scientists-bring-the-narrative-to-the-forefront/> (accessed on 16 April 2021).
15. June, R.D.A. Why the United States Needs a National Advanced Industry and Technology Agency [Electronic Resource]. 2021. Available online: <https://itif.org/publications/2021/06/17/why-united-states-needs-national-advanced-industry-and-technology-agency> (accessed on 17 June 2021).
16. Cozzens, S.; Gatchair, S.; Kang, J.; Kim, K.-S.; Lee, H.J.; Ordóñez, G.; Porter, A. Emerging technologies: Quantitative identification and measurement. *Technol. Anal. Strateg. Manag.* **2010**, *22*, 361–376. [CrossRef]
17. Takeda, Y.; Kajikawa, Y. Optics: A bibliometric approach to detect emerging research domains and intellectual bases. *Scientometrics* **2009**, *78*, 543–558. [CrossRef]
18. Small, H.; Boyack, K.W.; Klavans, R. Identifying emerging topics in science and technology. *Res. Policy* **2014**, *43*, 1450–1467. [CrossRef]
19. Lee, W.H. How to identify emerging research fields using scientometrics: An example in the field of Information Security. *Scientometrics* **2008**, *76*, 503–525. [CrossRef]
20. Serenko, A.; Bontis, N.; Booker, L.; Sadeddin, K.; Hardie, T. A scientometric analysis of knowledge management and intellectual capital academic literature (1994–2008). *J. Knowl. Manag.* **2010**, *14*, 3–23. [CrossRef]
21. Watts, R.J.; Porter, A.L. R and D cluster quality measures and technology maturity. *Technol. Forecast. Soc. Change* **2003**, *70*, 735–758. [CrossRef]
22. Jones, B.F.; Weinberg, B.A. Age dynamics in scientific creativity. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 18910–18914. [CrossRef] [PubMed]
23. Van Raan, A.F.J. On growth, ageing, and fractal differentiation of science. *Scientometrics* **2000**, *47*, 347–362. [CrossRef]
24. Jo, Y.; Lagoze, C.; Giles, C.L. Detecting research topics via the correlation between graphs and texts. *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.* **2007**, 370–379. [CrossRef]
25. Shibata, N.; Kajikawa, Y.; Takeda, Y.; Matsushima, K. Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation* **2008**, *28*, 758–775. [CrossRef]
26. Lee, J.; Lee, D. An improved cluster labeling method for support vector clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 461–464.
27. Lamirel, J.C.; Ta, A.P.; Attik, M. Novel labeling strategies for hierarchical representation of multidimensional data analysis results. In Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA), Innsbruck, Austria, 11–13 February 2008.
28. Osborne, F.; Scavo, G.; Motta, E. A Hybrid Semantic Approach to Building Dynamic Maps of Research Communities. In *Knowledge Engineering and Knowledge Management*; Janowicz, K., Schlobach, S., Lambrix, P., Hyvönen, E., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2014; Volume 8876. [CrossRef]
29. Van Lierde, H.; Chow, T.W. Query-oriented text summarization based on hypergraph transversals. *Inf. Process. Manag.* **2019**, *56*, 1317–1338. [CrossRef]
30. Du, S.; Xie, C. Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *J. Bus. Res.* **2021**, *129*, 961–974. [CrossRef]
31. Ouchchy, L.; Coin, A.; Dubljević, V. AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI Soc.* **2020**, *35*, 927–936. [CrossRef]
32. Ahmed, F.; Kousar, S.; Pervaiz, A.; Trinidad-Segovia, J.E.; Casado-Belmonte, M.D.P.; Ahmed, W. Role of green innovation, trade and energy to promote green economic growth: A case of South Asian Nations. *Environ. Sci. Pollut. Res.* **2022**, *29*, 6871–6885. [CrossRef]

**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.