sustainability

MDPI

Article

Identification of Hydrogen-Energy-Related Emerging Technologies Based on Text Mining

Yunlei Lin and Yuan Zhou *

School of Public Policy and Management, Tsinghua University, Beijing 100084, China; linyunlei@mail.tsinghua.edu.cn

* Correspondence: zhou_yuan@mail.tsinghua.edu.cn

Abstract: As a versatile energy carrier, hydrogen possesses tremendous potential to reduce greenhouse emissions and promote energy transition. Global interest in producing hydrogen from renewable energy sources and transporting, storing, and utilizing hydrogen is rising rapidly. However, the high costs of producing clean hydrogen and the uncertain application scenarios for hydrogen energy result in its relatively limited utilization worldwide. It is necessary to find new promising technological paths to drive the development of hydrogen energy. As part of technological innovation, emerging technologies have vital features such as prominent impact, novelty, relatively fast growth, etc. Identifying emerging hydrogen-energy-related technologies is important for discovering innovation opportunities during the energy transition. Existing research lacks analysis of the characteristics of emerging technologies. Thus, this paper proposes a method combining the latent Dirichlet allocation topic model and hydrogen-energy expert group decision-making. This is used to identify emerging hydrogen-related technology regarding two features of emerging technologies, novelty and prominent impact. After data processing, topic modeling, and analysis, the patent dataset was divided into twenty topics. Six emerging topics possess novelty and prominent impact among twenty topics. The results show that the current hotspots aim to promote the application of hydrogen energy by improving the performance of production catalysts, overcoming the wide power fluctuations and large-scale instability of renewable energy power generation, and developing advanced hydrogen safety technologies. This method efficiently identifies emerging technologies from patents and studies their development trends. It fills a gap in the research on emerging technologies in hydrogen-related energy. Research achievements could support the selection of technology pathways during the low-carbon energy transition.

Keywords: hydrogen energy; emerging technology; text mining; technology forecasting; reduce carbon

check for updates

Citation: Lin, Y.; Zhou, Y. Identification of Hydrogen-Energy-Related Emerging Technologies Based on Text Mining. *Sustainability* **2024**, *16*, 147. https://doi.org/10.3390/su16010147

Academic Editors: Zhaoyang Kong, Shaochao Ma and Yuhua Zheng

Received: 27 October 2023
Revised: 13 December 2023
Accepted: 20 December 2023
Published: 22 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

Hydrogen is the most abundant element in the universe [1,2], and accounts for approximately 90% of all atoms [3]. In recent years, green industries [4], especially hydrogen energy, have received widespread attention for their potential to combat climate change, support energy transition, and provide sustainable energy solutions [5–8]. Hydrogen can be derived from many resources and hence is named by different colors [9]. Fossil fuels produce grey hydrogen along with large amounts of carbon dioxide emissions [10]. Blue hydrogen is produced from natural gas, reducing greenhouse gas emissions via carbon capture, utilization, and storage (CCUS) [10]. It represents a transition from gray hydrogen to green hydrogen, which is produced from renewables [10]. Pink (or purple/red) hydrogen is produced by electrolysis through nuclear power [11].

In traditional industrial processes, hydrogen is mainly used as a raw chemical material [12]. With gradual breakthroughs in hydrogen fuel cells and water electrolysis, hydrogen has been regarded as a clean fuel used in transportation [13], construction [14], and industrial energy supply [15]. It is also regarded as a clean energy carrier with high

energy density to store and deliver energy [16,17]. Therefore, hydrogen energy is a critical pillar supporting the decarbonization of global energy systems [18], and many emerging technologies have sprung up in hydrogen production, storage, transportation, and utilization. According to Daniele Rotolo [19], emerging technology is characterized by radical novelty [20] (despite its abundance, hydrogen is a novelty in specific fields of application [21]), and it is also characterized by rapid growth that significantly impacts society [19]. Furthermore, emerging technology enables progressive development within an area and can lead to a paradigm shift in society [22]. It plays a vital role in industrial modernization, and researchers worldwide are developing new hydrogen technologies to preserve economic and energy security [23]. With the rapid development of technology, governments and enterprises have to decide which technology is the most appropriate for investment [24]. Predicting emerging technologies is important for governments and enterprises to seize opportunities in the new wave of technology [25,26].

The ability to predict emerging technology is a category of technology foresight. Technology foresight methods can be divided into exploratory technology forecasting methods, normative technology forecasting methods [25,27], and methods combining the first two methods (normative/exploratory). The exploratory forecasting method allows researchers to make realistic predictions of what will be the most advanced technologies in the near future [27], such as through bibliometrics [28] (including literature analysis and patent analysis [29]), the growth curve method, data mining (data mining tools include data mining [30], text mining, technology mining, and database tomography), analogies, system dynamics, technology forecasting using data envelopment analysis(TFDEA), etc. The normative forecasting method predicts what technology might be needed at a particular time [31]. It is a forecasting method that determines the organization's initial goal and then studies its feasibility, using methods such as relevance trees, the analytic hierarchy process (AHP), morphological analysis, backcasting, etc. The normative/exploratory approach combines the characteristics of the above two methods [31], as well as those of the Delphi method [32], the nominal group technique (NGT), scenario planning/writing, technology roadmapping [33], etc. The process of gaining technology foresight is complex and flexible. Different methods can be combined according to context to improve prediction accuracy [31]. With the big data industry rapidly developing [34,35] and the emergence of complex and social networks, a paradigm shift has occurred in data analysis [36]. Traditional manual coding possesses the disadvantages of efficiency [37], poor stability [38], and poor reproducibility [38], making it challenging to apply to large-scale text data analysis. Using data mining tools to explore texts such as documents and patents has become increasingly popular [39].

New patent information is the earliest signal of industrial innovation, and patent analysis has proven to be an effective method for identifying and studying emerging technology [40–42]. This study employed the latent Dirichlet allocation (LDA) topic model for text mining, adopted patent analysis methods, and combined technical experts' opinions to identify emerging technology in the field of hydrogen energy. Large quantities of patent data were analyzed to discover and define potential emerging technology topics that can provide valuable insights to support governments, research institutions, and industry stakeholders. As supporting and promoting innovation is challenging [43–45], the research framework and results can help decision-making, advance research and development, and accelerate the practical application of hydrogen-energy-related technology.

2. Methodology

The LDA model is a probabilistic model for generating a corpus proposed by David M. Blei and others in 2003 [46]. It is one of the most commonly used text mining techniques [47,48]. LDA is entirely unsupervised [49–51] and does not require the input of reference data [52]. In contrast, supervised methods require semantic annotation, which aims to assign semantics to the statement [53]. Semantic annotations often need manual input [53,54]. They are time-consuming, expensive [55], and unavailable in many application

scenarios. Furthermore, LDA provides an interpretable intermediate semantic representation of data in the form of topic graphs, thus leading to researchers making educated guesses and relying on initial information about the research domain [56]. Therefore, this paper uses a method based on the LDA model and integrates domain knowledge and expert group decision-making to conduct text mining of the hydrogen energy patent dataset, identify hydrogen energy technology topics, and conduct analyses of emerging technology. The research framework includes three modules: dataset construction, topic division, and technology foresight, as shwon in Figure 1.

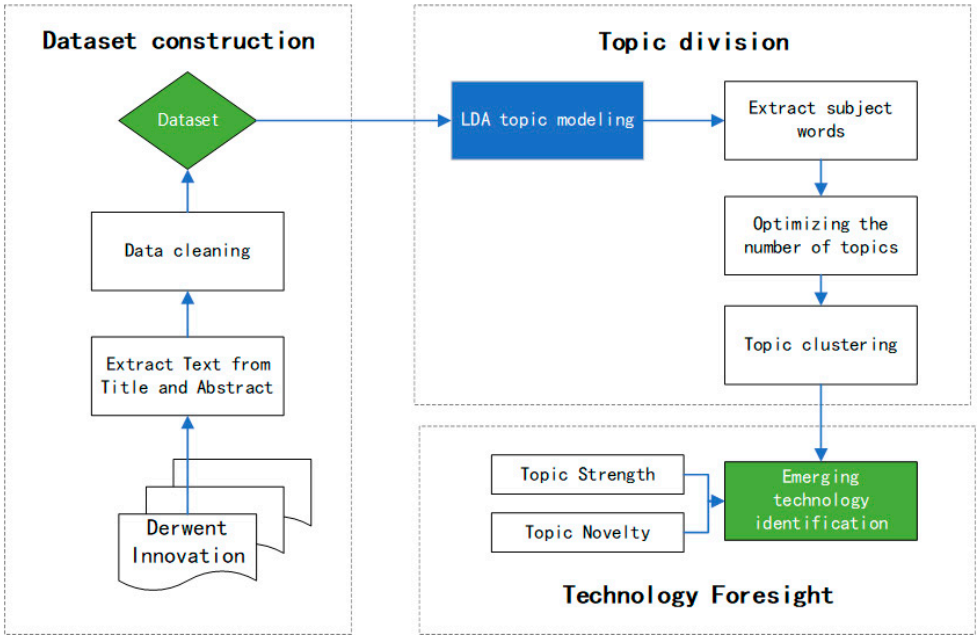

Figure 1. Overall process of the research framework.

2.1. Data Collection and Pre-Processing

Patents are a valuable source of knowledge about technological progress and innovation [57]. Thus, this study used patent texts to study and judge the emerging features of hydrogen-energy-related technological pathways. Derwent Innovation is a world-renowned comprehensive search platform dedicated to patent documents from more than 100 countries and regions worldwide, and the patent data used in this paper were derived from Derwent Innovation. The search query for patents was discussed and determined by the research group to embed experts' knowledge effectively and improve the rationality of the output. The expert group was composed of experts in the field of hydrogen energy from universities, state-owned enterprises, private companies, and research institutes in China.

Finally, the search query for patents was set as: "TAB=(hydrogen ADJ energy OR hydrogen ADJ generat* OR genarat* ADJ hydrogen OR generation ADJ of ADJ hydrogen OR hydrogen ADJ generation ADJ reactor OR hydrogen ADJ produc* OR produc* ADJ hydrogen OR hydrogen ADJ mak* OR mak* ADJ hydrogen OR production ADJ of ADJ hydrogen OR reform* ADJ to ADJ hydrogen OR reform* ADJ hydrogen OR hydrogen ADJ reform* OR hydrogen ADJ preparat* OR preparation ADJ of ADJ hydrogen OR prepar* ADJ hydrogen OR hydrogen ADJ manufacture OR manufacture ADJ of ADJ the ADJ hydrogen OR manufacture ADJ of ADJ hydrogen OR green ADJ hydrogen OR stor* ADJ hydrogen OR blue ADJ hydrogen OR hydrogen ADJ stor* OR stor* ADJ of ADJ hydrogen OR compressed ADJ hydrogen OR hydrogen ADJ transmission OR hydrogen ADJ transport* OR transport* ADJ of ADJ hydrogen OR transport* ADJ hydrogen OR hydrogen ADJ container OR hydrogen ADJ tank* OR hydrogen ADJ deliver* OR deliver* ADJ of ADJ hydrogen OR deliver* ADJ hydrogen OR hydrogen ADJ refuel* ADJ station OR hydrogen ADJ fuel* ADJ station OR hydrogen ADJ utilization OR hydrogen ADJ charging OR hydrogen ADJ fuel

ADJ cell* OR hydrogen ADJ burning OR hydrogen ADJ combustion OR hydrogen ADJ pip* OR hydrogen ADJ utilization OR hydrogen ADJ adoption) NOT TAB=(hydrogen ADJ peroxide) NOT TAB=(hydrogen ADJ sulfide) NOT TAB=(hydrogen ADJ chloride) NOT TAB=(hydrogen-rich ADJ water ADJ cup) NOT TAB=(hydrogen ADJ rich ADJ water ADJ cup) AND AD<=(20221231) AND AD>=(20080101)".

Valid and granted patents were selected in this paper. The collected patent information included application year and publication number. IPC number, title, abstract, inventor, DWPI patent family members, comprehensive influence, field influence, etc., were also included. The titles and abstract texts of the patents were used in the topic modeling stage.

The text had to be preprocessed before meaningful information was extracted from the collection of the title and abstract texts [58]. Text preprocessing converted the original text into precise semantic units and is applied to unstructured and structured documents. The specific steps used were text cleaning, tokenization, filtering, stemming, part-of-speech restoration, language processing, etc. [59,60]. This process was iteratively repeated through topic modeling and technology identification, ensuring that only terms that could be grouped into meaningful topics were retained.

2.2. Topic Modeling

The latent polynomial variables in the LDA model are called topics [61,62]. The basic idea is to treat the document as a random mixture of latent topics, and the word distribution determines the characteristics of each topic [63,64]. The LDA model contains a three-level Bayesian structure that includes documents, topics, and words [65]. This structure describes the topic probability distributions of documents and then considers the topic probability distributions of words to perform topic clustering or text classification based on the topic distribution [64]. The LDA model believes that a mixture distribution of topics generates each document. Every word in each document belongs to a topic with a certain probability. The LDA model allows the generated document to contain multiple topics, and each word in the document is generated by one of the topics [46,64]. The probability model is shown in Figure 2. LDA assumes that the prior distribution of the document topic is the Dirichlet distribution. The Gibbs sampling algorithm is used to iterate M times to calculate the topic distribution of each document and the distribution of words in each topic.

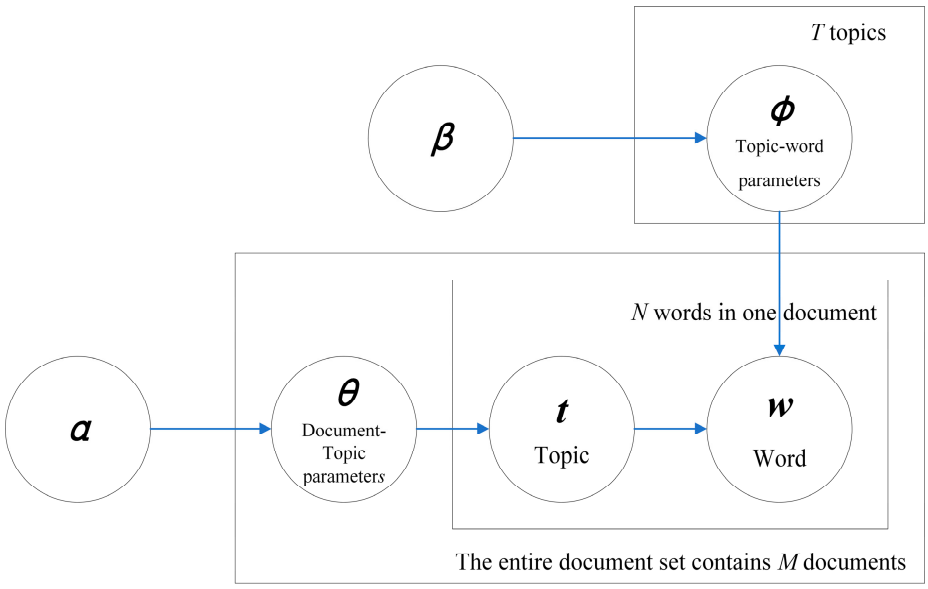

Figure 2. Schematic diagram of the latent Dirichlet allocation (LDA) model [46], where α and β are the parameters of the Dirichlet priors, θ is the document–topic generating variable, Φ is the topic–word generating variable, t is the topic assignment for the word, and w is the word. M represents the number of documents and T represents the number of topics.

The probability distribution of each word in the document is defined as follows [46]:

$$p(w|\theta,\beta) = \sum_{j=1}^{T} p(w|t,\beta) \times p(t|\theta) \quad (1)$$

where $\theta \sim \mathrm{Dir}(\alpha)$.

The number of topics (k), the number of iterations, α, and β are important parameters in constructing of the LDA model. The number of topics depends on the corpus's characteristics and size. The LDA model itself cannot determine the optimal number of topics [46]. The perplexity index can be used to evaluate the LDA model [66,67]. The perplexity index is the uncertainty regarding whether a document belongs to a specific topic [68]. This paper used the perplexity method to calculate the perplexity under different k values and combined it with matplotlib library in Python for visual display. The lower the perplexity, the better the clustering effect [69]. However, to avoid overfitting, k was selected using multiple experiments and combined with topic interpretability.

In this study, α = 1 and β = 0.01. In order to ensure the model effect, a large number of iterations was set, 1000. pyLDAvis was used for visual analysis [70].

2.3. Emerging Technology Identification

Novelty and prominent impact are essential for measuring emerging technology [19]. This paper uses topic strength and novelty indicators to determine whether each emerging technology has significant impact or novelty characteristics, respectively.

Topic strength shows the influence of the topic. It assesses whether a scientific topic is a research hotspot [71]. To provide a quantitative analysis indicator that can determine whether an emerging technology has a prominent impact, this paper applies the sum of the weights of the research topic in all documents to express the following equation [72]:

$$S_{t,j=} \frac{\sum_{i=1}^{M_t} \theta_j^i}{M_t} \quad (2)$$

where $S_{t,j}$ is the topic intensity of the jth topic in period t (the time unit is measured in years, for example, 2008), M_t is the total number of documents in period t, and θ_j^i is the probability of the ith document being assigned to the jth topic. As shown in Equation (1), the greater the topic intensity value, the more significant the proportion of the jth topic in the document set, and the more scholars pay attention to it. On this basis, a ThemeRiver chart of technology topics and a diagram of the evolutionary trends in technology topic strength were used to analyze the technical topic list further.

Emerging technology is radically novel [19] related to time. In this paper, topic novelty is defined as the index used to consider this evolution in the characteristics of technological development over time. The research cycle is divided into three stages, which measure whether the scientific research topic is novel. If a research topic accounts for less than 1/3 of the weight in the early stage (first stage), it is novel, and the following expression is used:

$$\theta_j^{(1)} \leq \frac{1}{3}\theta_j \quad (3)$$

where $\theta_j^{(1)}$ is the sum of the topic intensity values for the jth research topic in the first stage and θ_j is the sum of the topic intensity values for the jth research topic in the whole cycle.

3. Results

3.1. Hydrogen Energy Technology Topic Classification Based on Patent Analysis

After performing patent searching via Derwent Innovation, 37,261 records and 27,907 DWPI patent families dating from 1 January 2008 to 31 December 2022 were obtained. The DWPI

patent family has a strict definition in which one invention corresponds to a DWPI record and each DWPI patent family member is identical regarding its technical content. Hydrogen is widely used in human life. In the chemical industry, hydrogen has always been one of the primary raw materials used for synthesizing ammonia and methanol [73]. It is usually used in the refining of naphtha, gas, fuel, and petroleum via hydrogenation [12]. In the electronics industry, it is often used as a protective gas [74]. Hydrogen is fresh and environmentally friendly, making it suitable for use in the new field of green energy production [16,75]. This paper does not consider traditional hydrogen applications to ensure the study's consistency and reliability. After removing irrelevant records, 9611 DWPI patent families were obtained. As shown in Figure 3, the number of granted patents in hydrogen-energy-related regions has increased yearly since 2009 according to application year. Since there is a specific period between patent application and formal authorization, the number of granted patents for 2022 is relatively small. Research and the practical use of hydrogen-related energy is becoming increasingly popular.

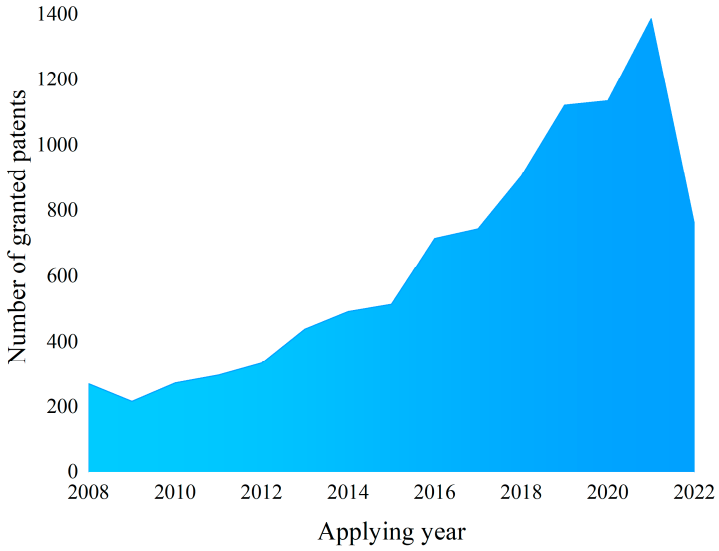

Figure 3. Numbers of granted patents for hydrogen-related technology, where the *x*-axis represents application year.

When cleaning the patent collection, it was considered that some words could be more descriptive of the topic in actual usage scenarios, and it is often necessary to associate the words to understand the topic's broader meaning. Therefore, a list of binary and multivariate phrases commonly used in the field of hydrogen energy (such as fuel cell, and generation of hydrogen) was created by the expert group to form a bag-of-word model containing binary and multivariate phrases. The NLTK English stopword list was used, and standard stopwords (e.g., the) and context-specific words (e.g., comprising) that frequently appeared in patent texts were removed, as were words that frequently appeared but did not have a representative meaning (e.g., hydrogen); after multiple modeling and iteration stages were conducted, a special deactivation word list for the hydrogen energy field was established. Following text preprocessing, each patent's title and abstract text were converted into a list of several words, which was used as the input of the LDA model. When carrying out topic modeling, the number of topics selected was 1–50. After 1000 iterations, it was concluded that the perplexity was relatively stable when the number of topics was 3–11 or 15–24. On this basis, the expert group had a meeting to consider the calculation results for different numbers of topics, believing that when the number of topics was 20, the technical topic clustering results were most consistent with the actual situation. The intertropical distance map [76,77] of the 20 topics is shown in Figure 4. The distances between different bubbles (they represent topics) reflect how similar the topics are to each other, and the overlap indicates the overlap between topics [78]. As shown in Figure 4, the

distribution of the 20 topics is relatively even, and the number of documents included in each topic is pretty reasonable. At the same time, there is an inevitable overlap between the topics.

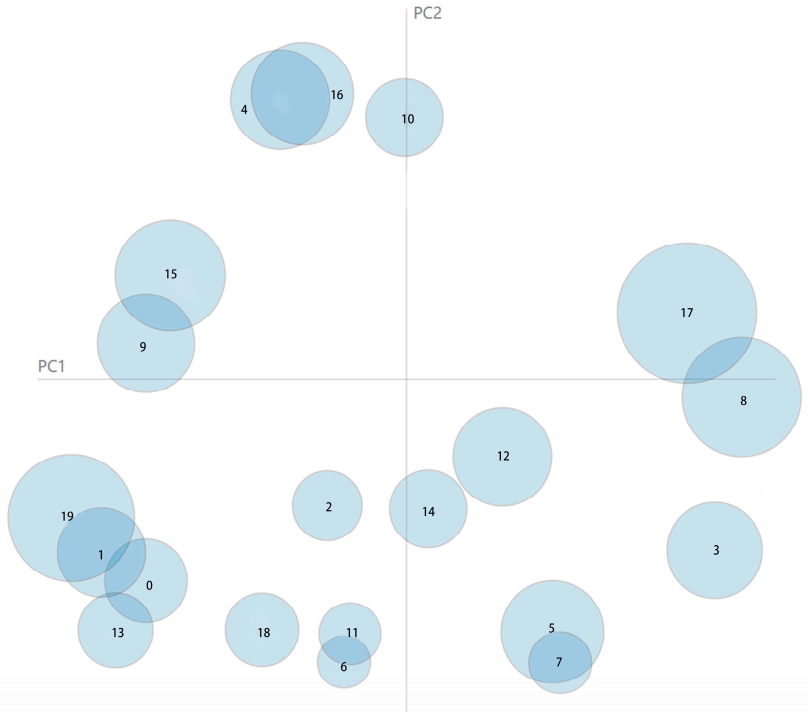

Figure 4. Inter-topic distance map when k = 20.

Given the number of topics, 20, the model extracted 100 keywords that best represented each topic. Each keyword possessed a probability that the topic belonged to this keyword. The top 20 keywords, with probability values, were selected to represent the topic. We defined every topic depending on the technical category, then a list of 20 technical topics and keywords in the field of hydrogen energy was obtained, as shown in Table 1.

Table 1. Hydrogen energy patent subject classification results.

Topic	Word (Probability, Top 20)
Topic 0 Hydrogen energy devices, systems, and detection methods	pipe, device, test, cylinder, sealing, groove, magnetic, sleeve, housing, winding, explosion, measuring, cylindrical, wire, feeding, connection, fire, screw, piston, and annular
Topic 1 Hydrogen fuel cells and applications	fuel cell, battery, power, fuel, hydrogen fuel cell, stack, electric, hybrid, circuit, air, galvanic, electrical, converter, starting, motor, ship, aquatic, capacitor, portable, and proton exchange membrane
Topic 2 Hydrogen storage containers	manufacture, hydrogen gas, container, steel, high pressure, fiber, device, metallic, material, hydrogen storage container, forming, hydrogen containing gas, pressure, produce, Si, permeation, stainless, hydrogen storage bottle, optical, and resistant
Topic 3 Hydrogen storage materials	metal, material, oxide, organic, film, compound, coat, Pd, P, nanometer, ceramic, preparation method, prospect, Zr, B, Mn, sintering, inorganic, forming, and solvent
Topic 4 Reforming hydrogen production	reform, steam, tube, methanol, water, hydrogen generation, conversion, purify, device, reformer, gas, reactor, vapour, hydrocarbon, raw material, vapor, burner, heat exchanger, combustion, and membrane
Topic 5 Hydrogen storage, alloy, etc.	alloy, storage, Mg, Ni, Al, hydrogen storage material, hydride, preparation method, battery, material, earth, metal, cadmium, ball milling, Li, discharge, releasing, La, Fe, and Mn

Table 1. *Cont.*

Topic	Word (Probability, Top 20)
Topic 6 Hydrogen-fuel-cell-related devices	vessel, vacuum, hydrogen generator, nozzle, bipolar, machine, spraying, terminal, pole, beam, belt, cabin, ethylene, rotor, width, marine, gear, activating, injector, and spray
Topic 7 Organic matter and composite materials used in hydrogen storage, etc.	composite material, Ti, cobalt, fibre, polymer, dioxide, hollow, salt, preparation method, resin, carbonate, glass, supercritical, soaking, poly, mesh, hydrogenase, trioxide, curing, and binder
Topic 8 Hydrogen production catalysts	catalyst, Ni, carrier, hydrogen generation, preparation method, Mo, metal, Pt, Fe, Cu, oxide, noble, Na, disulfide, hydroxide, borohydride, mesoporous, oxidation, ethanol, and Al
Topic 9 Hydrogen production from wind and solar and energy storage systems	power, electric, device, storage, generation, generate, hydrogen generation, wind, generator, water electrolysis, photovoltaic, electrolytic cell, control method, hydrogen production device, conversion, thermal, utilization, turbine, comprehensive, and store
Topic 10 Hydrogen production from biomass	reactor, hydrogen generation, methane, oil, biomass, gasification, hydrocarbon, generate, pyrolysis, produce, fermentation, biological, raw material, anaerobic, regeneration, bio, diesel, cl, conversion, and production of hydrogen
Topic 11 Hydrogen production units	cathode, anode, graphene, plasma, alkali, seawater, aqueous, sulfur, alumina, sulphur, dissolved, electrolyte, discharge, iodine, volatile, conducting, isotope, net, cation, and concentrated
Topic 12 Hydrogen production by water electrolysis	electrode, membrane, water electrolysis, electro, electrochemical, ion, electrolyte, tio2, alkaline, conductive, foam, electrolytic cell, electrolytic, ruthenium, electron, proton, bismuth, film, diffusion, and coat
Topic 13 Hydrogen fuel cells, vehicle equipment, and safety monitoring and control methods	vehicle, detect, sensor, ammonia, charging, hydrogenation, station, signal, controller, hydrogen fuel cell, device, hydrogen storage system, monitoring, automobile, mounted, motor, hydrogen leakage, hydrogen charging, control system, and transportation
Topic 14 Hydrogen production materials and methods and hydrogen purification	carbon, nitrogen, nitride graphite, adsorption, calcium, heterojunction, pressure swing adsorption, dry, microorganism, adsorbent, sieve, slag, splitting, oven, molten, raw material, sewage, methanation, and desorption
Topic 15 Preparation of hydrogen, synthesis, gas, hydrogen purification, and hydrogen-doped combustion	gas, fuel, carbon dioxide, device, combustion, generate, engine, stream, coal, synthesis, produce separating, purity, carbon monoxide, exhaust, mixing, purify, synthetic, recovery, and emission
Topic 16 Hydrogen production units	water, hydrogen generation, generate, device, electrolysis, solar, cell, produce, hydrogen production system, vanadium, bacteria, smelting, melting, sea, electrolyzer, electrocatalyst, generation, acetic, induction, and capable
Topic 17 Materials for photocatalytic hydrogen production	nano, preparation method, catalyst, material, photo, light, water, hydrogen generation, photocatalyst, decomposition, nanometre, zinc, photocatalytic, catalysis, carbide, hydrothermal, sulphide, solvent, Si, and mixing
Topic 18 Liquid hydrogen storage and transportation methods and devices	liquid, liquid hydrogen, liquefied, storage, heat exchanger, tank, filter, mos2, slurry, store, gaseous, hydrogen storage tank, nanotube, separator, dehydrogenation, lng, insulating, cryogenic, filtering, and cooled
Topic 19 Monitoring and safety control systems for hydrogen production, storage, and transportation equipment	pressure, tank, device, valve, air, pipeline, storage, pump, high pressure, filling, gas, compressor, pipe, discharging, hydrogen storage tank, bottle, discharge, store, exhaust, and conveying

Topic 0 includes devices, systems, and performance testing methods for hydrogen storage, hydrogen transportation, and hydrogen utilization equipment. Topic 1 includes hydrogen-fuel-cell systems, as well as systems and control methods for hydrogen-fuel-cell vehicles, drones, ships, and other equipment. Topic 2 involves hydrogen storage containers for hydrogen production. Topic 3 includes metal–organic framework materials that can be used as catalysts for photocatalytic-water-splitting hydrogen production, metal hydrogen absorption materials, metal-containing hydrogen evolution cathode materials, and hydrogen-fuel-cell anode materials, as well as ceramic-based hydrogen storage materials, etc. Topic 4 includes methods, devices, and systems for hydrogen production via the reforming of alcohols, natural gas, hydrocarbons, etc. Topic 5 includes hydrogen storage alloys, metal-based hydrogen storage materials, and hydride-containing hydrogen storage materials. Topic 6 includes hydrogen-fuel-cell systems and devices, as well as hydrogen-fuel-cell vehicle devices and systems. Topic 7 includes organic matter and composite materials for hydrogen storage and hydrogen production. Topic 8 includes ammonia decomposition hydrogen production, reforming hydrogen production, photocatalysis catalysts for hydrogen production, and water electrolysis. Topic 9 includes methods, devices, and systems for hydrogen production from wind and solar and energy storage systems. Topic 10 includes methods and apparatus for hydrogen production from biomass. Topic 11 includes hydrogen production equipment. Topic 12 includes materials, methods, and devices for hydrogen production by water electrolysis. Topic 13 includes hydrogen-fuel-cell vehicle equipment and safety monitoring and control methods. Topic 14 includes hydrogen production materials, methods, and purification. Topic 15 includes preparing hydrogen synthesis gas, hydrogen purification, and hydrogen-doped combustion. Topic 16 mainly involves hydrogen production devices. Topic 17 includes catalysts for photocatalytic hydrogen production. Topic 18 includes hydrogen liquefaction methods, systems, and devices, as well as liquid hydrogen storage and transportation methods, devices, and systems. Topic 19 includes hydrogen storage and transportation methods, as well as monitoring and safety control systems for hydrogenation equipment.

3.2. Identification of Emerging Hydrogen Energy Technologies

3.2.1. Topic Strength

Based on the lengths of the application periods for hydrogen energy patents, starting from 2008, each year was taken as a time slice, and the sum of the intensity weights of all patents belonging to different topics in each year was calculated. According to the "document-topic" distribution, the probability that each document belongs to each topic is obtained, and the topic strength is then calculated. For the period extending from 2008 to 2022, the sums of the topic strengths of the 20 topics are sorted from strong to weak. The results are shown in Figure 4. Since 2008, Topic 17 has been the most robust topic, and the top ten topics include Topic 19, Topic 15, Topic 8, Topic 5, Topic 16, Topic 4, Topic 3, Topic 12, and Topic 9. These technical topics have received widespread attention from scholars and are of high research interest.

Figure 5 reflects the overall influence of the hydrogen energy technology topics. The greater the power of the topic, the greater the research interest and the more attention it receives. Figure 6 describes the changing trends in the strengths of technical topics over time. There were also correlations between the 20 technical topics. For example, Topic 0 and Topic 19 both belong to the category of hydrogen safety. Therefore, the 20 technical topics were classified and compared according to their specialized technical fields.

Figure 7a shows the evolutionary trends in patent technology topic strengths in hydrogen production methods. As the figure shows, the topic strength values for hydrogen production using wind and solar and energy storage systems generally went upward. From 2017, the upward slope was the largest, indicating that hydrogen production using wind and solar technologies has recently received widespread attention, with research interest rapidly increasing. At the same time, the topic strengths for materials, methods, and devices for hydrogen production from water electrolysis continued to grow from 2014

to 2018 before decreasing slightly, and, overall, remained relatively stable. However, the topic strengths for technologies such as reforming hydrogen production, biomass hydrogen production, hydrogen synthesis gas preparation, and hydrogen purification generally showed a downward trend.

Figure 7b shows the evolutionary trends in patent technology topic strengths in hydrogen production materials and devices. Research interest in photocatalytic hydrogen production materials generally increased yearly from 2008 until 2018, after which point it declined. The topic strength for catalysts for hydrogen production increased slightly from 2014 to 2017 and then slowly decreased. The topic strengths for hydrogen production units and other topics were generally stable or somewhat fluctuating.

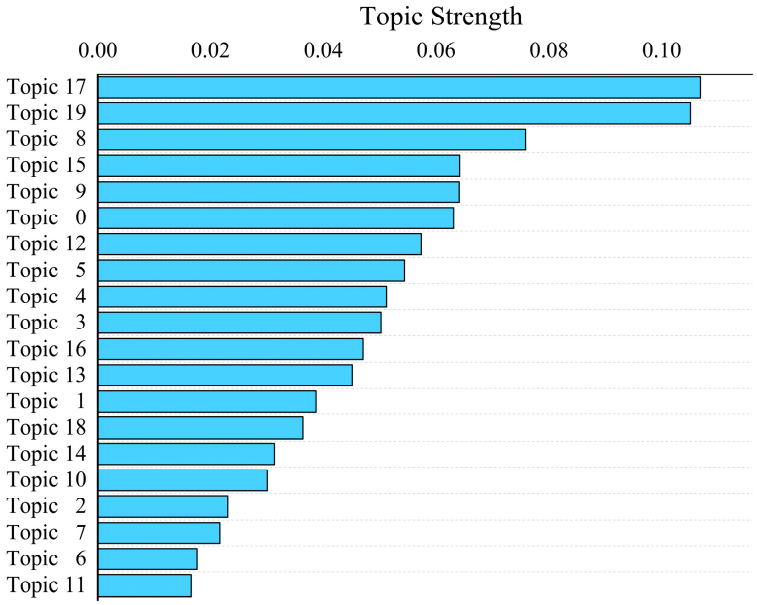

Figure 5. Overall ranking of the influence levels of 20 topics.

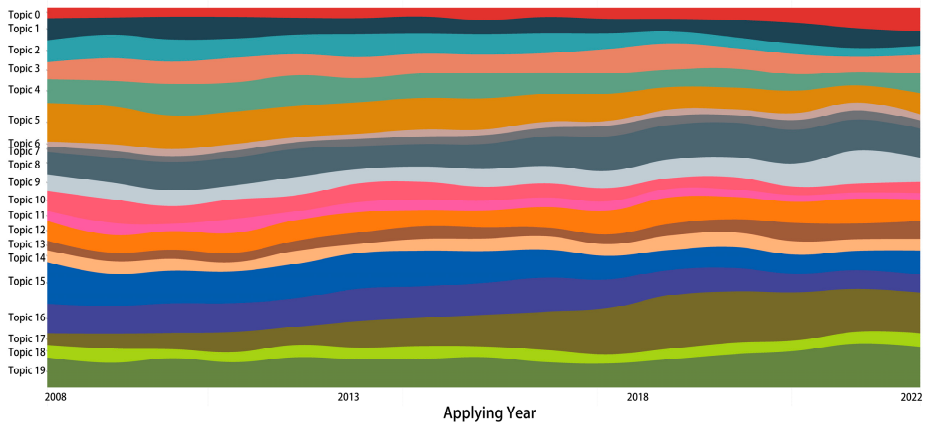

Figure 6. ThemeRiver for 20 hydrogen-energy-related technology topics, depicting topic strength changes from 2008 to 2022.

Figure 7c shows the evolutionary trends for the topic strengths for hydrogen storage and transportation patent technologies. The topic strength for hydrogen transportation and hydrogenation equipment monitoring and safety control systems fluctuated slightly from 2008 to 2017 before beginning to increase rapidly in 2017. The topic strength for organic matter and composite materials applied to hydrogen storage increased somewhat from 2008 to 2015, after which point the performance stabilized. The topic strength for hydrogen storage containers and hydrogen storage alloys was initially relatively high. However,

there was an obvious inflection point around 2016, and the topic strength dropped. As technology matures, research interest decreases. Figure 7d shows the evolutionary trends for the patent technology topic strength for hydrogen fuel cells and other hydrogen-energy-related equipment and systems. Figure 7 shows that the overall focus on safety monitoring methods and systems for hydrogen fuel cells and other hydrogen energy equipment has increased.

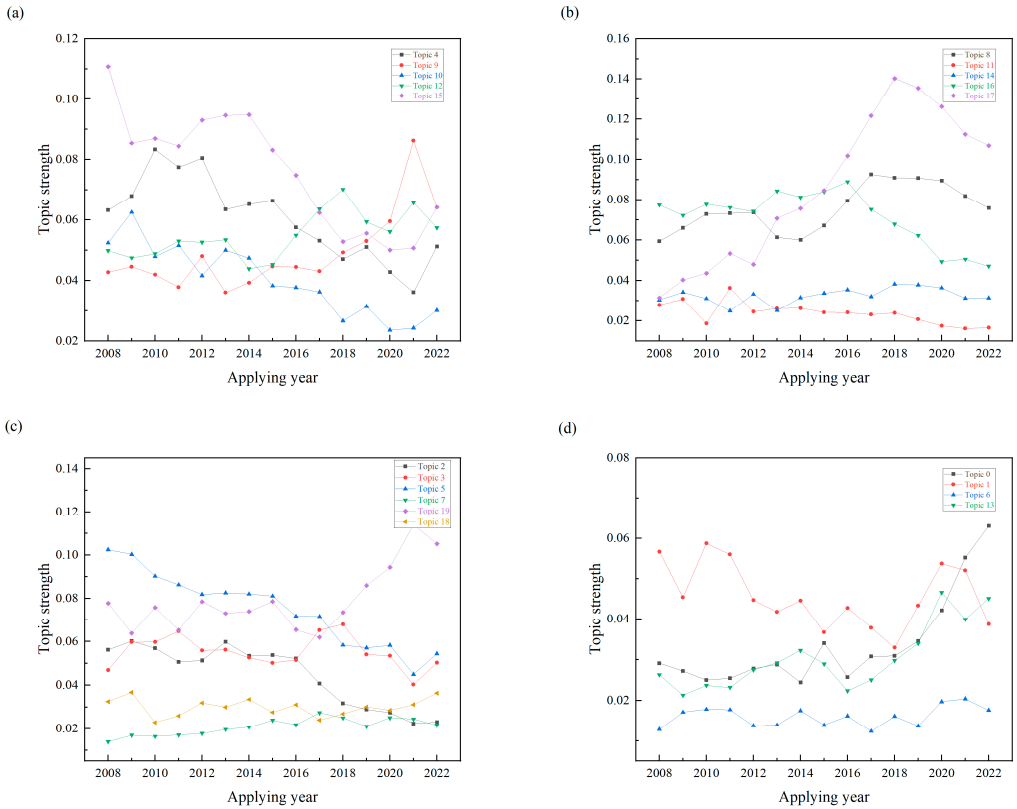

Figure 7. Evolutionary trends in topic strength in different technical fields: (**a**) hydrogen production methods; (**b**) hydrogen production materials and devices; (**c**) hydrogen storage and transportation; (**d**) hydrogen fuel cells and other hydrogen-energy-related equipment and systems.

Therefore, of the top ten technical topics with the highest topic intensity in recent years, Topic 15, Topic 5, Topic 16, and Topic 4 show an apparent downward trend. Although the overall topic intensity is high, their influence has gradually weakened. Topic 0 and Topic 13 are similar to the technical Topic 19 and belong to hydrogen-safety-related categories. Their topic intensity has rapidly increased since 2016, reaching sixth (Topic 0) and 12th (Topic 13) place in 2022. Considering everything, it is believed that Topic 0, Topic 3, Topic 8, Topic 9, Topic 12, Topic 15, Topic 17, and Topic 19 significantly impact the field of hydrogen energy.

3.2.2. Topic Novelty

According to this paper's definition of topic novelty, the entire period of study lasted from 2008 to 2022, divided into three stages. Each stage spanned five years. The first stage is from 2008 to 2012. The first stage's topic intensity was registered as Sum', and the topic intensities from 2008 to 2022 were recorded as Sum. Then, the weight of Sum' to Sum was calculated, as shown in Table 2. If the weight of a specific research topic in the first stage was less than 33.3%, the topic was considered novel. From this method, it was concluded that Topic 0, Topic 6, Topic 7, Topic 8, Topic 9, Topic 12, Topic 13, Topic 14, Topic 17, and Topic 19 were novel.

Table 2. Novelty degree of 20 topics.

Topic Strength	Topic 0	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9
Sum	0.504482	0.686213	0.668508	0.828601	0.906131	1.122623	0.240857	0.312875	1.134842	0.734414
Sum′	0.134586	0.261597	0.275174	0.286987	0.372306	0.460923	0.07934	0.082056	0.345163	0.214976
Sum′/Sum (%)	26.68	38.12	41.16	34.64	41.09	41.06	32.94	26.23	30.42	29.27
Novelty	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes
Topic strength	**Topic 10**	**Topic 11**	**Topic 12**	**Topic 13**	**Topic 14**	**Topic 15**	**Topic 16**	**Topic 17**	**Topic 18**	**Topic 19**
Sum	0.601183	0.356078	0.821414	0.45533	0.486081	1.143459	1.068837	1.292599	0.449126	1.186347
Sum′	0.255891	0.137928	0.251623	0.12203	0.15362	0.460563	0.377857	0.216521	0.150083	0.360778
Sum′/Sum (%)	42.56	38.74	30.63	26.80	31.60	40.28	35.35	16.75	33.42	30.41
Novelty	No	No	Yes	Yes	Yes	No	No	Yes	No	Yes

The calculation results were listed in a questionnaire for each expert individually. Then, in discussion through meetings of the expert group, it was determined that the emerging technology topics with both prominent impact and novelty were Topic 0 (hydrogen energy devices, systems, and detection methods), Topic 8 (hydrogen production catalysts), Topic 9 (hydrogen production from wind and solar and energy storage systems), Topic 12 (materials, methods, and devices for hydrogen production by water electrolysis), Topic 17 (materials for photocatalytic hydrogen production), and Topic 19 (hydrogen storage, monitoring, and safety control systems for hydrogen transportation and hydrogenation equipment). Hydrogen energy covers a wide range of technological topics and cross-integrates with other technological fields [79], increasing the difficulty of identifying emerging technology. Combining unsupervised topic models and expert judgment can, on the one hand, discover latent or abstract topics appearing in document collections and, on the other hand, improve the credibility of technical identification. Manuel W. Bickel [80] used the LDA model to research Scopus-indexed abstracts in the field of sustainable energy, and the 15 hot topics with the most vital positive trends were obtained; those related to hydrogen energy included fuel cells, photocatalytic hydrogen production, and chemical catalysis. There were similar views to this paper regarding research interests and development trends.

4. Discussion

This paper selected six emerging hydrogen technology topics for discussion, which involved the issues of hydrogen production, storage, and transportation.

Hydrogen production technology can be divided into four categories based on different raw materials: hydrogen production from fossil energy reforming, hydrogen production from industrial byproducts, hydrogen production from biomass, and hydrogen production from water electrolysis. Fossil energy hydrogen production consists of coal hydrogen production, natural gas hydrogen production, petroleum hydrogen production, etc.; industrial byproduct hydrogen production refers to hydrogen production from industrial byproducts such as coke oven gas, methanol, synthetic ammonia, alcohols, etc.; biomass hydrogen production refers to gasification and microbial catalytic dehydrogenation to produce hydrogen; and hydrogen production through water electrolysis is a process that uses electricity to split water into hydrogen and oxygen, making it a perfect way to produce hydrogen [81–83] and a promising option for producing carbon-free hydrogen using renewable and nuclear resources. Affected by production costs, hydrogen production from fossil fuels still dominates the global hydrogen supply [84]. However, the growing use of renewable resources has drawn attention to the production of green hydrogen [85]. Many scholars have compared the economic viability and environmental performance of different hydrogen production technologies [86–89]. Richa Kothari et al. [90] evaluated the environmental impact of varying hydrogen production processes. They found that, at 75% system efficiency, the carbon dioxide emissions per kilogram of hydrogen produced using traditional fuels are 7.33 to 29.33 kg. The carbon dioxide emissions produced using renewable energy sources such as wind, solar, and water to produce hydrogen are zero. Currently, producing hydrogen from fossil fuels is the cheapest option in much of the world.

The average cost of producing hydrogen from natural gas ranges from USD 0.50 to 1.70 per kilogram (kg), depending on regional natural gas prices. Using CCUS technology to reduce CO_2 emissions in hydrogen production could increase production costs to around USD 1 to 2 per kilogram. Producing hydrogen using renewable electricity costs between USD 3 and 8 per kilogram [91]. Although using water electrolysis to produce hydrogen is expensive, it is a potentially cost-effective method of producing hydrogen at a distributed scale at a cost appropriate to meeting the hydrogen supply needs of early fuel-cell vehicles [90].

From 2018 to 2022, the hydrogen production method involved in Topic 8 was hydrogen production via water electrolysis. The current mainstream water electrolysis technologies include alkaline water electrolyzers [92], proton-exchange membrane water electrolyzers [93], and solid-oxide water electrolyzers [94]. The catalysts for hydrogen production by water electrolysis are precious-metal catalysts, molybdenum-series catalysts, nickel-based catalysts, high-entropy alloys, Pt-based catalysts, nanoparticle composite materials, etc. They aim to improve electrocatalytic and hydrogen productive efficiency and promote industrial applications. Topic 8 includes hydrogen production methods, including formaldehyde, methane, formic acid, acetic acid, ammonia, sodium borohydride, and other industrial-byproduct hydrogen production methods. The catalysts involved include metal-based catalysts, nanocatalysts, etc. Their primary purpose is to improve catalytic activity and cyclic stability. In addition, this topic describes a small number of fossil energy reforming hydrogen production technologies and biomass hydrogen production technologies. Topic 12 includes electrolytic cells, proton exchange membranes, catalysts, electrodes, water electrolysis hydrogen production systems, and other related technologies for electrolyzing water to produce hydrogen. Electrolysis is a process that uses electricity to split water into hydrogen and oxygen. The size range of electrolytic cells is wide. Electrolysis hydrogen production from water is suitable for production ranging from small household appliances to large-scale facilities connected to renewable energy. From 2008 to 2022, this topic focused on improving the efficiency and stability of the hydrogen evolution reaction to reduce the cost of hydrogen production.

Photocatalytic hydrogen production uses the current generated by the light absorbed by the semiconductor material to decompose water, thereby producing hydrogen and oxygen. Initially, photocatalytic hydrogen production used TiO_2 as a catalyst. So far, various substances and their derivatives have been found to exhibit this performance. According to elemental composition, they can be divided into elemental photocatalysts, metal-oxide photocatalysts, metal–carbon–nitrogen photocatalysts, chemical photocatalysts, metal–organic-framework-compound photocatalysts, covalent-organic-framework compound photocatalysts, and graphite-phase carbon-nitride photocatalysts. As photocatalytic hydrogen production uses solar energy to provide energy for water splitting and is environmentally friendly, Topic 17 has received widespread attention from researchers in the field of hydrogen energy.

To improve the hydrogen yield and productive efficiency of green energy systems, many scholars have conducted studies regarding combining photovoltaic and wind energy systems with water-electrolysis technology [95]. From 2018 to 2022, the technical topics described in Topic 9 included wind-power stable hydrogen supply systems, wind-power water-electrolysis hydrogen-production-system collaborative control methods, offshore wind-power hydrogen production methods and systems, and methods and systems for hydrogen production from solar. As wind and solar are intermittent energy sources, hydrogen has been considered one of the most realistic alternatives for the long-term storage of renewable energy [96]. Research on collaborative systems combining wind, solar, and hydrogen storage has also been a trend in recent years. The above is intended to solve the problems of wide power fluctuation, a large-scale stable hydrogen supply, and the economic viability of photovoltaic power generation and wind-power hydrogen production.

Hydrogen is flammable, has an extensive explosion concentration range, has strong diffusivity, and is easy to leak. Gathering hydrogen within a closed space can quickly form an explosive environment. When hydrogen encounters a fire source, an intense

explosion will occur. Ventilation and leak detection are essential factors in the design of safe hydrogen systems. In addition, hydrogen is colorless and odorless, and its combustion flame is difficult to detect with the naked eye. Hydrogen flame detection is essential for hydrogen safety. In addition, hydrogen can cause hydrogen embrittlement in some metals, causing many difficulties regarding hydrogen storage. Therefore, choosing suitable materials is critical to ensure hydrogen's safe use [97]. It can be seen that hydrogen safety technology is required in the production, storage, transportation, and use of hydrogen, as it ensures its secure handling and use [98]. Issues that need to be considered include, but are not limited to, assessing where hydrogen should be stored, calculating where the hydrogen will go after it escapes, safety issues about the wide range of operation of electrolyzers [99], looking at specific sensor types and placements [100], safety monitoring of hydrogen utilization equipment and systems [101,102], safe distance requirements [103], assessing the risks associated with hydrogen refueling stations [104–106], and the risks and reliability of hydrogen storage and delivery systems [107,108]. In our study, Topic 0 and Topic 19 described the infrastructure, safety equipment, monitoring, and control systems involved in all aspects of hydrogen utilization. Hydrogen security is critical to hydrogen supply chain establishment, and, since 2016, researchers' interest in these two topics has significantly increased.

Through the discussion, the developmental status of and trends in emerging hydrogen technologies have been explored, and critical development directions to promote the development of green energy technology and energy transformation have been clarified.

5. Conclusions

Hydrogen energy is a green, low-carbon, widely used secondary energy source. It has enormous potential to reduce emissions in the industrial sector. As costs fall, hydrogen will become the primary low-carbon fuel in the transportation sector, and hydrogen is one of the ways to store renewable energy for the long term. This is why hydrogen is gradually becoming one of the essential carriers of the global energy transition. However, hydrogen energy technology has not been widely used because of the high cost of producing hydrogen from renewable energy, inadequate hydrogen infrastructure, incomplete standards and regulations regarding safety, etc. Whether new innovative solutions have the potential to impact future economic and social development is one of the critical factors in whether they can be successfully applied. Therefore, identifying emerging technologies with impactful and novel features is important for developing hydrogen energy and promoting the energy transition.

To identify emerging technologies objectively, efficiently, and accurately from big data, an emerging technology identification method based on text mining was conducted in this study using the LDA model and incorporating technical experts' knowledge. Hydrogen-energy-related patent texts from 2008 to 2022 were analyzed based on this method. A total of 20 topic clusters were formed from the dataset, consisting of 9611 patents. We proposed two indicators, topic strength and topic novelty, to evaluate emerging technologies.

The technology of the materials for photocatalytic hydrogen production displayed the most robust performance, followed by monitoring and safety control systems for hydrogen-energy-related and other apparatus, judging from the cumulative topic strength from 2008 to 2022. The overall trend indicates a growing research interest in hydrogen production from wind and solar, while some topics, such as hydrogen production from biomass, show a decline in interest. Photocatalytic hydrogen production materials rapidly increased in research interest until 2018, then decreased slightly. The evolution of hydrogen storage and transportation technologies shows fluctuations, with some topics experiencing a decline in interest as technology matures. Safety monitoring methods for hydrogen fuel cells and hydrogen-energy-related equipment have gained increased attention over the years. Among 20 topics, photocatalytic hydrogen production materials had the greatest novelty, followed by the technology of devices, systems, and detection methods for hydrogen energy.

Six key emerging technologies were obtained based on the characteristics of hydrogen energy technology and industrial development requirements. In order to reduce the production cost of green hydrogen, address renewables' volatility, reduce safety risks in the hydrogen industrial chain, and achieve a diversified energy supply, governments, enterprises, and researchers may focus on these emerging technologies: hydrogen production technology using renewable energy sources such as wind and light; water electrolysis; photocatalytic hydrogen production technology; research and development in hydrogen production catalysts; and monitoring and safety technology for hydrogen-energy-related equipment.

The identification method for emerging technology proposed by this paper, based on machine learning and expert knowledge, can perform topic clustering from a large amount of data and ensure the robustness of the results with professional technological understanding, which enriches the means of technology identification. Due to the rapid development and interdisciplinary nature of hydrogen energy, this is an effective means to study hydrogen energy technology. Fully exploring the connotations and features of emerging technologies is the key to identifying them. The discriminant index proposed in this article fills the gap in previous research on hydrogen energy technology developmental trends. It can provide more theoretical support for future options in technological paths for hydrogen energy.

This paper has several limitations. Firstly, there are overlapping topics, such as Topic 1 and 19, Topic 4 and 16, and Topic 5 and 7. Overlapping was caused by specific high-probability words, and leads to spurious correlations, making distinguishing subtopics hard and reducing topic interpretability. Secondly, this paper relied on the top 20 highest-weight words to determine the meaning of a topic, which results in a certain bias. Therefore, using an LDA model containing effective structural semantics analysis methods, especially for multiple languages, is recommended to improve the explainability of topics or to improve the LDA model to implement semi-supervised machine-learning methods.

Author Contributions: Conceptualization, Y.Z. and Y.L.; methodology, Y.Z. and Y.L.; supervision, Y.Z.; data curation, Y.Z. and Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.Z. and Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China [grant numbers 71974107, L2224059, L2124002, 91646102, L1924058]; the MOE (Ministry of Education in China) Project of Humanities and Social Sciences [grant number 16JDGC011]; the Construction Project of China Knowledge Center for Engineering Sciences and Technology [grant number CKCEST-2022-1-30]; the Tsinghua University Initiative Scientific Research Program [grant numbers: 2022THZWYY09, 2019Z02CAU]; and the Tsinghua University Project of Volvo-Supported Green Economy and Sustainable Development [grant number 20183910020].

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 corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Silvera, I.F.; Dias, R. Metallic hydrogen. *J. Phys. Condens. Matter* **2018**, *30*, 254003. [CrossRef] [PubMed]
2. Keshipour, S.; Asghari, A. A review on hydrogen generation by phthalocyanines. *Int. J. Hydrogen Energy* **2022**, *47*, 12865–12881. [CrossRef]
3. Grochala, W. First there was hydrogen. *Nat. Chem.* **2015**, *7*, 264. [CrossRef] [PubMed]
4. Zhou, Y.; Zhou, R.; Chen, L.; Zhao, Y.; Zhang, Q. Environmental Policy Mixes and Green Industrial Development: An Empirical Study of the Chinese Textile Industry from 1998 to 2012. *IEEE Trans. Eng. Manag.* **2022**, *69*, 742–754. [CrossRef]
5. Nowotny, J.; Hoshino, T.; Dodson, J.; Atanacio, A.J.; Ionescu, M.; Peterson, V.; Prince, K.E.; Yamawaki, M.; Bak, T.; Sigmund, W.; et al. Towards sustainable energy. Generation of hydrogen fuel using nuclear energy. *Int. J. Hydrogen Energy* **2016**, *41*, 12812–12825. [CrossRef]

6. Hennicke, P.; Fischedick, M. Towards sustainable energy systems: The related role of hydrogen. *Energy Policy* **2006**, *34*, 1260–1270. [CrossRef]
7. Hosseini, S.E.; Wahid, M.A. Hydrogen from solar energy, a clean energy carrier from a sustainable source of energy. *Int. J. Energy Res.* **2020**, *44*, 4110–4131. [CrossRef]
8. Elam, C.C.; Padró, C.E.G.; Sandrock, G.; Luzzi, A.; Lindblad, P.; Hagen, E.F. Realizing the hydrogen future: The International Energy Agency's efforts to advance hydrogen energy technologies. *Int. J. Hydrogen Energy* **2003**, *28*, 601–607. [CrossRef]
9. Newborough, M.; Cooley, G. Developments in the global hydrogen market: The spectrum of hydrogen colours. *Fuel Cells Bull.* **2020**, *2020*, 16–22. [CrossRef]
10. Noussan, M.; Raimondi, P.P.; Scita, R.; Hafner, M. The Role of Green and Blue Hydrogen in the Energy Transition—A Technological and Geopolitical Perspective. *Sustainability* **2021**, *13*, 298. [CrossRef]
11. Shirizadeh, B.; Quirion, P. Long-term optimization of the hydrogen-electricity nexus in France: Green, blue, or pink hydrogen? *Energy Policy* **2023**, *181*, 113702. [CrossRef]
12. Ramachandran, R.; Menon, R.K. An overview of industrial uses of hydrogen. *Int. J. Hydrogen Energy* **1998**, *23*, 593–598. [CrossRef]
13. Acar, C.; Dincer, I. The potential role of hydrogen as a sustainable transportation fuel to combat global warming. *Int. J. Hydrogen Energy* **2020**, *45*, 3396–3406. [CrossRef]
14. Endo, N.; Shimoda, E.; Goshome, K.; Yamane, T.; Nozu, T.; Maeda, T. Construction and operation of hydrogen energy utilization system for a zero emission building. *Int. J. Hydrogen Energy* **2019**, *44*, 14596–14604. [CrossRef]
15. Rezaee Jordehi, A.; Mansouri, S.A.; Tostado-Véliz, M.; Iqbal, A.; Marzband, M.; Jurado, F. Industrial energy hubs with electric, thermal and hydrogen demands for resilience enhancement of mobile storage-integrated power systems. *Int. J. Hydrogen Energy* **2023**, *50*, 77–91. [CrossRef]
16. Mazloomi, K.; Gomes, C. Hydrogen as an energy carrier: Prospects and challenges. *Renew. Sustain. Energy Rev.* **2012**, *16*, 3024–3033. [CrossRef]
17. Edwards, P.P.; Kuznetsov, V.L.; David, W.I. Hydrogen energy. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2007**, *365*, 1043–1056. [CrossRef]
18. Noyan, O.F.; Hasan, M.M.; Pala, N. A Global Review of the Hydrogen Energy Eco-System. *Energies* **2023**, *16*, 1484. [CrossRef]
19. Rotolo, D.; Hicks, D.; Martin, B.R. What is an emerging technology? *Res. Policy* **2015**, *44*, 1827–1843. [CrossRef]
20. Zhou, Y.; Dong, F.; Liu, Y.; Ran, L. A deep learning framework to early identify emerging technologies in large-scale outlier patents: An empirical study of CNC machine tool. *Scientometrics* **2021**, *126*, 969–994. [CrossRef]
21. Halaweh, M. Emerging technology: What is it. *J. Technol. Manag. Innov.* **2013**, *8*, 108–115. [CrossRef]
22. Dhar, S.; Tarafdar, P.; Bose, I. Understanding the evolution of an emerging technological paradigm and its impact: The case of Digital Twin. *Technol. Forecast. Soc. Chang.* **2022**, *185*, 122098. [CrossRef]
23. Sazali, N. Emerging technologies by hydrogen: A review. *Int. J. Hydrogen Energy* **2020**, *45*, 18753–18771. [CrossRef]
24. Kyebambe, M.N.; Cheng, G.; Huang, Y.; He, C.; Zhang, Z. Forecasting emerging technologies: A supervised learning approach through patent analysis. *Technol. Forecast. Soc. Chang.* **2017**, *125*, 236–244. [CrossRef]
25. Zhou, Y.; Dong, F.; Liu, Y.; Li, Z.; Du, J.; Zhang, L. Forecasting emerging technologies using data augmentation and deep learning. *Scientometrics* **2020**, *123*, 1–29. [CrossRef]
26. Xu, G.; Wu, Y.; Minshall, T.; Zhou, Y. Exploring innovation ecosystems across science, technology, and business: A case of 3D printing in China. *Technol. Forecast. Soc. Chang.* **2018**, *136*, 208–221. [CrossRef]
27. Roberts, E.B. Exploratory and normative technological forecasting: A critical appraisal. *Technol. Forecast.* **1969**, *1*, 113–127. [CrossRef]
28. Xu, G.; Zhou, Y.; Ji, H. How Can Government Promote Technology Diffusion in Manufacturing Paradigm Shift? Evidence from China. *IEEE Trans. Eng. Manag.* **2023**, *70*, 1547–1559. [CrossRef]
29. Zhou, Y.; Li, X.; Lema, R.; Urban, F. Comparing the knowledge bases of wind turbine firms in Asia and Europe: Patent trajectories, networks, and globalisation. *Sci. Public Policy* **2015**, *43*, 476–491. [CrossRef]
30. Kong, D.; Zhou, Y.; Liu, Y.; Xue, L. Using the data mining method to assess the innovation gap: A case of industrial robotics in a catching-up country. *Technol. Forecast. Soc. Chang.* **2017**, *119*, 80–97. [CrossRef]
31. Cho, Y. Investigating the merge of exploratory and normative technology forecasting methods. In Proceedings of the 2013 Proceedings of PICMET '13: Technology Management in the IT-Driven Services (PICMET), San Jose, CA, USA, 28 July–1 August 2013; pp. 2083–2092.
32. Marques, J.; Guillo, M.; Bas, E.; Ramazanova, M.; Albuquerque, H. Setting Research Priorities for Effective Climate Change Management and Policymaking: A Delphi Study in Bolivia and Paraguay. *Sustainability* **2023**, *15*, 14993. [CrossRef]
33. Zhou, Y.; Zang, J.; Miao, Z.; Minshall, T. Upgrading Pathways of Intelligent Manufacturing in China: Transitioning across Technological Paradigms. *Engineering* **2019**, *5*, 691–701. [CrossRef]
34. Chaovalitwongse, W.A.; Yuan, Y.; Zhang, Q.; Liu, J. Special issue: Innovative applications of big data and artificial intelligence. *Front. Eng. Manag.* **2022**, *9*, 517–519. [CrossRef]
35. Borgi, T.; Zoghlami, N.; Abed, M.; Naceur, M.S. Big Data for Operational Efficiency of Transport and Logistics: A Review. In Proceedings of the 2017 6th IEEE International Conference on Advanced Logistics and Transport (ICALT), Bali, Indonesia, 24–27 July 2017; pp. 113–120.
36. Thai, M.T.; Wu, W.; Xiong, H. *Big Data in Complex and Social Networks*; CRC Press: Boca Raton, FL, USA, 2016.

37. Basit, T. Manual or electronic? The role of coding in qualitative data analysis. *Educ. Res.* **2003**, *45*, 143–154. [CrossRef]
38. Hacking, C.; Verbeek, H.; Hamers, J.P.; Aarts, S. Comparing text mining and manual coding methods: Analysing interview data on quality of care in long-term care for older adults. *PLoS ONE* **2023**, *18*, e0292578. [CrossRef]
39. Kostoff, R.N.; Toothman, D.R.; Eberhart, H.J.; Humenik, J.A. Text mining using database tomography and bibliometrics: A review. *Technol. Forecast. Soc. Chang.* **2001**, *68*, 223–253. [CrossRef]
40. Huang, L.; Hou, Z.; Fang, Y.; Liu, J.; Shi, T. Evolution of CCUS Technologies Using LDA Topic Model and Derwent Patent Data. *Energies* **2023**, *16*, 2556. [CrossRef]
41. Joung, J.; Kim, K. Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. *Technol. Forecast. Soc. Chang.* **2017**, *114*, 281–292. [CrossRef]
42. Tian, C.; Zhang, J.; Liu, D.; Wang, Q.; Lin, S. Technological topic analysis of standard-essential patents based on the improved Latent Dirichlet Allocation (LDA) model. *Technol. Anal. Strateg. Manag.* **2022**. [CrossRef]
43. Sankaran, S.; Killen, C.P.; Pitsis, A. How do project-oriented organizations enhance innovation? An institutional theory perspective. *Front. Eng. Manag.* **2023**, *10*, 427–438. [CrossRef]
44. Jiang, S.; Yang, J.; Yu, M.; Lin, H.; Li, C.; Doty, H. Strategic conformity, organizational learning ambidexterity, and corporate innovation performance: An inverted U-shaped curve? *J. Bus. Res.* **2022**, *149*, 424–433. [CrossRef]
45. Lin, H.; Zeng, S.; Liu, H.; Li, C. Bridging the gaps or fecklessness? A moderated mediating examination of intermediaries' effects on corporate innovation. *Technovation* **2020**, *94–95*, 102018. [CrossRef]
46. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
47. Maskeri, G.; Sarkar, S.; Heafield, K. Mining business topics in source code using latent dirichlet allocation. In Proceedings of the 1st India Software Engineering Conference, Hyderabad, India, 19–22 February 2008; pp. 113–120.
48. Campbell, J.C.; Hindle, A.; Stroulia, E. Chapter 6—Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data. In *the Art and Science of Analyzing Software Data*; Bird, C., Menzies, T., Zimmermann, T., Eds.; Morgan Kaufmann: Boston, MA, USA, 2015; pp. 139–159.
49. Agarwal, D.; Chen, B.-C. fLDA: Matrix factorization through latent dirichlet allocation. In Proceedings of the Third ACM International Conference on Web Search and Data Mining, New York, NY, USA, 4–6 February 2010; pp. 91–100.
50. Zhou, Y.; Dong, F.; Kong, D.; Liu, Y. Unfolding the convergence process of scientific knowledge for the early identification of emerging technologies. *Technol. Forecast. Soc. Chang.* **2019**, *144*, 205–220. [CrossRef]
51. Zhou, Y.; Miao, Z.; Urban, F. China's leadership in the hydropower sector: Identifying green windows of opportunity for technological catch-up. *Ind. Corp. Chang.* **2021**, *29*, 1319–1343. [CrossRef]
52. Wang, D.; Thint, M.; Al-Rubaie, A. Semi-Supervised Latent Dirichlet Allocation and Its Application for Document Classification. In Proceedings of the 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Macau, China, 4–7 December 2012; pp. 306–310.
53. Dingli, A.; Ciravegna, F.; Wilks, Y. *Automatic Semantic Annotation Using Unsupervised Information Extraction and Integration*; CEUR Workshop Proceedings: Aachen, Germany, 2003.
54. Wu, H.; Ma, T.; Wu, L.; Manyumwa, T.; Ji, S. Unsupervised reference-free summary quality evaluation via contrastive learning. *arXiv* **2020**, arXiv:2010.01781.
55. Ciravegna, F.; Dingli, A.; Petrelli, D.; Wilks, Y. User-system cooperation in document annotation based on information extraction. In Proceedings of the Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web: 13th International Conference, EKAW 2002, Sigüenza, Spain, 1–4 October 2002; Proceedings 13. pp. 122–137.
56. Asiyabi, R.M.; Datcu, M. Earth Observation Semantic Data Mining: Latent Dirichlet Allocation-Based Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 2607–2620. [CrossRef]
57. Park, Y.; Yoon, B.; Lee, S. The idiosyncrasy and dynamism of technological innovation across industries: Patent citation analysis. *Technol. Soc.* **2005**, *27*, 471–485. [CrossRef]
58. Vijayarani, S.; Ilamathi, M.J.; Nithya, M. Preprocessing techniques for text mining-an overview. *Int. J. Comput. Sci. Commun. Netw.* **2015**, *5*, 7–16.
59. Naik, D.A.; Mythreyan, S.; Seema, S. Relevance Feature Discovery in Text Mining Using NLP. In Proceedings of the 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, 27–29 May 2022; pp. 1–6.
60. Abidin, D.Z.; Nurmaini, S.; Malik, R.F.; Jasmir; Rasywir, E.; Pratama, Y. A Model of Preprocessing for Social Media Data Extraction. In Proceedings of the 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), Jakarta, Indonesia, 24–25 October 2019; pp. 67–72.
61. Zhou, Z.; Qin, J.; Xiang, X.; Tan, Y.; Liu, Q.; Xiong, N.N. News Text Topic Clustering Optimized Method Based on TF-IDF Algorithm on Spark. *Comput. Mater. Contin.* **2020**, *62*, 217–231. [CrossRef]
62. Yuan, Y.; Du, J.; Lee, J.M. Tourism activity recognition and discovery based on improved LDA model. In Proceedings of the 2016 4th International Conference on Cloud Computing and Intelligence Systems (CCIS), Beijing, China, 17–19 August 2016; pp. 447–455.
63. Chien, J.-T.; Lee, C.-H.; Tan, Z.-H. Latent Dirichlet mixture model. *Neurocomputing* **2018**, *278*, 12–22. [CrossRef]
64. Jelodar, H.; Wang, Y.; Yuan, C.; Feng, X.; Jiang, X.; Li, Y.; Zhao, L. Latent Dirichlet allocation (LDA) and topic modeling: Models, applications, a survey. *Multimed. Tools Appl.* **2019**, *78*, 15169–15211. [CrossRef]

65. Anupriya, P.; Karpagavalli, S. LDA based topic modeling of journal abstracts. In Proceedings of the 2015 International Conference on Advanced Computing and Communication Systems, Coimbatore, India, 5–7 January 2015; pp. 1–5.
66. Kobayashi, H. Perplexity on reduced corpora. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Baltimore, MD, USA, 23–25 June 2014; pp. 797–806.
67. Wang, H.; Wang, J.; Zhang, Y.; Wang, M.; Mao, C. Optimization of Topic Recognition Model for News Texts Based on LDA. *J. Digit. Inf. Manag.* **2019**, *17*, 257. [CrossRef]
68. Bíró, I. Document Classification with Latent Dirichlet Allocation. Ph.D. Thesis, Eotvos Lorand University, Budapest, Hungary, 2009; Volume 4.
69. Hasan, M.; Rahman, A.; Karim, M.R.; Khan, M.S.I.; Islam, M.J. Normalized approach to find optimal number of topics in Latent Dirichlet Allocation (LDA). In *Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*; Springer: Singapore, 2021; pp. 341–354.
70. Hidayatullah, A.F.; Ma'arif, M.R. Road traffic topic modeling on Twitter using latent dirichlet allocation. In Proceedings of the 2017 International Conference on Sustainable Information Engineering and Technology (SIET), Malang, Indonesia, 24–25 November 2017; pp. 47–52.
71. Wang, J.; Fan, Y.; Zhang, H.; Feng, L. Technology Hotspot Tracking: Topic Discovery and Evolution of China's Blockchain Patents Based on a Dynamic LDA Model. *Symmetry* **2021**, *13*, 415. [CrossRef]
72. Li, C.; Feng, S.; Zeng, Q.; Ni, W.; Zhao, H.; Duan, H. Mining Dynamics of Research Topics Based on the Combined LDA and WordNet. *IEEE Access* **2019**, *7*, 6386–6399. [CrossRef]
73. Ausfelder, F.; Bazzanella, A. Hydrogen in the Chemical Industry. In *Hydrogen Science and Engineering: Materials, Processes, Systems and Technology*; Wiley: Hoboken, NJ, USA, 2016; pp. 19–40.
74. Shay, R.H.; Puerta, R.A.; Domchek, D.A. Advances in hydrogen usage in the metals and electronics industries. *Int. J. Hydrogen Energy* **1984**, *9*, 539–542. [CrossRef]
75. Kovač, A.; Paranos, M.; Marciuš, D. Hydrogen in energy transition: A review. *Int. J. Hydrogen Energy* **2021**, *46*, 10016–10035. [CrossRef]
76. Bahja, M.; Safdar, G.A. Unlink the link between COVID-19 and 5G networks: An NLP and SNA based approach. *IEEE Access* **2020**, *8*, 209127–209137. [CrossRef]
77. Sievert, C.; Shirley, K. LDAvis: A method for visualizing and interpreting topics. In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Baltimore, MD, USA, 27 June 2014; pp. 63–70.
78. Xiang, M.; Fu, D.; Lv, K. Identifying and Predicting Trends of Disruptive Technologies: An Empirical Study Based on Text Mining and Time Series Forecasting. *Sustainability* **2023**, *15*, 5412. [CrossRef]
79. Choi, H.; Woo, J. Investigating emerging hydrogen technology topics and comparing national level technological focus: Patent analysis using a structural topic model. *Appl. Energy* **2022**, *313*, 118898. [CrossRef]
80. Bickel, M.W. Reflecting trends in the academic landscape of sustainable energy using probabilistic topic modeling. *Energy Sustain. Soc.* **2019**, *9*, 49. [CrossRef]
81. Wang, S.; Lu, A.; Zhong, C.-J. Hydrogen production from water electrolysis: Role of catalysts. *Nano Converg.* **2021**, *8*, 4. [CrossRef] [PubMed]
82. Ahmed, S.F.; Mofijur, M.; Nuzhat, S.; Rafa, N.; Musharrat, A.; Lam, S.S.; Boretti, A. Sustainable hydrogen production: Technological advancements and economic analysis. *Int. J. Hydrogen Energy* **2022**, *47*, 37227–37255. [CrossRef]
83. Zhang, B.; Zhang, S.-X.; Yao, R.; Wu, Y.-H.; Qiu, J.-S. Progress and prospects of hydrogen production: Opportunities and challenges. *J. Electron. Sci. Technol.* **2021**, *19*, 100080. [CrossRef]
84. Megía, P.J.; Vizcaíno, A.J.; Calles, J.A.; Carrero, A. Hydrogen Production Technologies: From Fossil Fuels toward Renewable Sources. A Mini Review. *Energy Fuels* **2021**, *35*, 16403–16415. [CrossRef]
85. Hosseini, S.E.; Wahid, M.A. Hydrogen production from renewable and sustainable energy resources: Promising green energy carrier for clean development. *Renew. Sustain. Energy Rev.* **2016**, *57*, 850–866. [CrossRef]
86. Yukesh Kannah, R.; Kavitha, S.; Preethi; Parthiba Karthikeyan, O.; Kumar, G.; Dai-Viet, N.V.; Rajesh Banu, J. Techno-economic assessment of various hydrogen production methods—A review. *Bioresour. Technol.* **2021**, *319*, 124175. [CrossRef]
87. Kaiwen, L.; Bin, Y.; Tao, Z. Economic analysis of hydrogen production from steam reforming process: A literature review. *Energy Sources Part B Econ. Plan. Policy* **2018**, *13*, 109–115. [CrossRef]
88. Osman, A.I.; Mehta, N.; Elgarahy, A.M.; Hefny, M.; Al-Hinai, A.; Al-Muhtaseb, A.H.; Rooney, D.W. Hydrogen production, storage, utilisation and environmental impacts: A review. *Environ. Chem. Lett.* **2022**, *20*, 153–188. [CrossRef]
89. Acar, C.; Dincer, I. Review and evaluation of hydrogen production options for better environment. *J. Clean. Prod.* **2019**, *218*, 835–849. [CrossRef]
90. Kothari, R.; Buddhi, D.; Sawhney, R.L. Comparison of environmental and economic aspects of various hydrogen production methods. *Renew. Sustain. Energy Rev.* **2008**, *12*, 553–563. [CrossRef]
91. IEA. *Global Hydrogen Review 2021*; IEA: Paris, France, 2021.
92. David, M.; Ocampo-Martínez, C.; Sánchez-Peña, R. Advances in alkaline water electrolyzers: A review. *J. Energy Storage* **2019**, *23*, 392–403. [CrossRef]

93. Madheswaran, D.K.; Thangamuthu, M.; Gnanasekaran, S.; Gopi, S.; Ayyasamy, T.; Pardeshi, S.S. Powering the Future: Progress and Hurdles in Developing Proton Exchange Membrane Fuel Cell Components to Achieve Department of Energy Goals—A Systematic Review. *Sustainability* **2023**, *15*, 15923. [CrossRef]
94. Ni, M.; Leung, M.K.; Leung, D.Y. Technological development of hydrogen production by solid oxide electrolyzer cell (SOEC). *Int. J. Hydrogen Energy* **2008**, *33*, 2337–2354. [CrossRef]
95. Benghanem, M.; Mellit, A.; Almohamadi, H.; Haddad, S.; Chettibi, N.; Alanazi, A.M.; Dasalla, D.; Alzahrani, A. Hydrogen Production Methods Based on Solar and Wind Energy: A Review. *Energies* **2023**, *16*, 757. [CrossRef]
96. Egeland-Eriksen, T.; Hajizadeh, A.; Sartori, S. Hydrogen-based systems for integration of renewable energy in power systems: Achievements and perspectives. *Int. J. Hydrogen Energy* **2021**, *46*, 31963–31983. [CrossRef]
97. Ahad, M.T.; Bhuiyan, M.M.H.; Sakib, A.N.; Becerril Corral, A.; Siddique, Z. An Overview of Challenges for the Future of Hydrogen. *Materials* **2023**, *16*, 6680. [CrossRef]
98. Najjar, Y.S.H. Hydrogen safety: The road toward green technology. *Int. J. Hydrogen Energy* **2013**, *38*, 10716–10728. [CrossRef]
99. Haoran, C.; Xia, Y.; Wei, W.; Yongzhi, Z.; Bo, Z.; Leiqi, Z. Safety and efficiency problems of hydrogen production from alkaline water electrolyzers driven by renewable energy sources. *Int. J. Hydrogen Energy* **2023**, *54*, 700–712. [CrossRef]
100. Buttner, W.J.; Post, M.B.; Burgess, R.; Rivkin, C. An overview of hydrogen safety sensors and requirements. *Int. J. Hydrogen Energy* **2011**, *36*, 2462–2470. [CrossRef]
101. Klass, L.; Kabza, A.; Sehnke, F.; Strecker, K.; Hölzle, M. Lifelong performance monitoring of PEM fuel cells using machine learning models. *J. Power Source* **2023**, *580*, 233308. [CrossRef]
102. Folgado, F.J.; González, I.; Calderón, A.J. Data acquisition and monitoring system framed in Industrial Internet of Things for PEM hydrogen generators. *Internet Things* **2023**, *22*, 100795. [CrossRef]
103. West, M.; Al-Douri, A.; Hartmann, K.; Buttner, W.; Groth, K.M. Critical review and analysis of hydrogen safety data collection tools. *Int. J. Hydrogen Energy* **2022**, *47*, 17845–17858. [CrossRef]
104. Russo, P.; De Marco, A.; Mazzaro, M.; Capobianco, L. Quantitative risk assessment on a hydrogen refuelling station. *Chem. Eng. Trans.* **2018**, *67*, 739–744.
105. Li, Y.; Wang, Z.; Shi, X.; Fan, R. Safety analysis of hydrogen leakage accident with a mobile hydrogen refueling station. *Process Saf. Environ. Prot.* **2023**, *171*, 619–629. [CrossRef]
106. Tsunemi, K.; Kihara, T.; Kato, E.; Kawamoto, A.; Saburi, T. Quantitative risk assessment of the interior of a hydrogen refueling station considering safety barrier systems. *Int. J. Hydrogen Energy* **2019**, *44*, 23522–23531. [CrossRef]
107. Li, H.; Cao, X.; Liu, Y.; Shao, Y.; Nan, Z.; Teng, L.; Peng, W.; Bian, J. Safety of hydrogen storage and transportation: An overview on mechanisms, techniques, and challenges. *Energy Rep.* **2022**, *8*, 6258–6269. [CrossRef]
108. Zheng, J.; Liu, X.; Xu, P.; Liu, P.; Zhao, Y.; Yang, J. Development of high pressure gaseous hydrogen storage technologies. *Int. J. Hydrogen Energy* **2012**, *37*, 1048–1057. [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.