ENERGY RESOURCES OF THE EARTH'S CRUST: CHALLENGES OF THE MODERN ECONOMY

Original article
https://doi.org/10.29222/ipng.2078-5712.2025.01

Identification of relevant research topics based on pooled bibliometric data of the *International Journal of Hydrogen Energy* publications for 2022–2024 from The Lens and ScienceDirect platforms

Boris N. Chigarev ✉
Oil and Gas Research Institute, Russian Academy of Sciences, 3 Gubkin St., Moscow, 119333, Russia

Abstract. *Background.* Hydrogen is given the role of a clean energy carrier in the energy transition. *Objective.* To identify relevant research topics based on pooled bibliometric data of the *International Journal of Hydrogen Energy* publications from 2022–2024, collected from The Lens and ScienceDirect platforms. *Materials and methods*. 10,928 bibliometric records were exported from The Lens database and 10,857 records were exported from ScienceDirect. Keywords clustering and visual data analysis were performed using the following programs and algorithms: VOSviewer, Scimago Graphica, Inkscape, FP-growth utility. *Results*. The study showed the feasibility of merging bibliometric records from the open platforms ScienceDirect and The Lens, which complement each other. The "Fields of Study" data of The Lens was interpreted as system keywords similar to Scopus Index Keywords. The possibility of using "Fields of Study" data in a method similar to bibliographic coupling is shown. The practicality of using an alluvial diagram to show the co-occurrence of the four terms was demonstrated. The study also emphasizes the advisability of joint use of VOSviewer and Scimago Graphica programs for complex visualization of bibliometric analysis results. The study identified two dominant publication topics in the *International Journal of Hydrogen Energy* for 2022–2024, which can be described in terms of: "Catalysis, Hydrogen, Physical chemistry, Chemical engineering, Nanotechnology, Electrochemistry, Organic chemistry, Materials science" and "Electrical engineering, Hydrogen production, Renewable energy, Environmental science, Hydrogen economy, Hydrogen, Engineering". *Conclusions.* Joint use of metadata of different open abstract databases allows to compensate partial representation of their data.

Keywords: *International Journal of Hydrogen Energy*, The Lens, ScienceDirect, VOSviewer, Scimago Graphica, FP-growth

Funding: the work was funded by the Ministry of Science and Higher Education of the Russian Federation (State Assignment No. 125021302095-2).

For citation: Chigarev B.N. Identification of relevant research topics based on pooled bibliometric data of the *International Journal of Hydrogen Energy* publications for 2022–2024 from The Lens and ScienceDirect platforms. *Actual Problems of Oil and Gas*. 2025. https://doi.org/10.29222/ipng.2078-5712.2025.01

✉ Boris N. Chigarev, bchigarev@ipng.ru

Introduction

Motivation for the study

The lack of access to subscription abstract databases forces to turn to open access databases. The main issue with such databases is the incomplete coverage of bibliometric data compared to subscription databases. In this study, the author seeks a solution to this problem by combining data from several open access databases.

Objectives

The main objective of this study was to explore the possibility of joint utilization of bibliometric data from different open access abstract databases to merge data to mutually enrich missing fields in platform-specific data.

The secondary objective was to apply the aforementioned approach to identify relevant research topics based on pooled bibliometric data of the *International Journal of Hydrogen Energy* publications from 2022–2024 harvested from The Lens and ScienceDirect platforms.

Identification of publications on bibliometric analysis of metadata of the International Journal of Hydrogen Energy

In this section of the review, only publications with metadata in the ScienceDirect abstract database were considered; a detailed overview of other platforms was beyond the scope of this article. The aim was to identify typical bibliometric studies on topics related to hydrogen energy and to try to find articles that used aggregation of fields of bibliometric records from different abstract databases. It should be emphasized that such a task is relevant when using abstract databases with open access, in subscription databases such a task is not actual, as it is solved by the forces of such platforms.

A query to ScienceDirect: (bibliometric OR scientometric) AND “International Journal of Hydrogen Energy” in the fields “Title, abstract or author-specified keywords” returned 16 results. These include the number of publications by year: 2024 (6), 2023 (2), 2022 (4), 2021 (1), 2019 (1), 2016 (1), 2000 (1), the number of publications by type: Review articles (8), Research articles (7), Book chapters (1). This means that review articles dominate. Nine of these publications have been posted in the *International Journal of Hydrogen Energy* itself. The search results are summarized in Table 1. Records from ScienceDirect were exported in RIS format with T1, PY, KW fields corresponding to title, year; keywords fields are presented in the table.

Table 1. Sixteen publications matching the ScienceDirect query: (bibliometric OR scientometric) AND “International Journal of Hydrogen Energy” in the fields “Title, abstract or author-specified keywords”

Табл. 1. Шестнадцать публикаций, отвечающих запросу в ScienceDirect: (bibliometric OR scientometric) AND “International Journal of Hydrogen Energy” по полям “Title, abstract or author-specified keywords”

Title	Year	Keywords
1	2	3
Photosynthetic green hydrogen: Advances, challenges, opportunities, and prospects	2024	Green hydrogen; Photosynthetic production; Algae; Bibliometric analysis
Research on nickel-based catalysts for carbon dioxide methanation combined with literature measurement	2022	Methane; Catalyst; Hydrogen; Carbon dioxide; Nickel
Bibliometrical analysis of hydrogen storage	2019	Hydrogen storage; Bibliometrics; Visualization; Situation analysis

Table 1 continued

Продолжение табл. 1

1	2	3
Bibliometric analysis of the research on hydrogen economy: An analysis of current findings and roadmap ahead	2022	Hydrogen economy; Literature review; Hydrogen roadmap; Bibliometrics; Visualisation
Recent advancement in water electrolysis for hydrogen production: A comprehensive bibliometric analysis and technology updates	2024	Water electrolysis; Hydrogen production; Energy; Bibliometric; Green energy
Hydrogen-fuelled internal combustion engines - Bibliometric analysis on research trends, hotspots, and challenges	2024	Ammonia; Hydrogen; Injection strategy; Machine learning; Optimisation; Review
Mapping the knowledge domains of new energy vehicle safety: Informetrics analysis-based studies	2021	New energy vehicle (NEV); Safety; Research hotspot; Bibliometrics; Mapping knowledge domains
Sustainability and challenges in hydrogen production: An advanced bibliometric analysis	2023	Hydrogen production; Feedstocks; Research articles; Bibliometric analysis
Chapter 558 – Research activities in renewable energy in European countries, 1992–1998: A scientometric analysis	2000	
A bibliometric study on research trends in hydrogen safety	2022	Hydrogen safety; Bibliometric analysis; Knowledge structure; Visualisation
Progress in the production of hydrogen energy from food waste: A bibliometric analysis	2022	Bibliometric analysis; Food waste; Hydrogen energy; Research trends; Sustainability
Mapping the research on the spontaneous combustion of high-pressure hydrogen leakage: A bibliometric analysis	2024	High-pressure hydrogen; Spontaneous combustion; Research hotspots; Knowledge base; Bibliometrics analysis
An advanced bibliometric analysis and future research insights on safety of hydrogen energy	2024	Safety of hydrogen energy; Bibliometrics analysis; Visualization; Situation analysis
Chinese energy and fuels research priorities and trend: A bibliometric analysis	2016	Energy & fuels; Bibliometric; Bubble chart; Keyword analysis; Science Citation Index Expanded
Hydrogen energy storage integrated grid: A bibliometric analysis for sustainable energy production	2024	Hydrogen energy storage; Grid-connected; Highly-cited article; Sustainable production; Energy
Bibliometric analysis for research trends and hotspots in heat and mass transfer and its management of proton exchange membrane fuel cells	2023	Proton exchange membrane fuel cell; Heat and mass transfer; Water and heat management; Mapping Knowledge Domain; VOSviewer

Given the small number of articles and the quality of the journals in which they are published, they are easy to find by title and date of publication. The author considers it inexpedient to cite them in the reference list, since the current fashion for excessive citation leads to artificial citation boost.

The different combinations of the terms “Bibliometric analysis”, “Bibliometrics”, “Bibliometric”, “Bibliometrics analysis” occur 11 times, indicating the relevance of the topic at hand. “Visualisation” and “Visualization” are occurring 4 times and reflect the actuality of data analysis visualization. Among the most common terms related to hydrogen theme are “Energy”, “Hydrogen”, “Hydrogen production”, “Situation analysis”. The term “Situation analysis” is linked to the keywords “Hydrogen storage” and “Safety of hydrogen energy”, which may indicate the relevance of the topic of safe hydrogen storage, especially since the terms “Spontaneous combustion”, “Safety”, “Hydrogen safety” occur in the keywords.

Examples of publications that best reflect this topic.

This paper [1] analyzes 8,283 hydrogen safety research papers from 1992 to 2022, revealing a growing prominence in the field. China ranks first with 2,414 publications, while the United States, Germany and China have extensive international collaboration. The *International Journal of Hydrogen Energy* has the highest publication count. Research hotspots include storage safety, combustion and explosion safety, self-ignition, and embrittlement. However, understanding hydrogen autoignition mechanisms is limited, and further research is needed on prevention and control measures.

Hydrogen is crucial for addressing the greenhouse effect and energy crisis, as a clean fuel and storable energy carrier. Issues like hydrogen embrittlement, fire and explosion are of interest to researchers. The article [2] provides a knowledge structure of hydrogen safety publications from 1957 to 2021, with the USA having the most publications. The hot terms in hydrogen safety research are storage and detection, combustion and explosion, and ignition and propagation.

Publications that reflect bibliometric analysis methods of the topic under consideration.

The analysis of 1,275 articles published over the past five decades provides insights into research trends in the hydrogen economy [3]. The study found that the literature on the hydrogen economy has been increasing, particularly from 2016 to 2020. The *International Journal of Hydrogen Energy* contributes 40% of the overall publications. The analysis shows that the study on the hydrogen economy mostly deals with multidisciplinary aspects like hydrogen production, storage, transportation, application, and public policy formulation.

The research on heat and mass transfer in proton exchange membrane fuel cells (PEMFCs) has been growing rapidly since 2000 [4]. The field has gained significant attention in China, the United States and Canada, with high-quality publications from Canada. The *Journal of Power Sources* is the authoritative journal in this field. The field is divided into three stages: exploring gas-liquid two-phase flow characteristics, developing PEMFC heat and mass transfer and management, and optimizing water and heat management.

No articles were found that analyzed only the *International Journal of Hydrogen Energy* topics.

Bibliometric studies using a combination of metadata from different open referencing platforms

Below is a list of papers that share data from ScienceDirect AND The Lens platforms, presented in the format: paper reference → citation, consisting of a sentence taken from the abstract and containing the specified terms:

– Paper [5]: “Scopus, The Lens, PubMed, ScienceDirect and Microsoft academic were electronically searched. Peer-reviewed papers using sensors in post-stroke rehabilitation were included, for the period between 2015 to August 2021”.

– Paper [6]: “The search was conducted from January to July 2022, the descriptors used were ‘ADHD AND Physiotherapy’ being performed on the World Intellectual Property Organization (WIPO) patent database and articles and abstracts: Scopus Web of Science, PubMed, The Lens and ScienceDirect”.

– Paper [7]: “Methodology: a systematic review of articles in the databases Web of Science, dimensions, Scopus, ScienceDirect, Clinical Key, and The Lens was carried out between August 2022 – January 2023, identifying 150 records and through a selection process, 12 were included”.

– Paper [8]: “The technological search was carried out in the patent databases of Espacenet, Patentscope, Scopus, ScienceDirect, Google Patents, The Lens ORG, and USPTO from 2000 to 2021”.

– Paper [9]: “A search of ACM Digital Library, Cochrane Library, DARE, EBSCOHost, Medline, ProQuest, PubMed, ScienceDirect, Scopus and Web of Science produced 8,391 results”.

– Paper [10]: “This paper followed a systematic literature review of journal articles extracted from SCOPUS, Web of Science, ScienceDirect, and EBSCOhost electronic databases”.

– Paper [11]: “This research uses relevant Literature review which is reviewed following the search from 3 databases. They are ProQuest, ScienceDirect, and Google Scholar”.

The above publications can be categorized as systematic reviews.

There are publications comparing open abstract databases and subscription databases:

– Paper [12]: “The findings demonstrate that non-selective databases (Dimensions, OpenAlex, Scilit, and The Lens) index a greater amount of retracted literature than do databases that rely their indexation on venue selection (PubMed, Scopus, and WoS)”.

– Paper [13]: “Conversely, third-party databases (Dimensions, OpenAlex, Scilit, and The Lens) have more metadata quality and a higher completeness rate”.

– Paper [14]: “This study analyses the coverage of seven free-access bibliographic databases (CrossRef, Dimensions – non-subscription version, Google Scholar, The Lens, Microsoft Academic, Scilit, and Semantic Scholar) to identify the potential reasons that might cause the exclusion of scholarly documents and how they could influence coverage”.

– Paper [15]: “This communication aims to analyse the information that a large set of free-access databases (i.e., CrossRef, Dimensions, Microsoft Academic, OpenAlex, Scilit, Semantic Scholar, The Lens) provides about indexed publications in their databases”.

Summary conclusions from the literature review:

1. The most common references to free-access abstract databases are either systematic reviews or articles discussing comparative analysis of metadata populations. No publications were found in which fields from different databases were merged to improve the completeness of metadata.

2. The above publications may be of interest for systematic reviews because they not only list the most frequently used open access abstract databases, but also analyze the characteristics of their metadata.

Materials and methods

Data source

The data used in this paper were bibliometric records exported from The Lens and ScienceDirect platforms, which meet the following requirements:

– Source Title: International Journal of Hydrogen Energy;

– Publication Year: 2022–2024;

– The data are current as of 20 July 2024.

10,928 bibliometric records in CSV format were exported from The Lens database and 10,857 records in RIS format from ScienceDirect.

A comparison of DOI records yielded 10,551 matches. The difference is explained by the fact that the data exported from The Lens includes publications such as Short Communication and Letter to the Editor.

Programs and utilities

VOSviewer [16] was used to plot the keyword co-occurrence network and prepare a JSON file for visualization on app.vosviewer.com. The supplementary materials for this article (a ZIP archive) include JSON files that can be uploaded to this service for independent use[1].

Scimago Graphica [17] was used to construct the alluvial diagram and visualization of the co-occurrence network in coordinates Avg. pub. year (Average publication year) and Avg. norm. citations (Average normalized citations), as they are used in the VOSviewer program. Files of charts obtained using Scimago Graphica in SVG and interactive HTML formats are included in the attached archive.

Inkscape[2] was used for SVG files editing, for example, to correct the placement of the labels.

FP-growth (Frequent Pattern Growth algorithm) by Christian Borgelt [18] was used for the estimation of the co-occurrence of keywords.

Results and discussions

The subject matter of the journal's publications can be described by the systematization of the author keywords of individual articles. In The Lens platform data, the keywords field for this journal was empty. Therefore, the values were taken from the RIS records exported from ScienceDirect. On the other hand, the ScienceDirect records do not contain information about citations of publications. Therefore, they were taken from The Lens records. It should be emphasized that in this case the citation rate is based on information from a specific database, The Lens, and will differ from a citation rate based on data from another database, such as Scopus.

Author keywords clustering

Clustering of the author keywords was performed using VOSviewer software. Two iterations were performed: the first time the file my_thesaurus_terms.txt was built to bring some terms to a common form; this file was used in the second iteration.

The 500 keywords with the highest total link strength were clustered. The default output produced ten clusters, but the tenth cluster contained only one term – "carbon materials", so the minimum number of terms per cluster was increased to two, yielding nine clusters. Fig. 1 shows the general view of the author keyword clustering.

The data presented in Fig. 1 can be viewed in more detail using the AuthorKeywords-top500.json file located in the archive attached to this article by opening it in the online application https://app.vosviewer.com/.

Here are some findings that characterize the landscape of keyword distribution:

– Ten author keywords most frequently appearing in new publications: wind, sustainable energy, cushion gas, hydrogen supply chain, solar-to-hydrogen efficiency, techno-economic assessment, energy transition, flow field, electrochemical impedance spectroscopy, Tafel slope.

– Ten author keywords more frequently occurring in publications with high citations: clean energy, carbon capture, combined heat and power, bibliometric analysis, hydrogen economy, hybrid renewable energy system, s-scheme, sustainable development, cost, energy management system. As can be seen from this list, the term "bibliometric analysis" is often found in publications with high citation rates.

[1] Supplementary materials for this article are available via Figshare: https://doi.org/10.6084/m9.figshare.28468313.v1

[2] Inkscape: A free and open-source vector graphics editor. URL: https://inkscape.org (accessed 5 August 2024).

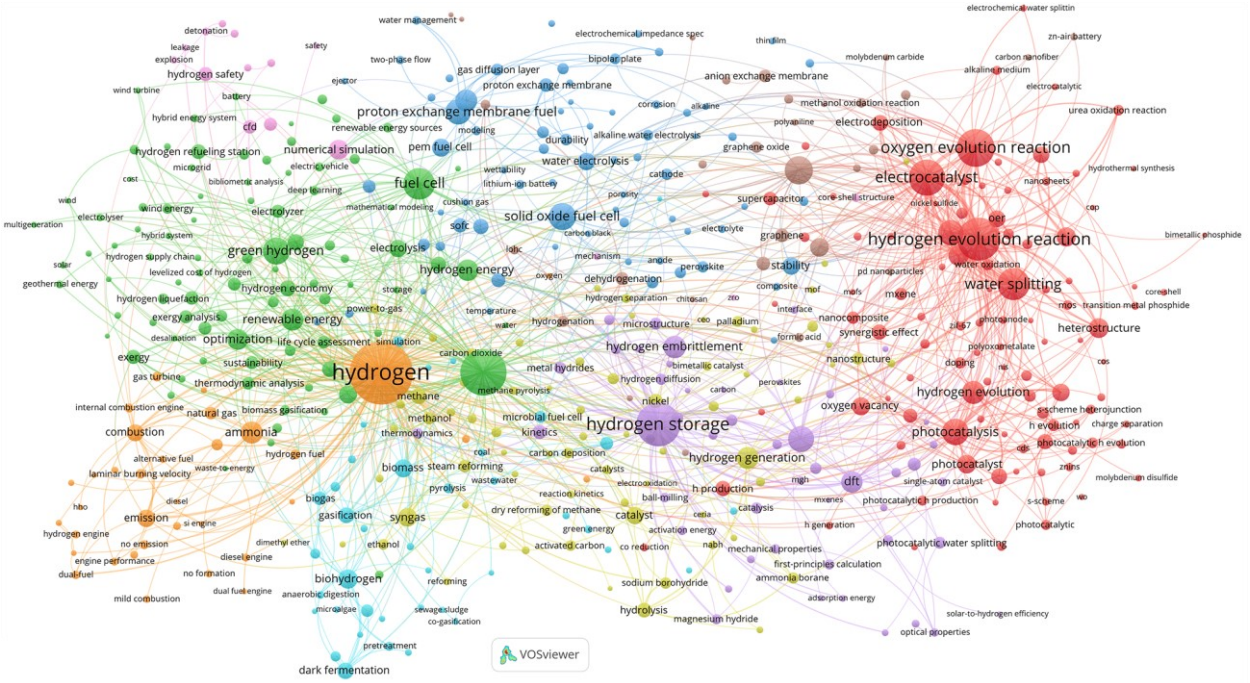

Fig. 1. Nine clusters of author keywords formed on the basis of their co-occurrence

Рис. 1. Девять кластеров авторских ключевых слов, сформированных на основе их совместной встречаемости

– Ten author keywords most frequently occurring in all publications: hydrogen, hydrogen production, hydrogen storage, hydrogen evolution reaction, oxygen evolution reaction, electrocatalyst, water splitting, fuel cell, oxygen reduction reaction, solid oxide fuel cell. These terms reflect well the dominant theme of the journal.

Clustering of the "Fields of Study" terms

As noted earlier, the terms "Fields of Study" can be considered as an analog of Index keywords in Scopus and used to describe the topics of publications.

Fig. 2 shows a similar picture to Fig. 1 but postulated for the "Fields of Study" terms.

The author keywords are more diverse (23,215 all terms; 1,390 meet >=5) than "Fields of Study" (5,152; 1,794 meet >=5), but due to their standardization, terms from "Fields of Study" are more often exceeding the threshold of five terms.

Here are some data characterizing the distribution of field terms "Fields of Study":

– Ten author keywords most frequently appearing in new publications: mechanism (biology), literature, dual (grammatical number), reduction (mathematics), scale (ratio), art, oxygen reduction, macroeconomics, production (economics), plasma.

– Ten author keywords more frequently occurring in publications with high citations: hydrogen technologies, hydrogen economy, software deployment, sustainability, energy carrier, greenhouse gas, climate change, fossil fuel, natural resource economics, risk analysis (engineering).

– Ten author keywords most frequently occurring in all publications: chemistry, materials science, engineering, organic chemistry, chemical engineering, hydrogen, catalysis, physics, physical chemistry, electrode.

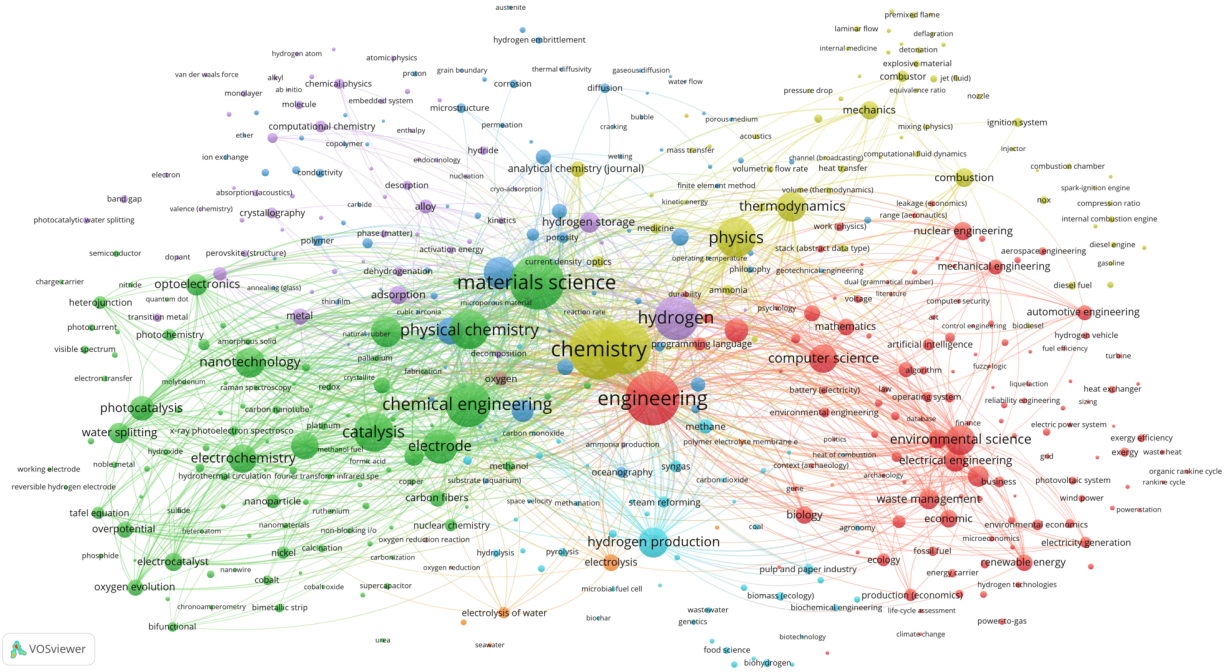

Fig. 2. Eight clusters of “Fields of Study” terms formed on the basis of their co-occurrence

Рис. 2. Восемь кластеров терминов “Fields of Study”, сформированных на основе их совместной встречаемости

The five most frequently occurring terms and the five most cited terms (the most cited term is identified by the citations of publications in which it appears) in the four largest clusters – green, red, khaki and purple:

1. *Green cluster:* materials science, chemical engineering, catalysis, physical chemistry, electrode; charge carrier, crystallinity, semiconductor, nanomaterials, nanocomposite.

2. *Red cluster:* engineering, environmental science, computer science, quantum mechanics, electrical engineering; hydrogen technologies, hydrogen economy, software deployment, sustainability, energy carrier.

3. *Khaki cluster:* chemistry, organic chemistry, physics, thermodynamics, combustion; hydrogen fuel enhancement, diesel engine, thermal efficiency, diesel fuel, biodiesel.

4. *Purple cluster:* hydrogen, hydrogen storage, adsorption, metal, alloy; ab initio, photocatalytic water splitting, magnesium hydride, Gibbs free energy, gravimetric analysis.

The selection of terms to build queries on the task of the researcher's interest can include many terms linked by the OR operator, but AND can rarely link more than 3–4 terms acting as filters, so the selection of terms linked by the AND operator requires more careful reasoning.

To select terms connected by the AND operator, it is useful to perform visualization of terms in the form of a network based on their co-occurrence. Most often, node placement algorithms for visualization seek to provide a good “readability” of the figure, but this approach is less informative than in the case of displaying a network of terms in specific coordinates.

According to the personal experience of the author of this article, such coordinates can be Avg. pub. year and Avg. norm. citations, as they are used in the VOSviewer program and described in its manual[3].

A similar visualization of the red and green cluster data is shown in Fig. 3 and Fig. 4.

The terms in each cluster were additionally grouped using an algorithm built into the Scimago Graphica program.

It is noteworthy that one cluster dominates in both figures, although the networks are built using a different algorithm than in Fig. 2, which may indicate the stability of the results obtained.

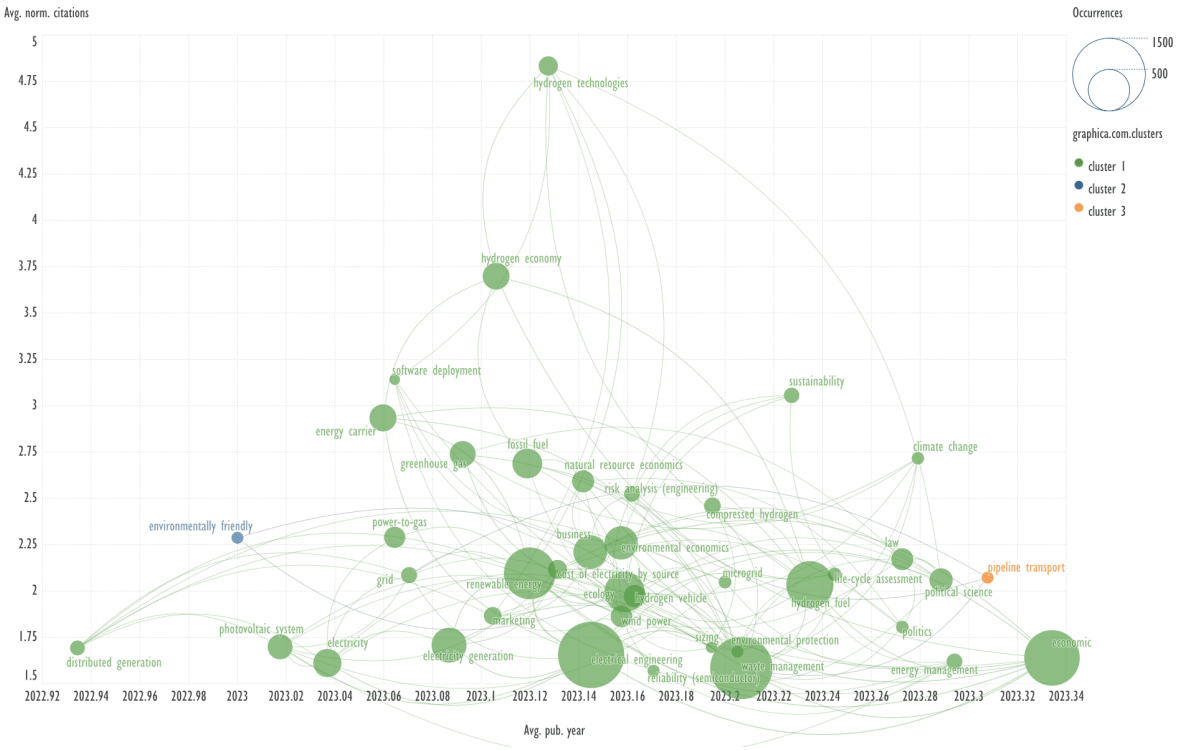

Fig. 3. Visualization of the co-occurrence network of the 40 most cited “Fields of Study” terms belonging to the red cluster of Fig. 2. The citation score is given by the value of “Avg. norm. citations”

Рис. 3. Визуализация сети совместной встречаемости 40 наиболее цитируемых терминов “Fields of Study”, входящих в красный кластер на рис. 2. Показатель цитируемости определяется значением “Avg. norm. citations”

[3] van Eck N.J., Waltman L. Manual for VOSviewer version 1.6.20. URL: https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf (accessed 5 August.2024).

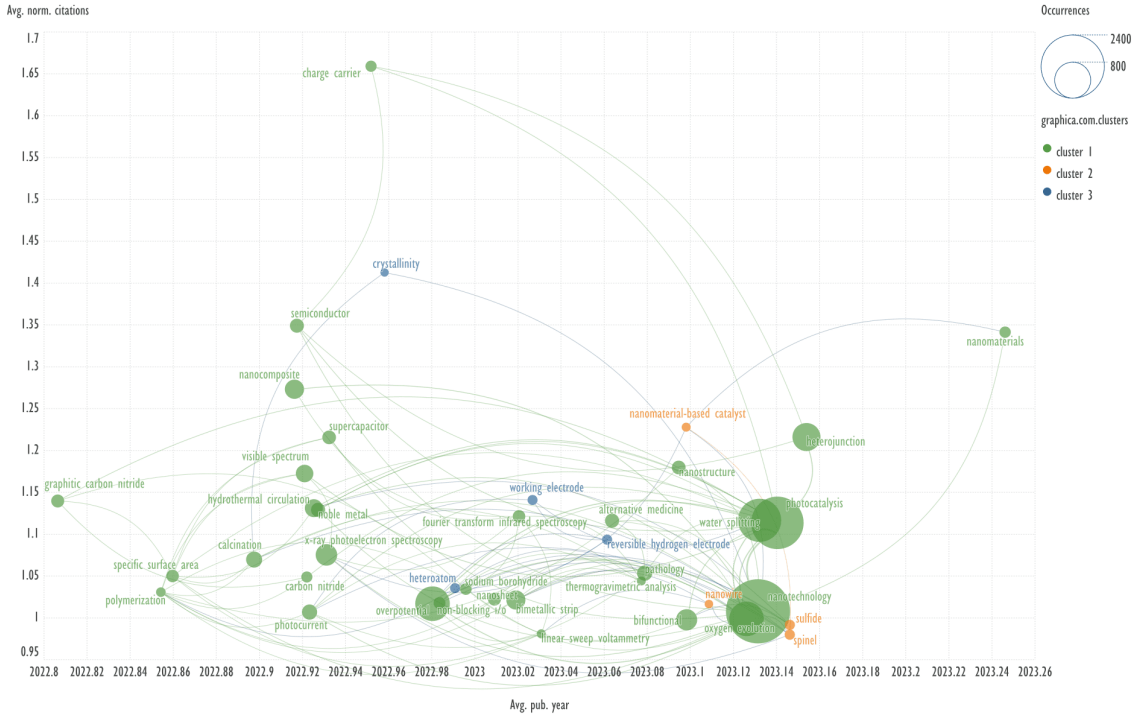

Fig. 4. Visualization of the co-occurrence network of
the 40 most cited “Fields of Study” terms belonging to the green cluster of Fig. 2.
The citation score is given by the value of “Avg. norm. citations”

Рис. 4. Визуализация сети совместной встречаемости
40 наиболее цитируемых терминов “Fields of Study”,
входящих в зеленый кластер на рис. 2.
Показатель цитируемости определяется значением “Avg. norm. citations”

Clustering of articles

The topics of publications presented in the journal can be described not only by keywords, but also by clustering the articles themselves. The Bibliographic Coupling methodology is used for this purpose.

The BibliographicCouplingDocuments.json file placed in the archive attached to this article allows you to open it in VOSviewer Online and see in detail the publication clustering network obtained by Bibliographic Coupling on cited Documents. The significant advantage of online browsing is the possibility to see in the tooltip not only general characteristics of the publication network, but also detailed data of the selected publication – its title, where and when it was published, with which documents it is included in the cluster. A copy of the JSON file view screen in VOSviewer Online is shown in Fig. 5.

Documents are organized into eight clusters, general characteristics of the document network: Items: 500; Links: 20,193; Total link strength: 39,785.

The problem with using “Bibliographic coupling” is that many publications nowadays cite very large reference lists. For example, in the data exported from The Lens used in our work, the average number of citations per publication is 575,389/10,928=52.6. Much of these citations are usually in the Introduction section and only indirectly reflect the content of the article. In my opinion, it is currently more appropriate to use tags (index keywords/labels) assigned by the abstract database platform to a particular publication to assess the similarity of publications.

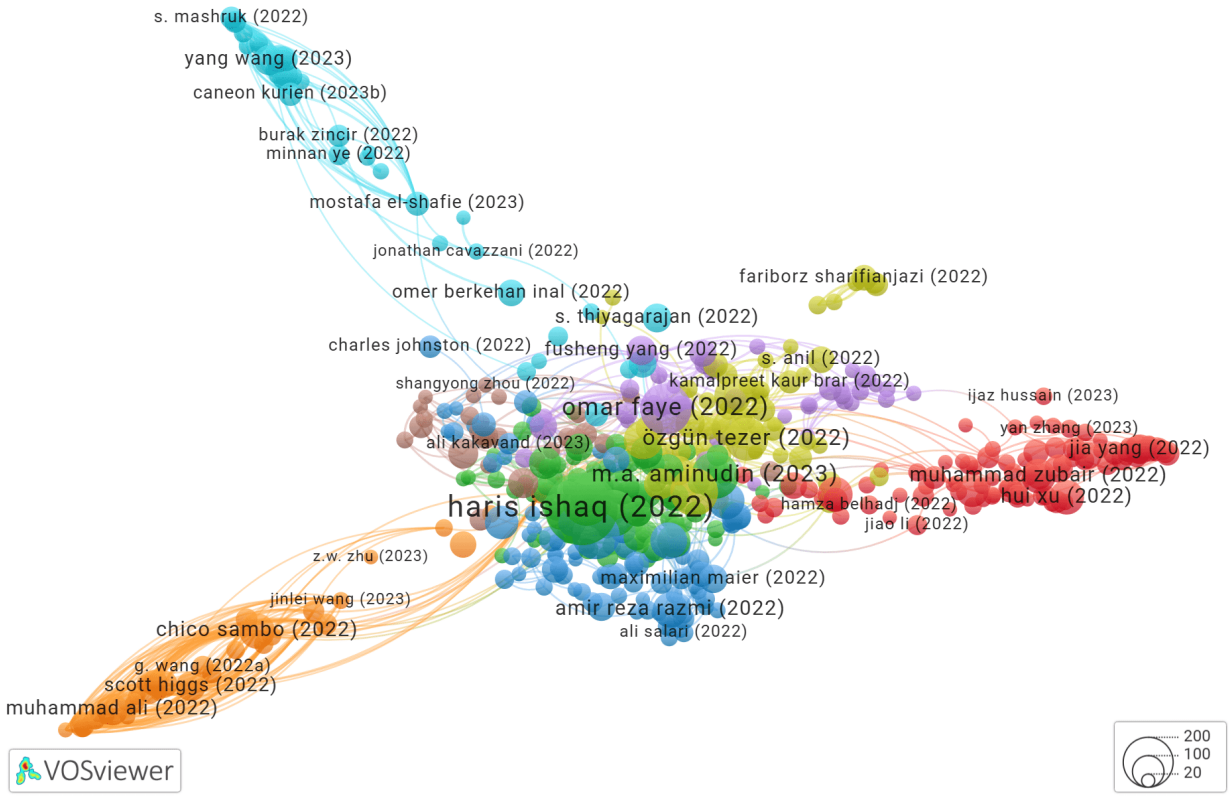

Fig. 5. Clustering of documents based on their Bibliographic Coupling by cited Documents

Рис. 5. Кластеризация документов на основе их библиографической связи по цитируемым документам

In The Lens system, such labels are placed in the “Fields of Study” column. In the data used in this paper, there was an average of 190,972/10,972=17.4 “Fields of Study” terms per publication. This is quite sufficient to compare the proximity of the topics of publications, all the more that the terms used in the “Fields of Study” field have the same spelling and therefore do not require additional normalization, and the field itself does not contain empty lines.

Based on the above, in the data table from The Lens analyzed by VOSviewer, the “Fields of Study” column was renamed to “References” and a similar analysis was performed to determine Bibliographic Coupling, the results of which are shown in Fig. 6.

The distinctive feature of this graph is the small number of clusters (2) when using VOSviewer with default parameters.

The following procedure was carried out to determine the current topics of the two clusters presented above:

– twenty records with the highest normalized citations were selected in each of the clusters; a map file expatriated from VOSviewer was used;

– for each publication of the twenty records, the corresponding “Fields of Study” terms were defined, merging was performed by DOI field using the INNER JOIN operator with the data of the table on which the network of publications was built, presented in Fig. 6;

– for each of the clusters, the co-occurrence of the four “Fields of Study” terms was plotted in the form of alluvial diagrams, as shown in Fig. 7 and Fig. 8.

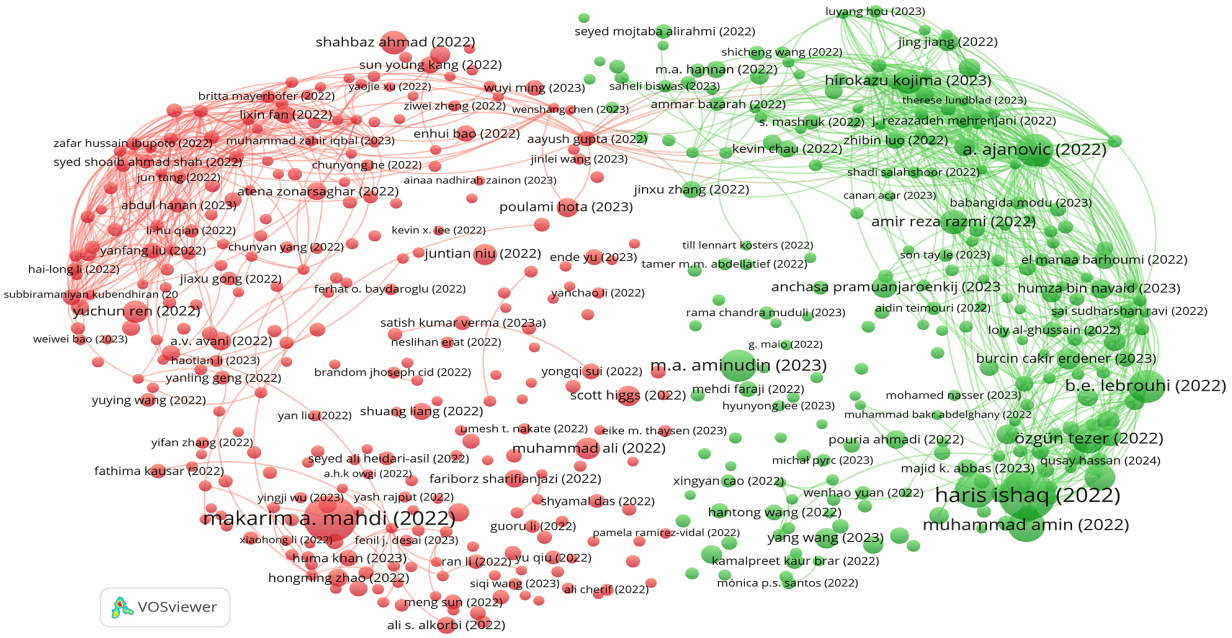

Fig. 6. Clustering of documents based on their similarity according to the “Fields of Study” field data

Рис. 6. Кластеризация документов на основе их сходства в соответствии с данными поля “Fields of Study”

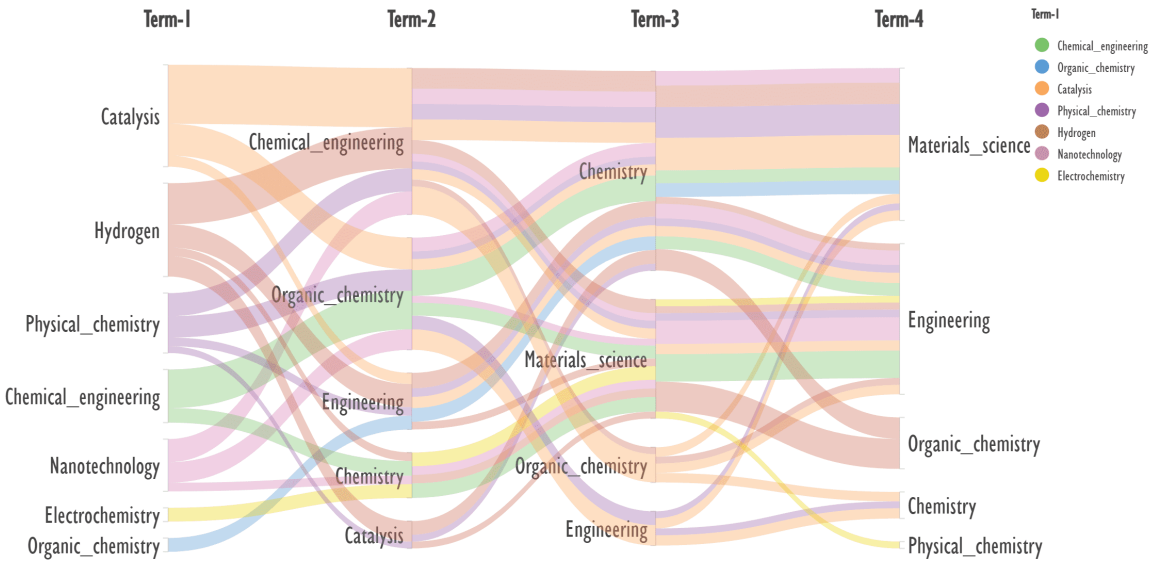

Fig. 7. Alluvial diagram for the 45 most frequently co-occurring four “Fields of Study” terms of the first cluster.
Colors are chosen according to the first term in the diagram

Рис. 7. Аллювиальная диаграмма для 45 наиболее часто встречающихся четырех терминов “Fields of Study” первого кластера.
Цвета выбраны в соответствии с первым термином на диаграмме

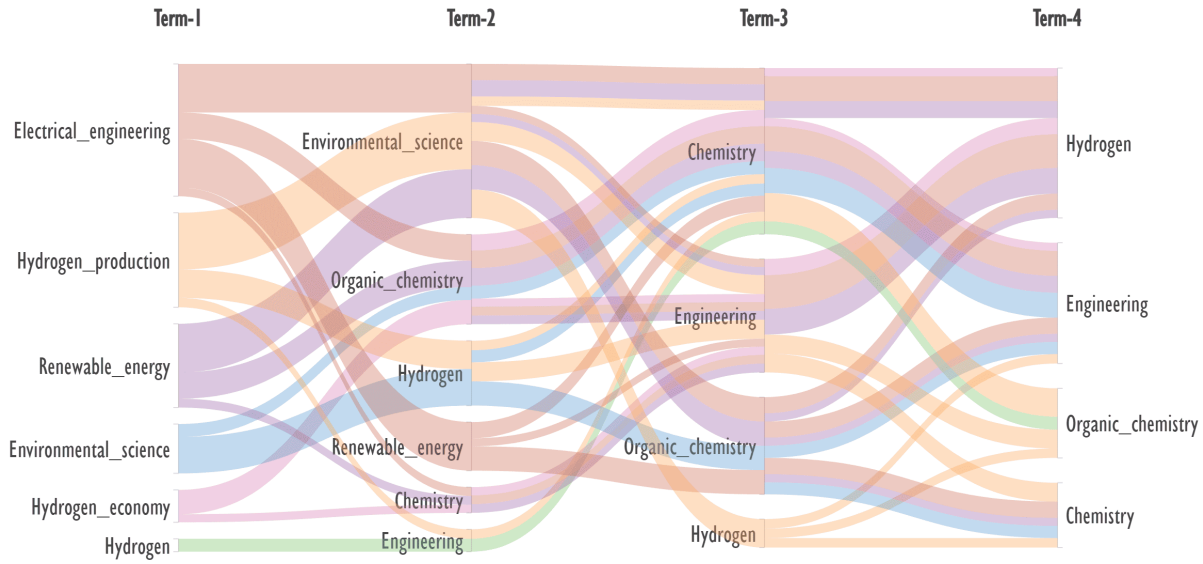

Fig. 8. Alluvial diagram for the 45 most frequently co-occurring four "Fields of Study" terms of the second cluster.
Colors are chosen according to the first term in the diagram

Рис. 8. Аллювиальная диаграмма для 45 наиболее часто совпадающих четырех терминов "Fields of Study" второго кластера.
Цвета выбраны в соответствии с первым термином на диаграмме

A total of 372 "Fields of Study" terms were contained in the 20 records of the first cluster and 460 in the second cluster. Of these, the first cluster contained 154 unique and the second cluster contained 135 unique "Fields of Study" terms.

The co-occurrence of terms was determined using FP-growth algorithm with parameters -s40m4n4. The s40 is a very high value of this parameter, which indicates the significant similarity of "Fields of Study" terms in different bibliometric records. This fact can be explained by the fact that all articles are published in one journal, the subject matter of which is significantly limited.

For the articles in the first cluster, 57 co-occurrence results were obtained for the four Fields of Study terms. For the articles in the second cluster, 199 co-occurrence results were obtained for the four Fields of Study terms.

The alluvial diagrams below were constructed using the Scimago Graphica program for the 45 most frequently occurring four terms.

Fig. 7 shows the most pronounced co-occurrence of terms: Catalysis, Chemical engineering, Chemistry, Materials science. The terms: Catalysis, Hydrogen, Physical chemistry, Chemical engineering, Nanotechnology, Electrochemistry, Organic chemistry, Materials science are the most reflective of the subject matter of this chart.

Fig. 8 shows the most pronounced co-occurrence of terms: Electrical engineering, Environmental science, Chemistry, Hydrogen. The terms: Electrical engineering, Hydrogen production, Renewable energy, Environmental science, Hydrogen economy, Hydrogen, Engineering are the most reflective of the subject matter of this chart.

The color selection options for each of the term layers are presented as four interactive web pages for both figures, available in the attached archive.

The “Fields of Study” terms can serve as filters when searching for information on The Lens platform. Knowing the co-occurrence of the terms significantly narrows down the exported sample of bibliometric records from The Lens platform, thus speeding up the search for the desired information.

Conclusions

The possibility of combined use of fields of bibliometric records of the abstract databases ScienceDirect and The Lens, complementing each other, is shown. For example, in The Lens the field of keywords is poorly filled, and in ScienceDirect records there are no fields of citations and reference lists. The “Fields of Study” field of The Lens platform can be interpreted as system keywords similar to Index Keywords in Scopus.

Given that widely used bibliometric analysis programs such as VOSviewer can use The Lens data, the easiest way to combine the data is to populate the Keywords and Abstract fields of the data exported from The Lens with the “AB” (abstract) and “KW” (keywords) fields of the RIS files exported from ScienceDirect. More generally, a merged data table in Scopus CSV format can be created by renaming the relevant fields and converting separator characters between terms. It is advisable to merge data from tables by DOI.

The feasibility of consistent use of VOSviewer and Scimago Graphica programs for more complete visualization of the results of bibliometric analysis is demonstrated. The feasibility of using alluvial diagram to map the co-occurrence of, for example, four keywords and to map the co-occurrence network of keywords in coordinates of average publication time and average normalized citation is shown.

The “Fields of Study” data, being normalized terms, provide good opportunities to analyze the topics of publications.

By analyzing the “Fields of Study” bibliometric data of International Journal of Hydrogen Energy for 2022–2024, two dominant publication themes are identified, which can be described in terms of: “Catalysis, Hydrogen, Physical chemistry, Chemical engineering, Nanotechnology, Electrochemistry, Organic chemistry, Materials science” and “Electrical engineering, Hydrogen production, Renewable energy, Environmental science, Hydrogen economy, Hydrogen, Engineering”.

Possible follow-up study: comparing different record grouping approaches – bibliographic coupling variants and Gibbs Sampling algorithm for Dirichlet Multinomial Mixture model.

Author contributions

Boris N. Chigarev – idea, data collection and analysis, text preparation and editing.

Conflict of interests

The author declares no conflict of interests.

References

1. Zhang D., Jiang M., Li G., Tang Y. An advanced bibliometric analysis and future research insights on safety of hydrogen energy. *Journal of Energy Storage.* 2024. Vol. 77. P. 109833. https://doi.org/10.1016/j.est.2023.109833

2. Wei R., Lan J., Lian L. et al. A bibliometric study on research trends in hydrogen safety. *Process Safety and Environmental Protection.* 2022. Vol. 159. P. 1064–1081. https://doi.org/10.1016/j.psep.2022.01.078

3. Kar S.K., Harichandan S., Roy B. Bibliometric analysis of the research on hydrogen economy: An analysis of current findings and roadmap ahead. *International Journal of Hydrogen Energy.* 2022. Vol. 47, No. 20. P. 10803–10824. https://doi.org/10.1016/j.ijhydene.2022.01.137

4. Chen X., Zhang Y., Xu S., Dong F. Bibliometric analysis for research trends and hotspots in heat and mass transfer and its management of proton exchange membrane fuel cells. *Applied Energy.* 2023. Vol. 333. P. 120611. https://doi.org/10.1016/j.apenergy.2022.120611

5. Boukhennoufa I., Zhai X., Utti V. et al. Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review. *Biomedical Signal Processing and Control.* 2022. Vol. 71. P. 103197. https://doi.org/10.1016/j.bspc.2021.103197

6. Conceição A.M.D., Carvalho M.B.S.D., Bastos V.H.D.V. et al. Physiotherapeutic approach to Attention Deficit Hyperactivity Disorder (ADHD). *Research, Society and Development.* 2022. Vol. 11, No. 15. P. e268111537138. https://doi.org/10.33448/rsd-v11i15.37138

7. Rodríguez-Cáceres A., Sánchez-Vera M.A., Alfonso Mora M. et al. Relationship between screen exposure, sedentary behavior and musculoskeletal pain in adolescents: a systematic review [Relación entre la exposición a pantallas, el comportamiento sedentario y el dolor musculoesquelético en adolescentes: revisión sistemática]. *Retos.* 2023. Vol. 50. P. 1064–1070. https://doi.org/10.47197/retos.v50.99865

8. Armenta E.E., Armenta J.M., Coronado M.A., Ayala J.R. Systems and methods for citrus essential oil isolation: a patent review. *Journal of Essential Oil Research.* 2024. Vol. 36, No. 2. P. 94–104. https://doi.org/10.1080/10412905.2023.2300999

9. Pflugfelder N.S. Knowledge management as a driver of performance in ambulatory healthcare – a systematic literature review through an intellectual capital lens. *Journal of Intellectual Capital.* 2021. Vol. 22, No. 2. P. 403–432. https://doi.org/10.1108/JIC-02-2020-0068

10. Magaiza G., Muchaku S. Curriculum enablement and posthumanism: Pathways for creating and implementing a community development curriculum. *Journal of Curriculum Studies Research.* 2023. Vol. 5, No. 1. P. 159–170. https://doi.org/10.46303/jcsr.2023.12

11. Herman H., Kusbaryanto K. The effect of family therapy toward prevention of violent behavior in child: a literature review. *Jurnal Keperawatan Respati Yogyakarta.* 2020. Vol. 7, No. 2. P. 117. https://doi.org/10.35842/jkry.v7i2.530

12. Ortega J.L., Delgado-Quirós L. The indexation of retracted literature in seven principal scholarly databases: a coverage comparison of dimensions, OpenAlex, PubMed, Scilit, Scopus, The Lens and Web of Science. *Scientometrics.* 2024. Vol. 129, No. 7. P. 3769–3785. https://doi.org/10.1007/s11192-024-05034-y

13. Delgado-Quirós L., Ortega J.L. Completeness degree of publication metadata in eight free-access scholarly databases. *Quantitative Science Studies.* 2024. Vol. 5, No. 1. P. 31–49. https://doi.org/10.1162/qss_a_00286

14. Delgado-Quirós L., Aguillo I.F., Martín-Martín A. et al. Why are these publications missing? Uncovering the reasons behind the exclusion of documents in free-access scholarly databases. *Journal of*

the Association for Information Science and Technology. 2024. Vol. 75, No. 1. P. 43–58. https://doi.org/10.1002/asi.24839

15. Delgado Quirós L.J., Ortega J.L. Comparing bibliographic descriptions in seven free-access databases. In: *27th International Conference on Science, Technology and Innovation Indicators (STI 2023),* Leiden, Netherlands, 27–29 September 2023. https://doi.org/10.55835/6436c590b3340c364be5b2c7

16. Van Eck N.J., Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics.* 2010. Vol. 84, No. 2. P. 523–538. https://doi.org/10.1007/s11192-009-0146-3

17. Hassan-Montero Y., De-Moya-Anegón F., Guerrero-Bote V.P. SCImago Graphica: a new tool for exploring and visually communicating data. *Profesional de la información.* 2022. Vol. 31, No. 5. P. e310502. https://doi.org/10.3145/epi.2022.sep.02

18. Borgelt C. Frequent item set mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* 2012. Vol. 2, No. 6. P. 437–456. https://doi.org/10.1002/widm.1074

Information about the author

Boris N. Chigarev – Cand. Sci. (Phys.-Math.), Leading Engineer on Scientific and Technical Information, Oil and Gas Research Institute, Russian Academy of Sciences, Moscow, Russia; https://orcid.org/0000-0001-9903-2800; e-mail: bchigarev@ipng.ru

Received 4 September 2024

Accepted 1 November 2024

ЭНЕРГЕТИЧЕСКИЕ РЕСУРСЫ ЗЕМНОЙ КОРЫ: ВЫЗОВЫ СОВРЕМЕННОЙ ЭКОНОМИКИ

Оригинальная статья
УДК [303.6+303.7]:001.8
https://doi.org/10.29222/ipng.2078-5712.2025.01

Выявление актуальных тем исследований на основе объединенных библиометрических данных публикаций International Journal of Hydrogen Energy за 2022–2024 гг. платформ The Lens и ScienceDirect

Б.Н. Чигарев ✉
Институт проблем нефти и газа РАН, Россия, 119333, Москва, ул. Губкина, д. 3

Аннотация. *Актуальность.* Водороду отводится роль чистого энергоносителя в рамках энергетического перехода. *Цель работы.* Выявление актуальных тем исследований на основе объединенных библиометрических данных публикаций International Journal of Hydrogen Energy за 2022–2024 гг., собранных на платформах The Lens и ScienceDirect. *Материалы и методы.* Для анализа было экспортировано 10 928 библиометрических записей из базы данных The Lens и 10 857 записей из ScienceDirect. Кластеризация ключевых слов и визуальный анализ данных проводились с помощью следующих программ и алгоритмов: VOSviewer, Scimago Graphica, Inkscape, утилита FP-growth. *Результаты.* Исследование показало возможность объединения библиометрических записей с открытых платформ ScienceDirect и The Lens, которые дополняют друг друга. Данные “Fields of Study” из The Lens могут быть рационально интерпретированы как системные ключевые слова, аналогичные Scopus Index Keywords. Показана возможность использования данных “Fields of Study” в методе, аналогичном библиографической связи. Продемонстрирована практическая целесообразность использования Аллювиальной диаграммы для отображения совместной встречаемости четырех терминов. В исследовании также отмечена целесообразность совместного использования программ VOSviewer и Scimago Graphica для комплексной визуализации результатов библиометрического анализа. Исследование выявило две доминирующие темы публикаций в International Journal of Hydrogen Energy на 2022–2024 гг., которые могут быть описаны в терминах: “Catalysis, Hydrogen, Physical chemistry, Chemical engineering, Nanotechnology, Electrochemistry, Organic chemistry, Materials science” и “Electrical engineering, Hydrogen production, Renewable energy, Environmental science, Hydrogen economy, Hydrogen, Engineering”. *Выводы.* Совместное использование метаданных различных открытых реферативных баз данных позволяет компенсировать неполное их представление.

Ключевые слова: International Journal of Hydrogen Energy, The Lens, ScienceDirect, VOSviewer, Scimago Graphica, FP-growth

Финансирование: работа выполнена в рамках государственного задания ИПНГ РАН (тема № 125021302095-2).

✉ Чигарев Борис Николаевич, bchigarev@ipng.ru

Для цитирования: *Чигарев Б.Н.* Выявление актуальных тем исследований на основе объединенных библиометрических данных публикаций International Journal of Hydrogen Energy за 2022–2024 гг. платформ The Lens и ScienceDirect // Актуальные проблемы нефти и газа. 2025. (На англ. яз.). https://doi.org/10.29222/ipng.2078-5712.2025.01

Вклад автора

Б.Н. Чигарев – концепция, сбор и анализ данных, подготовка и редактирование текста.

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Список источников

1. *Zhang D., Jiang M., Li G., Tang Y.* An advanced bibliometric analysis and future research insights on safety of hydrogen energy // Journal of Energy Storage. 2024. Vol. 77. P. 109833. https://doi.org/10.1016/j.est.2023.109833

2. *Wei R., Lan J., Lian L.* et al. A bibliometric study on research trends in hydrogen safety // Process Safety and Environmental Protection. 2022. Vol. 159. P. 1064–1081. https://doi.org/10.1016/j.psep.2022.01.078

3. *Kar S.K., Harichandan S., Roy B.* Bibliometric analysis of the research on hydrogen economy: An analysis of current findings and roadmap ahead // International Journal of Hydrogen Energy. 2022. Vol. 47, No. 20. P. 10803–10824. https://doi.org/10.1016/j.ijhydene.2022.01.137

4. *Chen X., Zhang Y., Xu S., Dong F.* Bibliometric analysis for research trends and hotspots in heat and mass transfer and its management of proton exchange membrane fuel cells // Applied Energy. 2023. Vol. 333. P. 120611. https://doi.org/10.1016/j.apenergy.2022.120611

5. *Boukhennoufa I., Zhai X., Utti V.* et al. Wearable sensors and machine learning in post-stroke rehabilitation assessment: A systematic review // Biomedical Signal Processing and Control. 2022. Vol. 71. P. 103197. https://doi.org/10.1016/j.bspc.2021.103197

6. *Conceição A.M.*D., *Carvalho M.B.*S.*D.*, *Bastos V.H.D.V.* et al. Physiotherapeutic approach to Attention Deficit Hyperactivity Disorder (ADHD) // Research, Society and Development. 2022. Vol. 11, No. 15. P. e268111537138. https://doi.org/10.33448/rsd-v11i15.37138

7. *Rodríguez-Cáceres A.,* Sánchez-*Vera M.A.*, *Alfonso Mora M.* et al. Relationship between screen exposure, sedentary behavior and musculoskeletal pain in adolescents: a systematic review [Relación entre la exposición a pantallas, el comportamiento sedentario y el dolor musculoesquelético en adolescentes: revisión sistemática] // Retos. 2023. Vol. 50. P. 1064–1070. https://doi.org/10.47197/retos.v50.99865

8. *Armenta E.E.*, *Armenta J.M.*, *Coronado M.A.*, *Ayala J.R.* Systems and methods for citrus essential oil isolation: a patent review // Journal of Essential Oil Research. 2024. Vol. 36, No. 2. P. 94–104. https://doi.org/10.1080/10412905.2023.2300999

9. *Pflugfelder N.S.* Knowledge management as a driver of performance in ambulatory healthcare – a systematic literature review through an intellectual capital lens // Journal of Intellectual Capital. 2021. Vol. 22, No. 2. P. 403–432. https://doi.org/10.1108/JIC-02-2020-0068

10. *Magaiza G., Muchaku S.* Curriculum enablement and posthumanism: Pathways for creating and implementing a community development curriculum // Journal of Curriculum Studies Research. 2023. Vol. 5, No. 1. P. 159–170. https://doi.org/10.46303/jcsr.2023.12

11. *Herman H., Kusbaryanto K.* The effect of family therapy toward prevention of violent behavior in child: a literature review // Jurnal Keperawatan Respati Yogyakarta. 2020. Vol. 7, No. 2. P. 117. https://doi.org/10.35842/jkry.v7i2.530

12. *Ortega J.L.*, *Delgado-Quirós L.* The indexation of retracted literature in seven principal scholarly databases: a coverage comparison of dimensions, OpenAlex, PubMed, Scilit, Scopus, The Lens and Web of Science // Scientometrics. 2024. Vol. 129, No. 7. P. 3769–3785. https://doi.org/10.1007/s11192-024-05034-y

13. *Delgado-Quirós L., Ortega J.L.* Completeness degree of publication metadata in eight free-access scholarly databases // Quantitative Science Studies. 2024. Vol. 5, No. 1. P. 31–49. https://doi.org/10.1162/qss_a_00286

14. *Delgado-Quirós L., Aguillo I.F.*, Martín-Martín A. et al. Why are these publications missing? Uncovering the reasons behind the exclusion of documents in free-access scholarly databases // Journal of the Association for Information Science and Technology. 2024. Vol. 75, No. 1. P. 43–58. https://doi.org/10.1002/asi.24839

15. *Delgado Quirós L.J.*, *Ortega J.L.* Comparing bibliographic descriptions in seven free-access databases // 27th International Conference on Science, Technology and Innovation Indicators (STI 2023), Leiden, Netherlands, 27–29 September 2023. https://doi.org/10.55835/6436c590b3340c364be5b2c7

16. *Van Eck N.J., Waltman L.* Software survey: VOSviewer, a computer program for bibliometric mapping // Scientometrics. 2010. Vol. 84, No. 2. P. 523–538. https://doi.org/10.1007/s11192-009-0146-3

17. *Hassan-Montero Y., De-Moya-Anegón F., Guerrero-Bote V.P.* SCImago Graphica: a new tool for exploring and visually communicating data // Profesional de la información. 2022. Vol. 31, No. 5. P. e310502. https://doi.org/10.3145/epi.2022.sep.02

18. *Borgelt C.* Frequent item set mining // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2012. Vol. 2, No. 6. P. 437–456. https://doi.org/10.1002/widm.1074

Информация об авторе

Борис Николаевич Чигарев – к.ф.-м.н., ведущий инженер по научно-технической информации, Институт проблем нефти и газа РАН, Москва, Россия; SPIN-код: 7610-8398, https://orcid.org/0000-0001-9903-2800; e-mail: bchigarev@ipng.ru

Поступила в редакцию 04.09.2024
Принята к публикации 01.11.2024