Saudi Arabia Energy Transition in the Context of Scholarly Publications

Boris Chigarev
Oil and Gas Research Institute, Russian Academy of Sciences, Moscow
E-mail: bchigarev@ipng.ru
https://orcid.org/0000-0001-9903-2800
https://figshare.com/authors/Boris_Chigarev/6474086

Abstract

The article is devoted to a bibliometric analysis of scientific publications reflecting trends in scientific publications related to the energy transition of leading universities in Saudi Arabia. The universities were selected based on the presence of joint publications with the Saudi Arabian Oil Company, they are: King Fahd University of Petroleum and Minerals, King Abdullah University of Science and Technology, and King Saud University. The importance of studying the experience of Saudi Arabia for Russia in the context of the energy transition was highlighted. The analysis was based on Scopus platform data related to Subject Areas: Physical Sciences. Scientists from Saudi Arabian universities are increasingly publishing with co-authors from China, India, Egypt, Pakistan, and Malaysia, while maintaining high interaction with co-authors from the United States, South Korea, and the United Kingdom. Authors from the Russian Federation publish poorly with co-authors from Saudi Arabia. The trends in the topics of scientific publications were investigated, and their connection with the renewable energy topic was shown. The VOSviewer software was used to determine the main research areas using the cluster analysis method based on the co-occurrence of key terms. The analysis was done for both Author's keywords and Index keywords of the Scopus system. Graphical representation of the final choice of terms to describe research trends was proposed.

Keywords: Saudi Arabia, energy transition, bibliometric analysis, Scopus, VOSviewer, research trends

Introduction

In the context of growing contacts between Russia and Saudi Arabia in the framework of OPEC+ decisions, it is reasonable to consider the priorities and trends in scientific research of this partner.¹

Since the 2010s, Russia has made significant progress in developing closer economic, political and security relations with the Gulf countries [1].

The initiative of national leaders to create a Charter for Cooperation among oil-producing countries will promote dialogue and cooperation among oil-producing countries at the ministerial and technical levels for the benefit of both oil-producing and consuming countries and the world economy.²

The Gulf countries need to diversify their economies away from dependence on oil and gas, which involves creating a next-generation energy market and improving access to markets in China, India and other developing countries. [2]. This trend can be traced not only at the political level, but also at the level of research cooperation, which can be assessed by analyzing bibliometric data on scientific publications.

Fossil fuels are the main source of economic income in the Gulf countries. On the other hand, climate change is closely linked to the use of fossil energy, and this has motivated the Gulf countries to look for alternative solutions, such as renewable energy technologies to comply with the Paris Agreement [3].

Countries have unequal natural resources, economic power and capacity. Achieving the Sustainable Development Goals is about making the best use of a country's resources and capabilities [4].

¹ Saudi Arabia and Russia Agree on Second Phase of Economic, Scientific, and Technical Cooperation

© 2021 by the author(s). Distributed under a Creative Commons CC BY license.
Economies that focus on hydrocarbons will have to determine the prospects and opportunities for a more sustainable future. The enormous untapped renewable energy (RE) potential in the Gulf Cooperation Council (GCC) countries, combined with the growing liquidity of money in new climate regimes, is paving the way for RE-based technologies [5].

The economy of Saudi Arabia as well as that of Russia is significantly dependent on hydrocarbon exports, and both countries have to adapt to the challenges of the energy transition. Therefore, it is of particular interest to identify changes in the direction of scientific research in Saudi Arabia in recent years in the field of natural sciences (or Physical Sciences as by Scopus).

Note: According to the categories of the Scopus abstract database, the subject area "Physical Sciences" includes: Chemical Engineering, Chemistry, Computer Science, Earth and Planetary Sciences, Energy, Engineering, Environmental Sciences, Materials Sciences, Mathematics, and Physics and Astronomy subject areas.

The objective of the study, the reasoning for the choice of bibliometric data and the main results of their analysis.

The aim of this paper is to identify trends in scientific research in Saudi Arabia in the context of the energy transition, based on an analysis of bibliometric data presented in abstract databases.

It is advisable to analyze trends in scientific research using the metadata of abstract databases, such as Web of Science, Scopus, The Lens, and Dimensions.

The Lens, and Dimensions abstract databases have a pronounced specificity in the structure of exported data, so it is advisable to use them in separate bibliometric studies beyond the scope of this article.

The data of the Scopus platform were used in this work. The choice of Scopus is due to the fact that the Index Keywords of the Scopus platform describe the publication topic in much more detail than the Keywords Plus of the WoS platform. The main emphasis of the article is on the analysis of changes in research topics, which are best described by keywords, both the author's own and those offered by the abstract base itself.

**Major Trends in Scientific Research at Saudi Aramco and Related Universities**

It is the source of about 80% of revenues to the Saudi budget. Therefore, it is of interest to analyze the direction of development of scientific research of this company in the context of the energy transition. The company's website announces the main priority of the company's development in these conditions - circular economy. "In 2020 Aramco set itself a path to become a leading circular economy company" [3]. The main tasks for the realization of the declared goals can be traced through scientific publications. However, the main sources of publications are universities. Therefore, it is useful to identify which universities in Saudi Arabia most frequently publish with Saudi Aramco and include the metadata of their articles.

**Selection of the abstract database**

On the Scopus platform, you can detect an organization identifier and then use it to filter publications whose authors are affiliated with that organization. Affiliation name: Saudi Arabian Oil Company, Scopus affiliation ID: 60025307.

Scopus indexed 6,093 documents for all years with LIMIT-TO (AF-ID, "Saudi Arabian Oil Company" 60025307) as of 2 September 2021.

But WoS indexes publications only since 1975, so we impose the PUBYEAR > 1974 restriction for comparison.

I got 6,768 results for AF-ID ("Saudi Arabian Oil Company" 60025307) AND PUBYEAR > 1974 of them Conference Paper (4,498) and Article (1,979).

---

On the WoS platform I get: 1,973 results for OG=(Saudi Aramco) Indexes=SCI-EXPANDED, ESCI Timespan=1975-2021 (In the classic WoS interface)

Thus, the number of articles indexed in Scopus and WoS for the period 1975-2021 is almost the same. The advantage of Scopus is the large number of indexed conference proceedings and the greater number of Index Keywords compared to the Keywords Plus of the WoS platform.

To understand which topics are of most interest to Saudi-Aramco, let's analyze Index Keywords of the 2000 most relevant conference papers obtained by the query: AF-ID (“Saudi Arabian Oil Company” 60025307) AND PUBYEAR > 2011 AND (LIMIT-TO (DOCTYPE, “cp”). The predominant Index Keywords is shown in Tab. 1.

Table 1. Top 40 Index Keywords for 2000 most relevance Conference Paper by Saudi Aramco in 2011-2021 (as by 2 sent 2021)

<table>
<thead>
<tr>
<th>Index Keywords</th>
<th>N</th>
<th>Index Keywords</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>gasoline</td>
<td>429</td>
<td>efficiency</td>
<td>93</td>
</tr>
<tr>
<td>oil wells</td>
<td>338</td>
<td>well stimulation</td>
<td>92</td>
</tr>
<tr>
<td>petroleum reservoir evaluation</td>
<td>286</td>
<td>well testing</td>
<td>92</td>
</tr>
<tr>
<td>petroleum reservoirs</td>
<td>269</td>
<td>hydraulic fracturing</td>
<td>91</td>
</tr>
<tr>
<td>infill drilling</td>
<td>254</td>
<td>hydrocarbons</td>
<td>90</td>
</tr>
<tr>
<td>boreholes</td>
<td>238</td>
<td>cost effectiveness</td>
<td>87</td>
</tr>
<tr>
<td>petroleum reservoir engineering</td>
<td>176</td>
<td>drilling fluids</td>
<td>87</td>
</tr>
<tr>
<td>gas industry</td>
<td>171</td>
<td>sulfur compounds</td>
<td>85</td>
</tr>
<tr>
<td>horizontal wells</td>
<td>167</td>
<td>crude oil</td>
<td>84</td>
</tr>
<tr>
<td>fracture</td>
<td>143</td>
<td>reservoir management</td>
<td>84</td>
</tr>
<tr>
<td>gases</td>
<td>134</td>
<td>oil well logging</td>
<td>82</td>
</tr>
<tr>
<td>oil field equipment</td>
<td>129</td>
<td>oil field development</td>
<td>81</td>
</tr>
<tr>
<td>offshore oil well production</td>
<td>125</td>
<td>floods</td>
<td>78</td>
</tr>
<tr>
<td>carbonation</td>
<td>118</td>
<td>coiled tubing</td>
<td>75</td>
</tr>
<tr>
<td>seismology</td>
<td>115</td>
<td>well completion</td>
<td>74</td>
</tr>
<tr>
<td>oil well flooding</td>
<td>114</td>
<td>secondary recovery</td>
<td>70</td>
</tr>
<tr>
<td>petroleum engineering</td>
<td>102</td>
<td>petroleum prospecting</td>
<td>69</td>
</tr>
<tr>
<td>enhanced recovery</td>
<td>100</td>
<td>natural gas wells</td>
<td>66</td>
</tr>
<tr>
<td>carbonate reservoir</td>
<td>99</td>
<td>porosity</td>
<td>65</td>
</tr>
<tr>
<td>oil fields</td>
<td>97</td>
<td>petroleum industry</td>
<td>64</td>
</tr>
</tbody>
</table>

Oil and gas companies stick to classic topics, even in conference proceedings, which reflect new topics more quickly than peer-reviewed articles.

Usually, universities develop new research areas. Therefore, it is useful to identify those that most often publish joint articles with Saudi Aramco, which can be done by analyzing the affiliation of the co-authors.

To identify such universities, I query the Scopus database with the following filters: years — 2012-2021; language — English; article type — journal article; Subject Areas — Physical Sciences; affiliation — the Saudi Arabian Oil Company organization, the following affiliations for co-authors in the format Affiliation number of publications are obtained:

- Saudi Arabian Oil Company 842
- King Fahd University of Petroleum and Minerals 156
- King Abdullah University of Science and Technology 112
- Korea Advanced Institute of Science and Technology 26
- Massachusetts Institute of Technology 19
For further analysis, we will limit it to the Arab universities. Below is a summary of four institutions (Saudi Arabian Oil Company, King Fahd University of Petroleum and Minerals, King Abdullah University of Science and Technology, and King Saud University), whose publication activities will be discussed later. A summary of the Subject Areas for each of the organizations is shown in Fig. 1.

**Fig. 1.** Main Subject Areas of the four organizations Saudi Arabian Oil Company, King Fahd University of Petroleum and Minerals, King Abdullah University of Science and Technology, King Saud University. Data exported from the Affiliation details section of Scopus.

Fig. 1 shows that the first three organizations are mostly published in natural sciences disciplines, so the number of joint publications of King Fahd University of Petroleum and Minerals, King Abdullah University...
of Science with Technology and Saudi Arabian Oil Company is greater than that of King Saud University, which pays more attention to research in the field of medicine.

But King Saud University is great, and the total number of publications on the topic of interest to us is also large:

- Engineering 12010 publications
- Chemistry 10570
- Computer Science 8921
- Energy 2656

Whereas King Fahd University of Petroleum has the following total number of publications in these disciplines:

- Engineering 10454
- Computer Science 5028
- Chemistry 4510
- Energy 3660

Thus, it is advisable to take into account the publications of King Saud University.

Creation the main query to the abstract database Scopus, the data on which will be further exploited

The assignment of a publication to a particular institution was determined by the affiliation ID in Scopus.

The affiliation ID for the four institutions will be expressed in the query as follows: (AF-ID (60025307) OR AF-ID (60009506) OR AF-ID (60092945) OR AF-ID (60013183))

The Scopus natural science area is listed as Subject Areas: Physical Sciences, and includes the following subjects:

- Chemical Engineering (CENG)
- Chemistry (CHEM)
- Computer Science (COMP)
- Earth and Planetary Sciences (EART)
- Energy (ENER)
- Engineering (ENGI)
- Environmental Science (ENVI)
- Materials Science (MATE)
- Mathematics (MATH)
- Physics and Astronomy (PHYS)

In a query to the database, it will be as follows: SUBJAREA (CENG OR CHEM OR COMP OR EART OR ENER OR ENGI OR ENVI OR MATE OR MATH OR PHYS)

Additional data filtering was used: type of publication article — LIMIT-TO (DOCTYPE, "ar"), in a journal — LIMIT-TO (SRCTYPE, "j"), in English — LIMIT-TO (LANGUAGE, "English") published after 2011 — PUBYEAR > 2011

The final request: "Your query : ((PUBYEAR > 2011) AND SUBJAREA(CENG OR CHEM OR COMP OR EART OR ENER OR ENGI OR ENVI OR MATE OR MATH OR PHYS) AND (AF-ID("Saudi Arabian Oil Company" 60025307) OR AF-ID("King Fahd University of Petroleum and Minerals" 60009506) OR AF-ID("King Abdullah University of Science and Technology" 60092945) OR AF-ID("King Saud University" 60013183)) AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (LANGUAGE,"English")) AND (LIMIT-TO (SRCTYPE,"j")))" gave out metadata of 45,988 documents.

The main parameters of the publication activity of the given organizations, which correspond to the request, are presented in Fig. 2 and 3.
Fig. 2. Number of publications by year. Scopus data: Analyze search results.

The decade can be conditionally divided into three periods: 2012-2015 — growth from a small number of publications, 2016-2018 - no growth, 2019-2021 rapid growth of the present (data as of August 06, 2021, not all 2021 documents are indexed yet). Next, I will analyze these intervals in more detail.

Fig. 3. Distribution of publications corresponding to the above query by Subject Areas. Scopus data — Analyze search results.

The main Subject Areas of research are: Chemistry, Engineering and Materials Science. Energy takes up a modest 4.5%. Computer Science (7.1%) and Mathematics (5.7%) suggest a significant role for digitalization in the scholarly research of these institutions.

The distribution of publication topics in the three previously noted intervals: 2012-2016; 2017-2019; and 2019-2021, is shown in Tab. 2.

Table 2. Distribution of publication topics by three-time intervals

<table>
<thead>
<tr>
<th>Subject Area</th>
<th>2012-2015</th>
<th>2016-2018</th>
<th>2019-2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemistry</td>
<td>5076</td>
<td>4555</td>
<td>5433</td>
</tr>
<tr>
<td>Engineering</td>
<td>4555</td>
<td>4310</td>
<td>5384</td>
</tr>
<tr>
<td>Materials Science</td>
<td>3899</td>
<td>4223</td>
<td>3940</td>
</tr>
<tr>
<td>Physics and Astronomy</td>
<td>3342</td>
<td>3181</td>
<td>5243</td>
</tr>
</tbody>
</table>
The relative number of Chemistry publications declined between 2019 and 2021, with Engineering and Materials Science coming in first and second during this period.

The number of publications in Computer Science has been increasing since the second interval, while the number of publications in Mathematics has been decreasing.

Recently, energy topics have been attracting more and more attention. It can be assumed that such changes reflect the process of the energy transition, which requires new engineering solutions and materials (especially for renewable energy storage systems), as well as methods for optimizing operations in energy systems, which requires new solutions in the field of computer science.

While Tab. 2 shows the ranking of topics by publication activity in a particular time interval, Tab. 3 shows the distribution of publications by year for each of the Subject Areas.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Engineering</td>
<td>409</td>
<td>423</td>
<td>581</td>
<td>802</td>
<td>834</td>
<td>845</td>
<td>770</td>
<td>912</td>
<td>1382</td>
<td>1074</td>
</tr>
<tr>
<td>Chemistry</td>
<td>1113</td>
<td>1058</td>
<td>1337</td>
<td>1568</td>
<td>1615</td>
<td>1541</td>
<td>1399</td>
<td>1549</td>
<td>2056</td>
<td>1638</td>
</tr>
<tr>
<td>Computer Science</td>
<td>370</td>
<td>450</td>
<td>529</td>
<td>569</td>
<td>586</td>
<td>671</td>
<td>745</td>
<td>910</td>
<td>1266</td>
<td>1014</td>
</tr>
<tr>
<td>Earth and Planetary Sciences</td>
<td>136</td>
<td>214</td>
<td>213</td>
<td>350</td>
<td>399</td>
<td>406</td>
<td>402</td>
<td>421</td>
<td>469</td>
<td>396</td>
</tr>
<tr>
<td>Energy</td>
<td>160</td>
<td>212</td>
<td>261</td>
<td>345</td>
<td>423</td>
<td>481</td>
<td>563</td>
<td>631</td>
<td>814</td>
<td>683</td>
</tr>
<tr>
<td>Engineering</td>
<td>809</td>
<td>867</td>
<td>1024</td>
<td>1255</td>
<td>1350</td>
<td>1574</td>
<td>1386</td>
<td>1618</td>
<td>2149</td>
<td>1686</td>
</tr>
<tr>
<td>Environmental Science</td>
<td>313</td>
<td>367</td>
<td>432</td>
<td>540</td>
<td>617</td>
<td>572</td>
<td>549</td>
<td>633</td>
<td>974</td>
<td>951</td>
</tr>
<tr>
<td>Materials Science</td>
<td>836</td>
<td>798</td>
<td>1009</td>
<td>1256</td>
<td>1303</td>
<td>1504</td>
<td>1416</td>
<td>1565</td>
<td>2175</td>
<td>1644</td>
</tr>
<tr>
<td>Mathematics</td>
<td>480</td>
<td>471</td>
<td>481</td>
<td>534</td>
<td>562</td>
<td>528</td>
<td>570</td>
<td>627</td>
<td>761</td>
<td>631</td>
</tr>
<tr>
<td>Physics and Astronomy</td>
<td>773</td>
<td>718</td>
<td>846</td>
<td>1005</td>
<td>1101</td>
<td>1065</td>
<td>1015</td>
<td>1150</td>
<td>1519</td>
<td>1271</td>
</tr>
</tbody>
</table>

The following list reflects the increase in the relative number of publications in 2020 compared to 2012 for Subject Areas that I believe are particularly important for the energy transition:

- Chemical Engineering 1382/409=3.38
- Computer Science 1266/370=3.42
- Energy 814/160=5.09
- Engineering 2149/809=2.66
- Materials Science 2175/836=2.60
- Environmental Science 974/313=3.11

While Engineering and Materials Science have grown in absolute terms, Energy has seen the highest relative growth, followed by Computer Science, Chemical Engineering, and Environmental Sciences—the picture is not contradicted by the growing interest in the energy transition to clean energy.

For clarity, the data in Table 3 are shown graphically in Fig. 4.
In order to catch trends in the development of research, it is crucial to know which countries are cooperating with, thus determining the potential for expanding the geographical scope of the topic.

Table 4 shows the distribution by year of co-authorship for 10 countries. Data as of August 6, 2021.

Table 4. The distribution by year of co-authorship for 10 countries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>214</td>
<td>165</td>
<td>261</td>
<td>354</td>
<td>437</td>
<td>502</td>
<td>563</td>
<td>734</td>
<td>1079</td>
<td>872</td>
</tr>
<tr>
<td>Egypt</td>
<td>293</td>
<td>364</td>
<td>515</td>
<td>598</td>
<td>669</td>
<td>574</td>
<td>449</td>
<td>586</td>
<td>1034</td>
<td>716</td>
</tr>
<tr>
<td>India</td>
<td>192</td>
<td>214</td>
<td>265</td>
<td>463</td>
<td>443</td>
<td>453</td>
<td>472</td>
<td>453</td>
<td>809</td>
<td>870</td>
</tr>
<tr>
<td>Malaysia</td>
<td>235</td>
<td>174</td>
<td>230</td>
<td>266</td>
<td>236</td>
<td>227</td>
<td>186</td>
<td>232</td>
<td>335</td>
<td>289</td>
</tr>
<tr>
<td>Pakistan</td>
<td>149</td>
<td>115</td>
<td>203</td>
<td>298</td>
<td>338</td>
<td>341</td>
<td>344</td>
<td>508</td>
<td>811</td>
<td>726</td>
</tr>
<tr>
<td>Russian Federation</td>
<td>11</td>
<td>6</td>
<td>8</td>
<td>29</td>
<td>40</td>
<td>44</td>
<td>73</td>
<td>71</td>
<td>72</td>
<td>57</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>2925</td>
<td>3093</td>
<td>3703</td>
<td>4387</td>
<td>4724</td>
<td>4720</td>
<td>4576</td>
<td>5195</td>
<td>6964</td>
<td>5433</td>
</tr>
<tr>
<td>South Korea</td>
<td>97</td>
<td>112</td>
<td>129</td>
<td>197</td>
<td>185</td>
<td>173</td>
<td>163</td>
<td>192</td>
<td>307</td>
<td>330</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>118</td>
<td>129</td>
<td>170</td>
<td>227</td>
<td>233</td>
<td>267</td>
<td>259</td>
<td>315</td>
<td>373</td>
<td>324</td>
</tr>
<tr>
<td>United States</td>
<td>382</td>
<td>435</td>
<td>500</td>
<td>605</td>
<td>596</td>
<td>595</td>
<td>626</td>
<td>633</td>
<td>703</td>
<td>523</td>
</tr>
</tbody>
</table>

Table 4 shows that prior to 2018, the majority of co-authors were from the U.S. Beginning in 2019, co-authors from China dominate, with an increase in the number of publications with U.S. co-authors continuing.

Authors from India, Egypt, Pakistan, and Malaysia are consistently in print along with authors from Saudi Arabia, a rate that increases significantly in 2020.

Publications with co-authors from South Korea and the United Kingdom steadily increasing over a whole ten-year period.

Authors from the Russian Federation publish poorly with co-authors from Saudi Arabia, even though there has been a relative increase over the years. The potential for cooperation between the Russian Federation and Saudi Arabia in the field of joint research remains unrealized, despite the extensive cooperation between the countries within the framework of OPEC+.

Using VOSviewer for Analysis of Scientific Publications Landscape
Bibliometric data of the 2,000 most-cited publications for each year were used to analyze trends in scientific research based on co-occurrence of key terms. Data as of September 09, 2021.

Why it makes sense to consider Author and Index keywords separately

Here is the Scopus definition for Author and Index keywords:

- **Author keywords**: These are keywords chosen by the author(s) which, in their opinion, best reflect the contents of their document.
- **Indexed keywords**: These are keywords chosen by content suppliers and are standardized based on publically available vocabularies. Unlike Author keywords, the Indexed keywords take into account synonyms, various spellings, and plurals.\(^4\)

For example, for the 343 records of Elsevier B.V., I obtained an average of 54.7 Indexed Keywords per record, for the 115 records of Academic Press Inc. - an average of 55.57 Indexed Keywords per entry, and for the 48 entries of the Royal Society of Chemistry, an average of 29.3. While for all of these publishers, I got an average of 5.5 Author keywords per entry.

To elaborate, here are two records:

**Ref. [6]**

Indexed keywords: larvicidal agent; nanocrystal; plant extract; Quisqualis indica extract; silver nanoparticle; unclassified drug; virus vector; insecticide; metal nanoparticle; plant extract; silver; acute toxicity; Aedes aegypti; Anisops bouvieri; Anopheles stephensi; aquatic species; Arbovirus; Article; atomic force microscopy; Combretaceae; crystal structure; Culex quinquefasciatus; Diplonychus indicus; Gambusia affinis; infrared spectroscopy; LC50; mosquito; nanofabrication; nonhuman; one pot synthesis; particle size; plant leaf; Plasmodium malariae; priority journal; Quisqualis indica; scanning electron microscopy; transmission electron microscopy; ultraviolet spectroscopy; X ray diffraction; Zika virus; animal; Anopheles; aquatic species; chemistry; Combretaceae; drug effects; growth, development and aging; larva; malaria; metabolism; mosquito vector; spectrometry; ultrastructure; ultraviolet spectrophotometry; Zika virus; Animals; Anopheles; Aquatic Organisms; Combretaceae; Insecticides; Larva; Malaria; Metal Nanoparticles; Microscopy, Atomic Force; Microscopy, Electron, Scanning; Mosquito Vectors; Particle Size; Plant Extracts; Plant Leaves; Silver; Spectrometry, X-Ray Emission; Spectrophotometry, Ultraviolet; Zika Virus – 73 terms.

Author keywords: AFM; Arbovirus; Green synthesis; Mosquito-borne diseases; Nanobiotechnology.

**Ref. [7]**

Indexed keywords: Carbon nitride; Chemical detection; Electrocatalysts; Electrodes; Electron transport properties; Fabrication; Green manufacturing; Iron compounds; Manganese compounds; Nanosensors; Nanostructured materials; Neurophysiology; Sonochemistry; Ultrasonic applications; Electrocatalytic activity; Electron transfer rates; High intensity ultrasonic; Neurotransmitter detection; Powder X ray diffraction; Real-time application; Sonochemical synthesis; Synthesis parameters; Electrochemical sensors; carbon; ferrous chloride; manganese sulfate; nanomaterial; neurotransmitter; serotonin; sodium dihydrogen phosphate; agents interacting with transmitter, hormone or drug receptors; cyanogen; ferric ion; manganese derivative; manganese ferrite; nitrile; aqueous solution; Article; chemical composition; controlled study; electrochemical detection; electron transport; elemental analysis; high intensity ultrasound; human; limit of detection; nanofabrication; nonhuman; priority journal; surface property; synthesis; X ray powder diffraction; animal; blood; brain; catalysis; chemistry; electrochemical analysis; green chemistry; metabolism; procedures; rat; ultrasound; X ray diffraction; Animals; Brain; Catalysis; Electrochemical Techniques; Ferric Compounds; Green Chemistry Technology; Humans; Limit of Detection; Manganese Compounds; Neurotransmitter Agents; Nitriles; Rats; Sonication; X-Ray Diffraction – 78 terms.

Author keywords: Bimetal oxides; Graphitic carbon nitride; Hybrid materials; Neurotransmitter detection; Sonochemical synthesis.

\(^4\) [https://service.elsevier.com/app/answers/detail/a_id/21730/supporthub/scopus/](https://service.elsevier.com/app/answers/detail/a_id/21730/supporthub/scopus/)
The validity of this expansion of the list of Indexed Keywords requires a separate study, which is beyond the scope of this publication. In this case, we can only assume that the Indexed Keywords broadly reflect the potential fields of research to which a particular article may be assigned.

**Analysis of Scientific Publications Landscape by Index Keywords for 20,000 Records**

As noted earlier, the main query yielded metadata of 45,988 documents.

For constructing and visualizing Index keywords networks, 2,000 metadata records of the most cited publications were taken for each year (20,000 records in total). Then, using the VOSviewer software, a network was constructed using the specified metadata, based on the co-occurrence of Index keywords.

The total number of Index keywords in this sample was 79295, of which 12598 occur more than 5 times and 6318 occur more than 10 times; 1000 Index keywords out of 6318 with maximum link strength are further used to construct a term co-occurrence network.

After an initial review of the results, a partial replacement of the most frequent terms was performed, e.g., terms can occur in plural or singular form in different records, or have a hyphen between words or not, etc. After this substitution, the total number of terms was 79250 of which 6273 occurred more than 10 times. The condition more than 10 in cluster was used to enlarge the clusters; this yielded 4 clusters, which are shown in Figure 5.

![Fig. 5. Clustering results based on the co-occurrence of Index keywords](image)

The following list shows the frequency data for each of the clusters in the format: number of unique terms/total number of terms:

1. 303/50164
2. 272/35835
3. 239/33723
4. 185/19733
Table 5: The most frequent Index keywords for each of the four clusters. Here: label cl I - name of the term in the I cluster, N - occurrence of the term in the sample

<table>
<thead>
<tr>
<th>label cl 1</th>
<th>N</th>
<th>label cl 2</th>
<th>N</th>
<th>label cl 3</th>
<th>N</th>
<th>label cl 4</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>scanning electron microscopy</td>
<td>1239</td>
<td>chemistry</td>
<td>1667</td>
<td>priority journal</td>
<td>1193</td>
<td>procedures</td>
<td>508</td>
</tr>
<tr>
<td>nanoparticle</td>
<td>1073</td>
<td>controlled study</td>
<td>1417</td>
<td>temperature</td>
<td>639</td>
<td>saudi arabia</td>
<td>422</td>
</tr>
<tr>
<td>x ray diffraction</td>
<td>1029</td>
<td>human</td>
<td>1155</td>
<td>carbon dioxide</td>
<td>637</td>
<td>ph</td>
<td>391</td>
</tr>
<tr>
<td>adsorption</td>
<td>935</td>
<td>unclassified drug</td>
<td>1063</td>
<td>polymer</td>
<td>526</td>
<td>bacteria</td>
<td>360</td>
</tr>
<tr>
<td>transmission electron microscopy</td>
<td>705</td>
<td>nonhuman</td>
<td>966</td>
<td>surface properties</td>
<td>473</td>
<td>copper</td>
<td>308</td>
</tr>
<tr>
<td>fourier transform infrared spectroscopy</td>
<td>665</td>
<td>metabolism</td>
<td>738</td>
<td>optimization</td>
<td>450</td>
<td>nitrogen</td>
<td>300</td>
</tr>
<tr>
<td>carbon</td>
<td>635</td>
<td>animal</td>
<td>697</td>
<td>energy efficiency</td>
<td>411</td>
<td>concentration (composition)</td>
<td>297</td>
</tr>
<tr>
<td>synthesis (chemical)</td>
<td>582</td>
<td>synthesis</td>
<td>687</td>
<td>kinetics</td>
<td>411</td>
<td>mass spectrometry</td>
<td>285</td>
</tr>
<tr>
<td>nanocomposites</td>
<td>571</td>
<td>particle size</td>
<td>581</td>
<td>algorithm</td>
<td>377</td>
<td>water treatment</td>
<td>254</td>
</tr>
<tr>
<td>density functional theory</td>
<td>555</td>
<td>chemical structure</td>
<td>542</td>
<td>desalination</td>
<td>363</td>
<td>zinc</td>
<td>254</td>
</tr>
<tr>
<td>efficiency</td>
<td>533</td>
<td>metal nanoparticles</td>
<td>482</td>
<td>membranes</td>
<td>358</td>
<td>aluminum</td>
<td>241</td>
</tr>
<tr>
<td>graphene</td>
<td>525</td>
<td>infrared spectroscopy</td>
<td>390</td>
<td>oxidation</td>
<td>349</td>
<td>genetics</td>
<td>239</td>
</tr>
<tr>
<td>catalysis</td>
<td>473</td>
<td>drug effect</td>
<td>348</td>
<td>water</td>
<td>347</td>
<td>&quot;water pollutants, chemical&quot;</td>
<td>239</td>
</tr>
<tr>
<td>electrode</td>
<td>473</td>
<td>human cell</td>
<td>330</td>
<td>porosity</td>
<td>337</td>
<td>water pollutant</td>
<td>231</td>
</tr>
<tr>
<td>energy gap</td>
<td>425</td>
<td>drug effects</td>
<td>319</td>
<td>thermodynamics</td>
<td>320</td>
<td>wastewater treatment</td>
<td>229</td>
</tr>
<tr>
<td>catalysts</td>
<td>416</td>
<td>drug synthesis</td>
<td>317</td>
<td>chemical analysis</td>
<td>316</td>
<td>sea water</td>
<td>227</td>
</tr>
<tr>
<td>metal</td>
<td>414</td>
<td>in vitro study</td>
<td>313</td>
<td>combustion</td>
<td>310</td>
<td>indian ocean</td>
<td>212</td>
</tr>
<tr>
<td>morphology</td>
<td>386</td>
<td>male</td>
<td>313</td>
<td>silica</td>
<td>284</td>
<td>lead</td>
<td>201</td>
</tr>
<tr>
<td>chlorine compounds</td>
<td>368</td>
<td>molecular structure</td>
<td>266</td>
<td>performance assessment</td>
<td>263</td>
<td>isolation and purification</td>
<td>199</td>
</tr>
<tr>
<td>solar cells</td>
<td>366</td>
<td>antioxidant</td>
<td>265</td>
<td>membrane</td>
<td>262</td>
<td>biomass</td>
<td>192</td>
</tr>
<tr>
<td>zinc oxide</td>
<td>366</td>
<td>hydrogen bond</td>
<td>264</td>
<td>energy utilization</td>
<td>258</td>
<td>water purification</td>
<td>192</td>
</tr>
<tr>
<td>high resolution transmission electron microscopy</td>
<td>354</td>
<td>escherichia coli</td>
<td>263</td>
<td>ethylene</td>
<td>252</td>
<td>physiology</td>
<td>191</td>
</tr>
<tr>
<td>crystal structure</td>
<td>347</td>
<td>silver</td>
<td>262</td>
<td>hydrophobicity</td>
<td>252</td>
<td>waste water</td>
<td>190</td>
</tr>
<tr>
<td>catalyst activity</td>
<td>346</td>
<td>enzyme activity</td>
<td>238</td>
<td>heat transfer</td>
<td>247</td>
<td>concentration (parameters)</td>
<td>188</td>
</tr>
<tr>
<td>titanium dioxide</td>
<td>345</td>
<td>oxidative stress</td>
<td>237</td>
<td>mixtures</td>
<td>240</td>
<td>water management</td>
<td>187</td>
</tr>
<tr>
<td>oxygen</td>
<td>342</td>
<td>antiinfective agent</td>
<td>236</td>
<td>internet of things</td>
<td>239</td>
<td>bacterium</td>
<td>185</td>
</tr>
<tr>
<td>perovskite</td>
<td>326</td>
<td>animal experiment</td>
<td>230</td>
<td>solar energy</td>
<td>236</td>
<td>organic compound</td>
<td>183</td>
</tr>
<tr>
<td>nickel</td>
<td>325</td>
<td>metal nanoparticle</td>
<td>228</td>
<td>forecasting</td>
<td>235</td>
<td>aqueous solution</td>
<td>178</td>
</tr>
<tr>
<td>thin films</td>
<td>312</td>
<td>female</td>
<td>227</td>
<td>fuel</td>
<td>234</td>
<td>iron</td>
<td>178</td>
</tr>
<tr>
<td>heterojunctions</td>
<td>306</td>
<td>antineoplastic agent</td>
<td>225</td>
<td>polymerization</td>
<td>233</td>
<td>limit of detection</td>
<td>167</td>
</tr>
</tbody>
</table>

The first cluster theme:
adsorption, catalysis, morphology, solar cells, thin films, heterojunctions
nanoparticle, carbon, nanocomposites, graphene, metal, chlorine compounds, zinc oxide, titanium dioxide, perovskite
scanning electron microscopy, x ray diffraction, transmission electron microscopy, fourier transform infrared spectroscopy

The second cluster theme:
chemistry, metabolism, synthesis, drug synthesis, chemical structure, enzyme activity, oxidative stress, antiinfective agent, male, female
unclassified drug, animal, human cell, metal nanoparticles, antioxidant, escherichia coli
infrared spectroscopy, controlled study, in vitro study, animal experiment

The third cluster theme:
optimization, energy efficiency, energy utilization, internet of things, solar energy,
carbon dioxide, polymer, water, membranes, silica, ethylene, fuel
kinetics, algorithm, thermodynamics, chemical analysis, performance assessment, forecasting

The fourth cluster theme:
Saudi Arabia, wastewater treatment, sea water, waste water, water management
bacteria, copper, nitrogen, zinc, aluminum, lead, biomass, organic compound
mass spectrometry, Ph, concentration, isolation and purification, limit of detection

Terms do not occur evenly over time; the simplest estimate of this is to use the average year of occurrence of the term in the metadata of publications, <Avg. pub. Year>, the results for some of the terms in the four clusters are shown in Table 6. This parameter will be used later in the construction of the final graphs: Fig. 15. and Fig. 16.

Table 6: Average year of occurrence of some terms in publication metadata for the four clusters

<table>
<thead>
<tr>
<th>label cl 1</th>
<th>Avg. pub. year</th>
<th>label cl 2</th>
<th>Avg. pub. year</th>
<th>label cl 3</th>
<th>Avg. pub. year</th>
<th>label cl 4</th>
<th>Avg. pub. year</th>
</tr>
</thead>
<tbody>
<tr>
<td>photocatalytic activity</td>
<td>2019.7</td>
<td>hep-g2 cell line</td>
<td>2018.4</td>
<td>deep learning</td>
<td>2019.7</td>
<td>concentration (parameter)</td>
<td>2020.0</td>
</tr>
<tr>
<td>surface morphology</td>
<td>2019.5</td>
<td>tumor necrosis factor</td>
<td>2018.3</td>
<td>machine learning</td>
<td>2019.1</td>
<td>detection method</td>
<td>2019.0</td>
</tr>
<tr>
<td>x ray photoemission spectroscopy</td>
<td>2019.3</td>
<td>targeted drug delivery</td>
<td>2018.1</td>
<td>internet of things</td>
<td>2018.8</td>
<td>inorganic compound</td>
<td>2018.7</td>
</tr>
<tr>
<td>reduction (chemistry)</td>
<td>2019.3</td>
<td>malonaldehyde</td>
<td>2018.0</td>
<td>sustainable development</td>
<td>2018.6</td>
<td>plant growth</td>
<td>2018.5</td>
</tr>
<tr>
<td>supercapacitor</td>
<td>2019.1</td>
<td>photon correlation spectroscopy</td>
<td>2018.0</td>
<td>learning systems</td>
<td>2018.5</td>
<td>soil pollutant</td>
<td>2018.1</td>
</tr>
<tr>
<td>nanocatalysts</td>
<td>2019.0</td>
<td>enzyme</td>
<td>2017.9</td>
<td>diagnosis</td>
<td>2018.3</td>
<td>soil pollution</td>
<td>2018.0</td>
</tr>
<tr>
<td>potassium hydroxide</td>
<td>2019.0</td>
<td>brain</td>
<td>2017.8</td>
<td>network security</td>
<td>2018.2</td>
<td>photosynthesis</td>
<td>2018.0</td>
</tr>
<tr>
<td>green synthesis</td>
<td>2018.9</td>
<td>histopathology</td>
<td>2017.8</td>
<td>composite</td>
<td>2018.2</td>
<td>soil pollutants</td>
<td>2018.0</td>
</tr>
<tr>
<td>silver compounds</td>
<td>2018.9</td>
<td>diseases</td>
<td>2017.8</td>
<td>aluminum oxide</td>
<td>2018.0</td>
<td>adsorbent</td>
<td>2018.0</td>
</tr>
<tr>
<td>crystallinity</td>
<td>2018.9</td>
<td>tumors</td>
<td>2017.8</td>
<td>nanofluidics</td>
<td>2018.0</td>
<td>antibiotics</td>
<td>2017.9</td>
</tr>
<tr>
<td>layered semiconductors</td>
<td>2018.8</td>
<td>controlled drug delivery</td>
<td>2017.8</td>
<td>surface charge</td>
<td>2017.9</td>
<td>biochar</td>
<td>2017.8</td>
</tr>
<tr>
<td>energy storage</td>
<td>2018.8</td>
<td>histology</td>
<td>2017.7</td>
<td>flame</td>
<td>2017.8</td>
<td>plant root</td>
<td>2017.8</td>
</tr>
<tr>
<td>binary alloys</td>
<td>2018.8</td>
<td>glutathione reductase</td>
<td>2017.7</td>
<td>membrane distillation</td>
<td>2017.8</td>
<td>wastewater treatment</td>
<td>2017.6</td>
</tr>
<tr>
<td>lead compounds</td>
<td>2018.7</td>
<td>drug effect</td>
<td>2017.7</td>
<td>fluorne compounds</td>
<td>2017.6</td>
<td>water pollution</td>
<td>2017.6</td>
</tr>
<tr>
<td>oxide minerals</td>
<td>2018.7</td>
<td>chitosan</td>
<td>2017.7</td>
<td>wetting</td>
<td>2017.6</td>
<td>soil</td>
<td>2017.6</td>
</tr>
<tr>
<td>manganese compounds</td>
<td>2018.6</td>
<td>green chemistry</td>
<td>2017.7</td>
<td>thermal conductivity</td>
<td>2017.6</td>
<td>hydrogen peroxide</td>
<td>2017.5</td>
</tr>
</tbody>
</table>
To identify trends in the topics of publications described by keywords, in addition to the average year of appearance of the term, it is important to assess the citation rate of publications in which the term is found. Given that the citation rate of publications increases over time, it is advisable to use citation normalization, e.g., over a certain period of time. I used the parameter <Avg. norm. Citations>, which is available in the map files exported by VOSviewer. The results for the terms with the highest Avg. norm. Citations for the four clusters are shown in Table 7.

Table 7. Average normalized citation of terms for a variety of Index keywords from each of the four clusters
The data given in Table 7 will be used to build the final graphs shown in Fig. 15 and Fig. 16.

The gradation of key terms can be very different, depending on the task at hand. For example, for a more detailed description of cluster topics, keywords can be assigned to different categories, conventionally labeled: research subject, materials, methods. In the lists below for each of the clusters, I have tried to intuitively divide the key terms into these categories. This is solely the author's categorization, given only to show that this division of terms can improve understanding of the topics of each of the clusters. To reduce biased judgments, a reasonable formalization of such categorization is necessary, which is beyond the scope of this paper, but can be accomplished by compiling for each of the categories its own vocabulary based on the occurrence of the terms in the corresponding sections of the full texts of the publications: introduction, materials, methods.

When selecting keywords to describe the cluster, it is important to consider what criteria were used to rank keywords. For example, in Table 5 they were ranked by occurrence, in Table 6 there was an additional ranking of frequently occurring keywords by <Avg.pub. Year>, and in Table 7 they were ranked by <Avg. norm. citation>. The leading keyword samples will be somewhat different, as will the cluster descriptions. At the same time, some commonality in the subject matter of the clusters is retained. This can be seen from the cluster descriptions in Table 5 below and the following cluster descriptions in Tables 7 and 6:

1 cluster, conditional division of key terms into categories: subject of research, materials, methods.

- power conversion efficiencies, perovskite solar cells, water splitting, solar power generation, fuel cell
- visible light, methylene blue, metal organic framework, transition metals, perovskite, monolayers, activated carbon, quantum dot, lead compounds
- open circuit voltage, photobleach spectroscopy, photocatalytic activities, photolysis, electrochemistry, photodegradation

2 cluster
• biomedical applications, green chemistry technology, chemical phenomena, nanoencapsulation
• ferric compounds, cancer cell, ferric ion, hela cells, zinc oxide nanoparticle, neoplasms, coordination compound
• nuclear magnetic resonance imaging, "microscopy, electron, transmission", ultraviolet spectrophotometry, "microscopy, electron, scanning", x-ray diffraction

3 cluster
• photovoltaic system, solar power, sustainable development, photovoltaic cells, seawater desalination, internet of things
• film, solvent, gas, composite, water
• stoichiometry, water filtration, decomposition, separation, evaporation

4 cluster
• water treatment, water purification, biotechnology, ecology, water supply, environmental impact, waste water management
• sulfate, biochar, charcoal, nanotubes, inorganic compound, copper, potassium, lead
• sorption, adsorption kinetics, pollutant removal, oxidation reduction reaction, waste component removal

The most affordable method of viewing the change in occurrence of terms over time is to use the VOSviewer Overlay graph for the time parameter. An example is shown in Fig. 6.

![VOSviewer Overlay graph](image)

**Fig. 6.** Changes in the occurrence of Index keywords over time. Nodes of yellow color occur more often in the red cluster, if we compare with Fig. 5.

The developers of the VOSviewer program currently offer the ability to build and view graphs online. To do this, you need to upload graph files exported from the desktop VOSviewer at [https://app.vosviewer.com/](https://app.vosviewer.com/), the json format is the most suitable. Readers of this preprint can use the attached files:
Index KWs 2012-2021 JSON.json and Author KWs 2012-2021.json. The online interface is user-friendly and allows a more detailed review and study of the graphs presented in this publication in Figs. 5-10.

Analysis of Scientific Publications Landscape by Author Keywords for 20,000 Records

The data processing procedure for Author Keywords was the same as for Index keywords, described in the previous section.

The total number of Author Keywords in this sample was 40369, which is significantly less than the Index Keywords — 79295, of which 2455 occurred more than 5 times.

In the absence of cluster size restrictions, I got 16 clusters. Therefore, a limit of at least 120 terms per cluster was implemented, resulting in 6 clusters.

The connectivity of clusters for Author Keywords is significantly lower than for Index Keywords (Fig. 7.). There may be several reasons: a much smaller total number of Author Keywords compared to Index keywords, authors having their own opinion on the choice of keywords, for example, the terms: deep leaning, internet of things, and multi-step methods (green cluster) are more often used by authors than in Index keywords.

Fig. 7. Clustering results based on co-occurrence of Author's keyword

For better viewing, the central part of Fig. 7. Presented in enlarged format in Fig. 8.
Fig. 8. Clustering results based on the co-occurrence of the Author's keywords, zoomed in of the central part of the graph in Fig. 7.

Fig. 9 shows the evolution of the occurrence of terms over time using the VOSviewer Overlay graph for the time variable.

Figure 9 shows that Author Keywords are likely to be indicative for terms that appear more frequently in new publications. I believe that it is the Author Keywords that are better suited for identifying emerging trends in scientific publications, and the Index Keywords are more interesting for identifying potential directions for expanding the application of results obtained in emerging trends.
VOSviewer allows us to detect not only the change of key terms over time, but also to show the distribution of citations of publications in which a particular term occurs. In my opinion, in the available publications on bibliometrics this feature is rarely used, but it serves as an indirect expert assessment of the topics described by the terms given in this graph (Fig. 10).

**Fig. 10.** Average Citation Rates of publications containing various Author keywords

The conjoint use of Figures 9 and 10 makes it possible to find keywords whose use is more common in new publications, and whose publications are highly cited. Figure 11 shows fragments of Figures 9 and 10 related to the term Visible light, which satisfies the above requirements.

**Fig. 11.** The term Visible light. On the left is the occurrence over time, on the right the average citation rate of publications with the given keyword.

The figure shows that the term Visible light is used in the context of the keywords Photocatalysis and Nanocomposites.

Figure 12 shows fragments of Figures 9 and 10 related to the term Chitosan, which satisfies only one requirement.
As can be seen in Figures 11 and 12, using the average publication time of articles containing a given term (score_Avg_pub_year) and their average citation rate (score_Avg_citations) allows us to clearly assess possible research trends in scientific publications. The simplest formalization of the term selection process for such consideration may look as follows: we set a query to a file containing, for example, the Author's keywords and their score parameters. At the same time, we set restrictions on the threshold of occurrence of a key term, the average year of occurrence, and the average citation rate.

Let's look at an example result of such a query: select * from "Author KWs map csv" where "weight_Occurrences" > 10 AND "score_Avg_pub_year" > 2018 AND "score_Avg_citations" > 50.

Table 8: Some terms that satisfy the above query.

<table>
<thead>
<tr>
<th>label</th>
<th>cluster</th>
<th>Weight Links</th>
<th>Weight Occurrences</th>
<th>Score_Avg_pub_year</th>
<th>Score_Avg_citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>mxene</td>
<td>1</td>
<td>23</td>
<td>21</td>
<td>2018.19</td>
<td>109.43</td>
</tr>
<tr>
<td>bending</td>
<td>2</td>
<td>10</td>
<td>16</td>
<td>2018.69</td>
<td>69.38</td>
</tr>
<tr>
<td>biochar</td>
<td>5</td>
<td>37</td>
<td>36</td>
<td>2018.03</td>
<td>68.00</td>
</tr>
<tr>
<td>co2 reduction</td>
<td>1</td>
<td>17</td>
<td>13</td>
<td>2018.15</td>
<td>67.77</td>
</tr>
<tr>
<td>vibration</td>
<td>2</td>
<td>8</td>
<td>20</td>
<td>2018.40</td>
<td>61.80</td>
</tr>
<tr>
<td>buckling</td>
<td>2</td>
<td>8</td>
<td>18</td>
<td>2018.89</td>
<td>58.89</td>
</tr>
<tr>
<td>nonfullerene</td>
<td>1</td>
<td>7</td>
<td>17</td>
<td>2018.71</td>
<td>57.53</td>
</tr>
<tr>
<td>acceptors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dyes</td>
<td>5</td>
<td>28</td>
<td>17</td>
<td>2018.06</td>
<td>54.59</td>
</tr>
<tr>
<td>free vibration</td>
<td>2</td>
<td>7</td>
<td>18</td>
<td>2018.67</td>
<td>52.33</td>
</tr>
<tr>
<td>perovskite</td>
<td>1</td>
<td>11</td>
<td>17</td>
<td>2018.76</td>
<td>51.12</td>
</tr>
<tr>
<td>solar cells</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zinc</td>
<td>3</td>
<td>22</td>
<td>12</td>
<td>2018.17</td>
<td>50.75</td>
</tr>
<tr>
<td>composite pcm</td>
<td>2</td>
<td>14</td>
<td>11</td>
<td>2019.36</td>
<td>50.09</td>
</tr>
</tbody>
</table>

Some values are highlighted in yellow, by which terms can be selected for further, more detailed consideration of their role in describing research trends. The semantic significance of the selected terms can be determined using the description of the topics on the ScienceDirect platform:

- MXenes [https://www.sciencedirect.com/topics/materials-science/mxene]
- Biochar [https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/biochar]
- Dyes [https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/dye]

The term composite pcm (phase change material (PCM)) is found in the most recent publications (Avg_pub_year 2019.36). Phase change materials are used to store thermal energy, which is the second most important task after electrical energy storage. Examples of papers exposing this topic are [8, 9]
The disadvantage of the considered approach is the use of Avg. citations indicator, which increases over time, to assess the importance of key terms it is more appropriate to use a normalized estimate of the average citation (Avg. norm. citations). It is also necessary to propose a graphical representation that is convenient for the final assessment of the importance of terms for the description of research trends.

**Graphical representation of the final choice of terms to describe research trends**

The structure of scientific research landscapes well reflects the overall picture of publication activity, clustering based on the co-occurrence of key terms makes it possible to identify the main areas in scientific research. The use of an overlay makes it possible to trace changes in the use of terms over time and determine the citation rate of publications containing certain terms. When analyzing trends in scientific publications, VOSviewer allows you to browse flexibly through particular areas of the scientific research landscape, as shown in Figure 5, and 7. But when compiling an analytical report based on bibliometric data, the interactivity provided by analytical platforms and programs is lost; moreover, the report should provide summaries highlighting the main conclusions of the bibliometric study. Therefore, it is advisable to supplement the overall picture of publication activity, relevant to the main query to the abstract database, with illustrations reflecting the main results.

In my opinion, in order to identify trends in the topics of scientific publications, it is advisable to reflect the following data on the final graph: key terms and their occurrence over time; the average normalized citation rate of publications containing the selected key terms. The above parameters can be retrieved from the data export files provided by VOSviewer. In such files, they are stored in the record fields: label, weight<Occurrences>, score<Avg. pub. year> and score<Avg. norm. citations>.

Thus, the final graph should reflect the four parameters and be two-dimensional for readability. Bubble charts meet these requirements and are further used to present the final results in Figures 13, and 14.

Given that different time intervals are characterized by the use of different key terms, and that too many bubbles in the graph make it unreadable, the bibliometric data were divided into two periods - before 2017 and after 2017 (horizontal axis).

The second parameter I used to limit the choice of terms for the final graph was the average normalized citation (vertical axis), which was taken above 1.0. Next, the 15 most frequent key terms were selected. Each term was represented by its color, and the size of each term corresponded to its frequency of occurrence in the sample.

The final graphs for Author keywords and Index keywords are shown in Figures 13 - 14 and Figures 15 - 16, correspondingly.
Water splitting, activated carbon and hydrogen production were the main terms in the publications and had an average normalized citation rate. Photocatalytic activity, taxonomy and phylogeny had the highest citation in the given time interval.

Examples of publications containing the Author keywords—Water Splitting:

- Photocatalytic Water Splitting: Quantitative Approaches toward Photocatalyst by Design [10]
- Vertically aligned Ta₃N₅ nanorod arrays for solar-driven photoelectrochemical water splitting [13]
- Tungsten carbide nanoparticles as efficient cocatalysts for photocatalytic overall water splitting [14]

Definition: «In photoelectrochemical (PEC) water splitting, hydrogen is produced from water using sunlight and specialized semiconductors called photoelectrochemical materials, which use light energy to directly dissociate water molecules into hydrogen and oxygen»[^5]

Examples of publications containing Author keyword — taxonomy. The context for the term taxonomy is not obvious, but the titles of the articles disclose it:

- Constructing Features for Detecting Android Malicious Applications: Issues, Taxonomy and Directions [16]
- Internet of Things Architecture: Recent Advances, Taxonomy, Requirements, and Open Challenges [17]
- Survey of mobile device virtualization: Taxonomy and state of the art [18]
- Securing software defined networks: Taxonomy, requirements, and open issues [19]

In Deep learning, Features extraction, anomaly detection — taxonomy is a structuring and conceptualization of data in a particular subject area, which allows to reduce the dimensionality of the task and increase the stability of the analysis results.

**Fig. 14.** The 15 most frequent and highly cited Author keywords for the time interval > 2017.

This time interval is most characterized by topics described by the keywords visible light, solar cells, perovskite, and biochar.

For identifying emerging trends in scientific publications, the terms Mxene and composite PCM are of most interest. The first is characterized by a high citation rate and the middle of the time interval, the second by an average citation rate, but the prevalence in newer publications.

Examples of publications containing the key term Mxene:
- Versatile N-Doped MXene Ink for Printed Electrochemical Energy Storage Application [20]
- Asymmetric Flexible MXene-Reduced Graphene Oxide Micro-Supercapacitor [21]
- Large Dielectric Constant Enhancement in MXene Percolative Polymer Composites [22]
- MXene Ti$_3$C$_2$: An Effective 2D Light-to-Heat Conversion Material [23]
- MXene-on-Paper Coplanar Microsupercapacitors [24]

Examples of publications containing the key term Composite PCM:
- Silica fume/capric acid-palmitic acid composite phase change material doped with CNTs for thermal energy storage [25]
- Preparation, characterization and thermal regulation performance of cement based-composite phase change material [26]
- Thermal characteristics of expanded perlite/paraffin composite phase change material with enhanced thermal conductivity using carbon nanotubes [27]
- Thermal energy storage characteristics of bentonite-based composite PCMs with enhanced thermal conductivity as novel thermal storage building materials [28]
Fig. 15. The 15 most frequent and highly cited Index keywords for the time interval < 2017.

In this case, the main topics are described with the words metal organic frameworks, electrochemistry, monolayers, decomposition and crystallization.

Sample publications:

- A microporous metal-organic framework with naphthalene diimide groups for high methane storage [29]
- A microporous metal-organic framework with basic sites for efficient C₂H₂/CO₂ separation [30]
- Effects of structural crystallinity and defects in microporous Al-MOF filled chitosan mixed matrix membranes for pervaporation of water/ethanol mixtures [31]
- Light Hydrocarbon Adsorption Mechanisms in Two Calcium-Based Microporous Metal Organic Frameworks [32]

Electron and electron transport – terms that appear in the most cited articles.

Sample publications:

- Phototuning Selectively Hole and Electron Transport in Optically Switchable Ambipolar Transistors [33]
- Solution-Processed In₂O₃/ZnO Heterojunction Electron Transport Layers for Efficient Organic Bulk Heterojunction and Inorganic Colloidal Quantum-Dot Solar Cells [34]
- Hydrothermally synthesized titania nanotubes as a promising electron transport medium in dye sensitized solar cells exhibiting a record efficiency of 7.6% for 1-D based devices [35]
Fig. 16. The 15 most frequent and highly cited Index keywords for the time interval > 2017.

Perovskite, perovskite solar cells and power conversion efficiencies – highly cited topics well described by these Index keywords.

Sample titles of publications that reflect well on perovskite solar cells:

- Managing grains and interfaces via ligand anchoring enables 22.3%-efficiency inverted perovskite solar cells [36]
- Single-Crystal MAPbI₃ Perovskite Solar Cells Exceeding 21% Power Conversion Efficiency [37]
- Inorganic CsPbI₂Br Perovskite Solar Cells: The Progress and Perspective [38]
- Phase Transition Control for High Performance Ruddlesden–Popper Perovskite Solar Cells [39]
- Highly efficient perovskite solar cells based on a nanostructured WO₃-TiO₂ core-shell electron transporting material [40]

Transition metals - it is advisable to disclose the use of this term in more detail, since its context is not obvious from the above chart. For this purpose, the publications in which the term Transition metals appears both in the Index keywords and in the titles of the articles have been picked.

Samples of article titles that give a good context for the term Transition metals:

- Synthesis and characterization of binary selenides of transition metals to investigate its photocatalytic, antimicrobial and anticancer efficacy [41]
- Heteroatom-doped magnetic hydrochar to remove post-transition and transition metals from water: Synthesis, characterization, and adsorption studies [42]
- Unraveling the role of entropy in tuning unimolecular vs. bimolecular reaction rates: The case of olefin polymerization catalyzed by transition metals [43]
- Complexation of trichlorosalicylic acid with alkaline and first row transition metals as a switch for their antibacterial activity [44]
- Molecule-Level g-C₃N₄ Coordinated Transition Metals as a New Class of Electrocatalysts for Oxygen Electrode Reactions [45]
From the list above we can see that the context for transition metals is the terms photocatalytic, antimicrobial and anticancer efficacy, olefin polymerization, and remove post-transition and transition metals from water which agrees well with the term photolytic degradation in Fig. 16.

Conclusions

A bibliometric analysis of scientific publications showing trends in scientific publications from leading universities in Saudi Arabia related to the energy transition was conducted.

The following universities were selected based on the presence of joint publications with the Saudi Arabian Oil Company they are: King Fahd University of Petroleum and Minerals, King Abdullah University of Science and Technology, and King Saud University.

According to data from the Scopus platform related to the subject areas: Physical Sciences, scientists from Saudi Arabian universities are increasingly publishing with co-authors from China, India, Egypt, Pakistan, and Malaysia, while maintaining a high level of interaction with co-authors from the US, South Korea, and the UK. Authors from the Russian Federation publish poorly with co-authors from Saudi Arabia.

A significant difference between the Scopus platform Index keywords and Author keywords has been shown, which should be taken into account when identifying trends in scientific publications.

The main subject areas of research are: Chemistry, Engineering, and Materials Science. Energy occupies a modest 4.5%. Computer Science (7.1%) and Mathematics (5.7%) indicate a significant role for digitalization in research at these institutions.

In recent years, the relative number of chemistry publications has declined, while engineering and materials science has taken the lead.

There has been a relative increase in attention to energy topics, reflecting the energy transition process, which requires new engineering solutions and materials (especially for renewable energy storage systems), as well as methods for optimizing performance in energy systems, which requires new computer science solutions.

Based on the co-occurrence of Index keywords, four sub-themes are identified, which can be briefly described by the terms:

- adsorption, catalysis, morphology, solar cells, thin films, heterojunctions, nanoparticle, carbon, nanocomposites, graphene, metal, chlorine compounds, zinc oxide, titanium dioxide, perovskite
- chemistry, metabolism, synthesis, drug synthesis, chemical structure, enzyme activity, oxidative stress, anti-infective agent, unclassified drug, animal, human cell, metal nanoparticles, antioxidant
- optimization, energy efficiency, energy use, internet of things, solar energy, carbon dioxide, polymer, water, membranes, silica, ethylene, fuel
- Saudi Arabia, wastewater treatment, seawater, wastewater, water management, bacteria, copper, nitrogen, zinc, aluminum, lead, biomass, organic compound

Shows the feasibility of using additional keyword ranking for <Avg.pub. Year> and <Avg. norm. citation> to identify promising areas of research.

The use of bubble charts to present the final results of bibliometric studies to identify emerging research topics has been proposed.

Based on the analysis, two areas of research that are advisable to study in more detail are: composite phase change materials (PCM) and MXenes - carbides and nitrides of transition metals, a fast-growing and already very large family of 2D materials.

In methodological terms, it is advisable to formalize in more detail the approaches proposed in the article to the graphical representation of the final data on the identification of promising areas of research and additional separation of keywords in the clusters by the categories Subject, Materials, and Methods.
References


