Identifying Promising Domains of Decarbonization Technologies: An Improved Methodology

Paulo Moisés Costa
Electrical Department
Institute Polytechnical of Viseu
Viseu, Portugal
paulomoises@estgv.ipv.pt

Paulo Tomé
Computer Science
Institute Polytechnical of Viseu
Viseu, Portugal
ptome@estgv.ipv.pt

António Duarte
Electrical Department
Institute Polytechnical of Viseu
Viseu, Portugal
estgv17088@alunos.estgv.ipv.pt

Nuno Bento
Instituto Universitario de Lisboa (ISCTE-IUL)
DINAMIA ´CET
Lisbon, Portugal
Nuno.Bento@iscte.pt

Bruno Almeida
Electrical Department
Institute Polytechnical of Viseu
Viseu, Portugal
estgv16783@alunos.estgv.ipv.pt

***Abstract*—Humanity has sought ways to decarbonize the economies and mitigate the effects of global warming. The scientific community has been producing intensive research on decarbonizing technologies and strategies. This article introduces an improved methodology, based on two available text-mining software, to identify the most relevant domains of decarbonizing technologies presented in scientific papers, research projects and patents. The improvements include the definition of an approach to obtain a final rank of the technologies' domains and an enhanced procedure to create the semantic dictionary, a critical component for the bibliometric analysis. Moreover, a more robust and comprehensive search string was defined, reconciling the particularities of the used text-mining software. The methodology was applied to the 2011 to 2021 period. The results evidence the twenty better-ranked domains, where energy efficiency/management, hydrogen, generic renewable energy, carbon capture, energy storage and biofuel occupy the six first positions. Hydrogen, carbon capture and energy storage are the domains with a more accelerated increase in the scientific community's relative interest. In contrast, energy efficiency and management, biofuel, and low-carbon transportation have lost relative interest. Moreover, results show that electrification, hydrogen, and geothermal energy domains significantly increase their relevance/occurrence from 2019 to 2021.**

***Index Terms*—Decarbonisation, Text-mining, Bibliometric analysis, Global warming, Energy system**

Acknowledgment

The authors acknowledge the financial support from FCT – Foundation for Science and Technology, I.P., to the research project Sus2Trans (PTDC/GESAMB/0934/2020). Paulo Costa and Paulo Tome also acknowledges funding from FCT of the project Refª UIDB/05583/2020 and thanks the Research Centre in Digital Services (CISeD) and the Polytechnic of Viseu for their support.

I. Introduction

In recent years there has been increasing pressure to decarbonize human activities. The Paris Agreement is the most significant achievement to date concerning climate change, where it was established that emissions must be halved by 2030 and that 'net zero' should be reached by 2050. Those objectives appear necessary to ensure warming of 1.5 °C or less compared to pre-industrial levels and avoid the most harmful effects of climate change.

A recent Intergovernmental Panel on Climate Change (IPCC) report indicates that the observed global warming is worse and faster than feared and that the implementation of the Paris Agreement is slow and insufficient to achieve the proposed goals [1], [2]. Thus, the urgency in implementing procedures and technologies that allow accelerating decarbonization, aiming to mitigate the worst effects of climate change.

An extensive amount of scientific work on pathways and technologies for decarbonization has been published over the last few years. As an illustration, given the high number of publications, some recent work on paths to decarbonization can be found in [3]–[9]. Concerning technologies, some recent publications can be found in [10]–[14]. As discussed in a previous work by the authors, the high volume of research on decarbonization technologies and procedures can make understanding the relevance of the presented proposals challenging [15]. The use of methodologies based on text-mining tools can facilitate the understanding of decarbonization technologies that, over the years, have received greater attention from the scientific community.

In previous work, the authors of this paper introduced a methodology based on using two free text-mining tools to perceive the domains of technologies that, according to the scientific community, appear more promising for decarbonization [15]. That methodology is now improved by adopting an approach that allows comparing the results of the two used text-mining software (which are not directly comparable, as explained in [15]) and obtaining a global rank of the domains of technologies for decarbonization. An improved version of the procedure used to construct the semantic dictionary, a critical component needed for filtering and eliminating terms/words retrieved by the text-mining software that are irrelevant to the intended study and aggregating various versions of the same relevant meaning, is also presented. Moreover, applying the proposed methodology allows for obtaining a more comprehensive set of conclusions about the domains of decarbonising technologies.

II. Related works

Some works in the scientific literature use text-mining software to track decarbonization technologies and policies. Reza et al. [16] searched the Scopus database for articles indexed between 2011 and 2021 on energy storage and decarbonization. The authors conclude that integrating storage systems into grids can be essential in decarbonizing the electricity sector. Indeed, when associated with the integration of renewable energies, the storage systems allow replacing conventional fuels used in electricity generation and provide the necessary flexibility to face problems related to the dependence of renewable generation on weather conditions.

The impact of the diffusion of energy efficiency technologies on decarbonization in European residential buildings is assessed in [17]. The authors use text-mining software to explore the Web of Science (WoS) database for 2008-2018. They concluded an asymmetrical research activity exists in Europe since certain member states show an active research activity and others have little or no publications. They also conclude that, although energy efficiency has seen exponential growth in the number of publications, more research is needed to quantify and monetize its impacts.

A bibliometric perspective on research on carbon capture, storage and use (CCSU) in China is done in [18]. The Citespace software is used to map the information of 1202 scientific documents taken from the WoS database, concerning the research focuses, most productive organizations and principal authors for 2002-2019. The authors point out that the establishment of government policies around CCSU influenced the trend of research in China.

In [19], the authors use the VOSviewer software, a free tool, to perform a bibliometric analysis of research trends around green Hydrogen (H_2), analyzing the publications extracted from the Scopus database for the 2016 to 2021 period. The authors highlight the number of articles published in the adopted period, the most productive organizations and countries, and the most relevant research items about green H_2.

A revision of the literature about the main concepts related to climate change mitigation and its progress over time is presented in [20]. The authors use the Scopus database to obtain the scientific content and the VOSviewer software to analyze it, extracting conclusions about the geographic and sectoral focus and determining the interrelationships between the key concepts. This study stands out from the others for studying decarbonization in general and not a specific technology or sector, even though it uses only one scientific database and one text-mining software.

CiteSpace software is used in [21] to conduct a bibliometric analysis of developments in low-carbon energy generation. The research is based on 1419 articles from 1983 to 2021, retrieved by the Scopus database. The authors concluded that the number of publications has increased over time and that China and the United States lead publications on this topic. They also conclude that renewable energy resources and that energy storage are crucial for decarbonizing.

At [22] the Science Citation Index Expanded and Social Sciences Citation Index databases are used to gather the literature on the decarbonization of the electricity system on a global scale. Using a quantitative analysis of article titles, abstracts and keywords, the authors conclude that dependence on coal and nuclear fuels in the 1990s stimulated the search for cleaner alternatives, resulting in more publications on Carbon Capture and Storage and wind energy in the first decade of the 21st century. Furthermore, the authors conclude that the constant focus on energy efficiency will continue.

III. Methodology for Bibliometric Analysis

Fig. 1 summarizes the proposed approach, which starts with obtaining the set of documents (scientific publications, patents) that obey a carefully designed boolean search string, prepared to capture the desired information. It is essential to point out that the sets of documents obtained to be analyzed by VOSviewer and by TIM are different since different databases are used (in this case, the WoS database for VOSviewer and the SCOPUS, CORDIS, and PATSTAT databases for TIM). Note that the TIM platform automatically retrieves the set of documents respecting the condition defined by the search string, not imposing a previous extraction procedure as in the case of VOSviewer.

The obtained documents are then processed to attain the ”Author Keywords” and their occurrence in the VOSviewer case and the ”Relevant Keywords” and their relevance in the TIM software (the readers are invited to refer to [15] for details on this issue). The processing work is carried out in the documents' titles, abstracts, and keywords. Moreover, the ”Author Keywords” and the ”Relevant Keywords” may correspond to terms (two or more words).

The obtained lists of ”Author Keywords” and ”Relevant Keywords” includes many irrelevant words and terms to the intended study. Moreover, various versions of the same meaning (e.g. PV system, photovoltaic, photovoltaics, solar PV, etc.) also appear, which need to be aggregated. Therefore, it is necessary to process the lists of obtained ”Author Keywords” and ”Relevant Keywords” to eliminate the irrelevant terms/words and aggregate the ones with the same meaning.

A filtering process based on a custom semantic dictionary is adopted to obtain a clean set of ”Author Keywords” and ”Relevant Keywords”. A phyton code was developed to ensure the filtering procedure, substituting the Excel-based approach used by the authors in [15].

Nonetheless, before the filtering procedure, a text normalization procedure [23] is used to clean and consolidate the

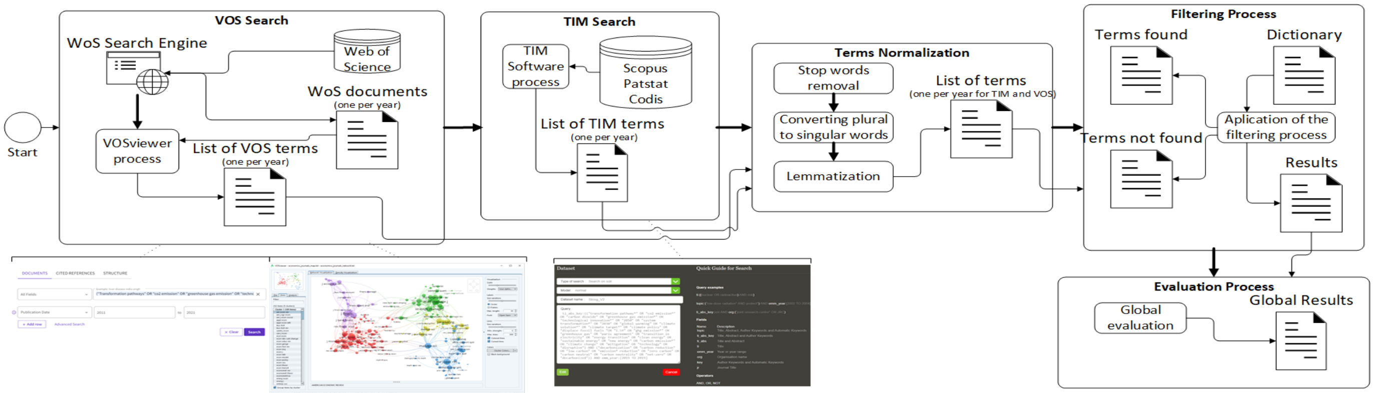

Fig. 1. Overall procedure for bibliometric analysis

retrieved lists of "Author Keywords" and "Relevant Keywords", improving the results. A phyton code was also written to implement the text normalization procedure, which is a novelty face to the methodology introduced in [15]. This procedure includes returning plural nouns to the singular form, turning inflected forms of verbs back to their stem, returning comparative adjectives to their primary form, and eliminating connectors and stop words. The unnecessary acronymous/abbreviatures are also eliminated, which contributes to the filtering process's success. However, once not all acronyms/abbreviations can be removed (e.g. the H_2 in the term "H2 storage"), a procedure has been implemented to avoid such situations by replacing the acronym/abbreviature with the text it represents.

After the normalisation process, a filtering procedure is used to, in each list of words returned by TIM and VOSviewer, separate the terms that make sense for the intended study from those that do not make sense. These last terms are discarded, while the others have their relevance/occurrence aggregated in the item (domain/technology) they fit. A previously defined semantic dictionary supports this filtering procedure, being an essential component to filter the retrieved keywords, aggregating several versions of the same word/term and discarding the irrelevant ones. The semantic dictionary is created using a semiautomatic approach [24], as shown in Fig. 2. Based on the Levenshtein approach [23], the automatic part is executed through a python based code, proposing new terms for each dictionary entry according to the already existent terms in the dictionary (a pre-dictionary is therefore needed). The automatic suggestions are then assessed by a specialist who decides on the ones that should integrate a revised dictionary version. Then, the new dictionary is used in the previously described filtering procedure with the retrieved lists of "Author Keywords" and "Relevant Keywords", allowing one to obtain a list of not found terms. Based on this list and on the specialist knowledge, some new terms may be added to the current dictionary. The process can then be repeated, starting with the automatic part of the algorithm. The number of repetitions is established by considering a criterion deemed appropriate. Therefore, the dictionary's construction involves some iterations to ensure a better set of terms/words that must incorporate its final version.

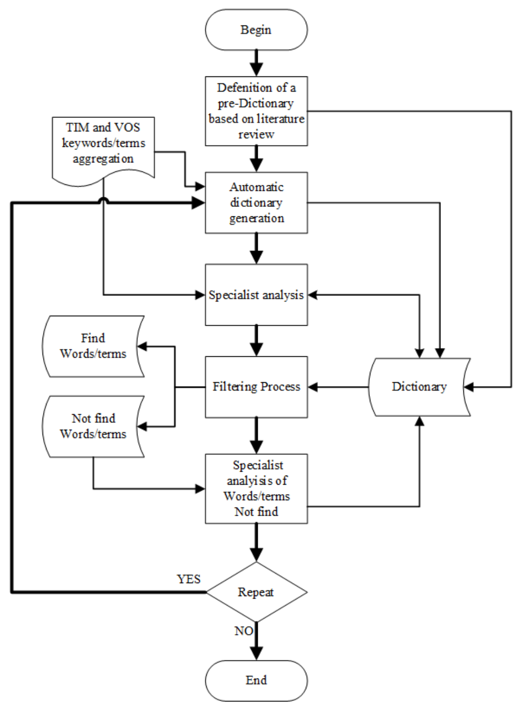

Fig. 2. Construction process of the semantic dictionary

Following the filtering process, an approach based on ranking the results from TIM and VOS is adopted. This procedure, implemented to create a fair comparison among domains of technologies, is also a novelty face to the methodology introduced in [15]. As previously explained, the relevance value retrieved by TIM software should not be directly compared with the occurrence value obtained in the VOSviewer software. Therefore, a two-step ranking procedure is adopted. In the first step, the results retrieved by TIM and VOS software are ranked according to their relevance or occurrence, respectively. Thus, the most relevant term in TIM occupies position one on the rank, the second most relevant position two, and so on. A similar procedure is used for the case of VOS software, but

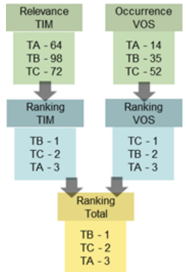

Fig. 3. Global ranking obtention

the rank is based on the occurrence. In the second step, a global rank is defined based on the sum of the values of the individual ranks, as shown in Fig. 3. Note that the term that occupies the first position in the final rank is the one with the lower sum of these values.

IV. Application

The methodology proposed in the previous section was applied considering the search string presented below. This string is an evolution of the one used in [15], being more robust and comprehensive, accounting for trying to reconcile the particularities of the WoS and TIM search engines (for example, the consideration of words in the plural or singular form is automatic in TIM, but not in WoS). Moreover, a more exhaustive literature review of energy sector decarbonisation allowed adding additional decarbonisation-related terms to the string. The search string has two parts linked through a logical AND, forcing each document in the retrieved dataset to contain at least one of the terms of both string parts. Therefore, the string was designed to capture the most relevant domains of technologies being studied to decarbonise.

Algorithm 1 String design

((”Transformation pathways” OR” co2 emission” OR” greenhouse gas emission” OR” technological innovations” OR” 2050” OR” System transformation” OR” 2030” OR” global warming” OR” climate solution” OR” climate target” OR” climate policy” OR” displace fossil fuels” OR” 1.5º” OR” ghg emission” OR” green house gas” OR” Paris Agreement” OR” transition in electricity” OR” energ transition” OR” clean energ*” OR” sustainable energy” OR” new energ*” OR” carbon emission” OR” climate change” OR” mitigation” OR” technology” OR” disruptive”) AND (” decarbonization” OR” carbon reduction” OR” low carbon” OR” emission reduction” OR” zero carbon” OR” decarbonized”))*

Fig. 4 evidences the increasing number of documents respecting the search string retrieved by TIM and WoS over 2011 to 2021. A total of 87212 documents were retrieved, 59411 by TIM and 27801 by WoS.

The bibliometric analysis done by TIM and VOSviewer software retrieved 11 sets (one per year) of ”Relevant Keywords” and their ”relevance” (TIM) and another 11 sets of ”Author Keywords” and their occurrence (VOSviewer).

Fig. 5 shows the number of keywords retrieved by each text-mining software. A total of 793700 keywords (689075 from TIM and 104625 from VOSviewer) were obtained. Since the

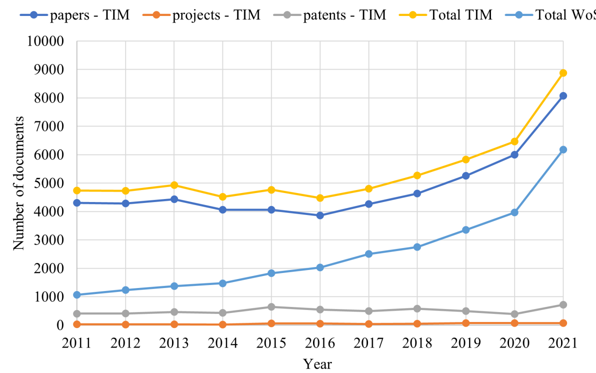

Fig. 4. Number of documents retrieved by TIM and WoS

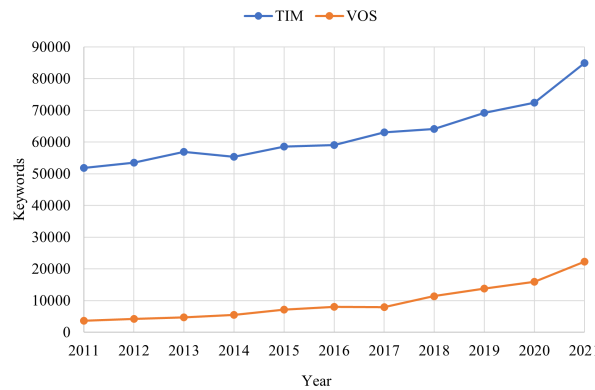

Fig. 5. Number keywords retrieved from TIM and VOSviewer

keywords were obtained by year, multiple repetitions of terms occur. After repeated keywords were eliminated, 196129 terms remained, 155778 from TIM and 40351 from VOSviewer. However, the total number of non-repeated keywords equals 176029, since there were also some repeated terms in the sets of keywords returned by TIM and VOSviewer.

As explained in section III, the retrieved keyword sets were subjected to a text-processing procedure.

The construction of the semantic dictionary started with the definition of an initial set of 102 terms related to technologies/domains of technologies for decarbonization. This initial dictionary was based on a preliminary bibliographical review, considering the authors' knowledge in the area. Then, the procedure described in section III was used.

Fig. 7 shows the evolution of the number of terms included in the dictionary. It is important to note that the terms returned by TIM and VOSviewer referring to only two years (2021 and 2020, respectively) were considered in the first seven iterations. The first Levenshtein ratio was equal to 0.75 and then increased by 0.05 until the maximum value of 0,9. The following iterations considered all the terms returned by TIM and VOSviewer for the study period (11 years - 176029 terms). This fact justifies the variation from the seventh to the eighth iteration. Moreover, the Levenshtein ratio was redefined as equal to 0.8 and increased by 0.05 on the tenth iteration. The

1	renewable energy system	renewable energy	hybrid renewable energy system	...
4	photovoltaic	photovoltaic energy	photovoltaic panel	...
13	perovskite solar cell	perovskite silicon tandem solar cell	perovskite	...
16	solar chimney power plant	solar chimney	hybrid solar chimney power plant	...
22	photovoltaic thermal	building photovoltaic thermal	photovoltaic thermal collector	...
28	building integration solar based te	solar thermal integration	solar energy integration	...
38	ocean	ocean energy	ocean renewable energy	...
55	biomethane	biochemical methane potential	biomethanation	...
69	biomass	crop production	wood production	...
104	nuclear fission	fission power system	fission energy	...
115	renewable hydrogen production	sustainable hydrogen production	biohydrogen	...
139	hydrogen turbine	hydrogen engine	hydrogen internal combustion engine	...
162	electricity transmission	inter regional power transmission	interprovincial electricity transmission	...
175	air source heat pump	air heat pump	air water heat pump	...

Fig. 6. Extract of dictionary

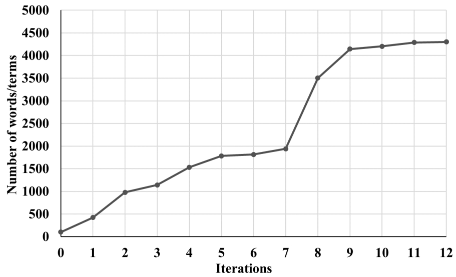

Fig. 7. Dictionary evolution

iterative process was stopped when the list of not found terms, resulting from the filtering process, did not contain terms to be added to the dictionary with occurrence greater than three or relevance greater than 5. The final dictionary has 4300 terms divided by 426 sets with similar semantic meanings. Note that this dictionary allows for a broader accounting of technology domains of technologies than the one adopted in [15]. For example, a global domain was considered regarding energy storage, and particular subdomains was assumed for electrical and thermal storage. Fig. 6 shows an extract of the obtained semantic dictionary.

After obtaining the semantic dictionary, the information obtained using text-mining software was processed. The previous 426 groups were aggregated into 38 technologies domains. Fig. 8 and 9 show the per-year relevance/occurrence for the 20 first-ranked domains of technologies presented in fig.

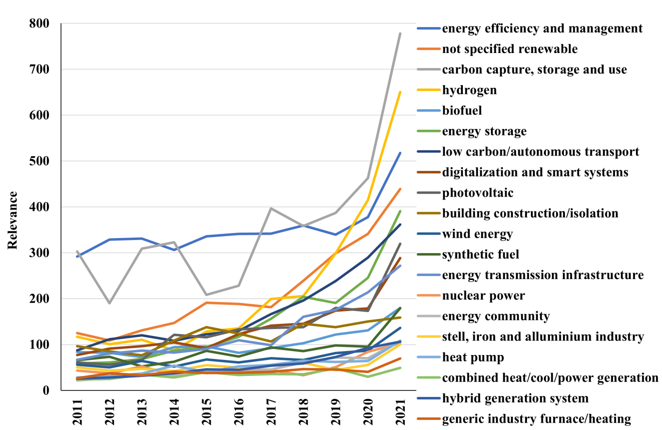

Fig. 8. Relevance values of the 20 first-ranked domains

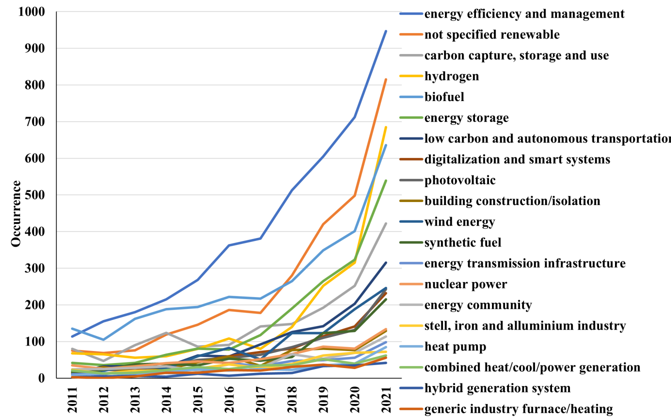

Fig. 9. Occurrence values of the 20 first-ranked domains

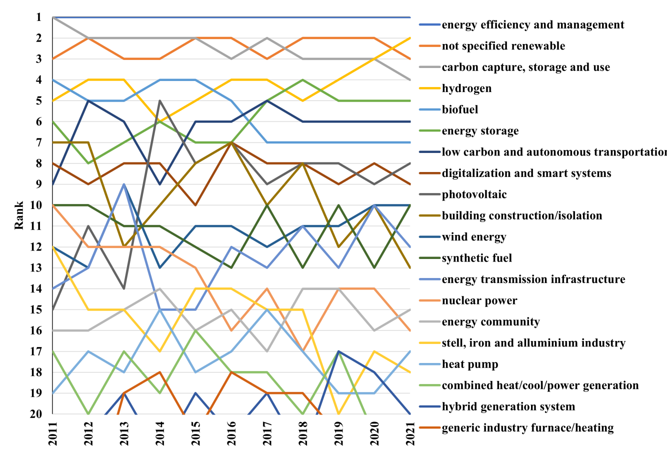

Fig. 10. Rank of technologies domains

10. A general increase in relevance/occurrence materialises for all domains. However, the most relevant domains do not necessarily have a higher occurrence. The methodology described in section III allows for defining a global ranking, as fig. 10 shows. Concerning relevance/occurrence values, energy efficiency/management, hydrogen technology, generic renewable energy, carbon capture, energy storage and biofuels appear as the six better-ranked domains.

The normalisation of the relevance/occurrence values, dividing the individual values by their sum, allows an understanding of the relative weight of each technology domain over time, as shown in figs. 11 and 12. The information from these figs. makes it possible to perceive the domains that, in the last years, have seen a more accelerated increase in the scientific community's interest. This is the case of hydrogen, carbon capture and energy storage domains in the last three years. On the other hand, it is also possible to perceive which domains have lost relative interest, such as energy efficiency and management, biofuel and low-carbon transports. Note that a lower relative relevance/occurrence value does not necessarily mean reducing the absolute relevance/occurrence values. For example, in the case of energy efficiency and management, it is possible to notice an increase in relevance/occurrence values

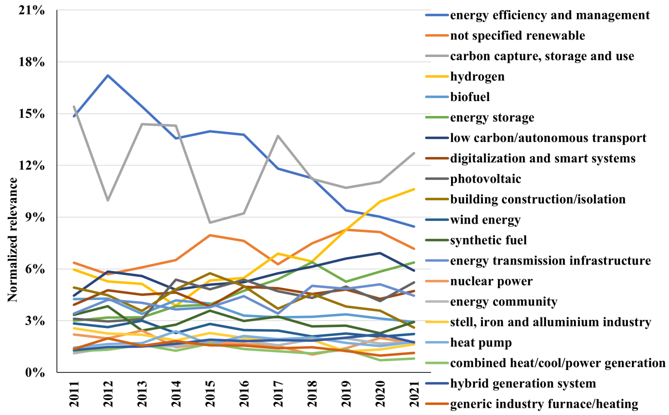

Fig. 11. Normalized relevance

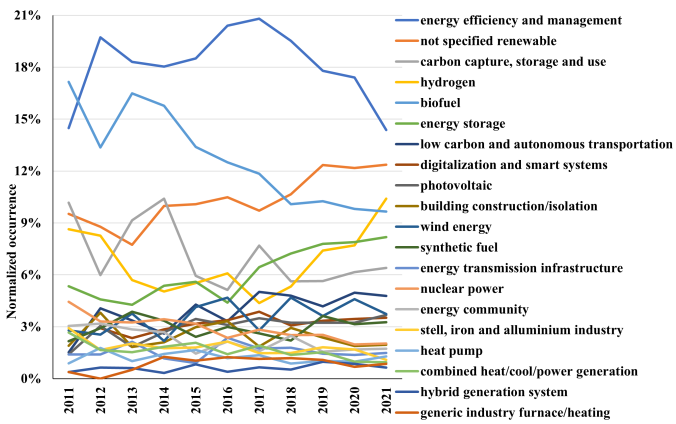

Fig. 12. Normalized occurrence

in recent years (Figs. 8 and 9), despite the decrease in relative relevance/occurrence values (normalised). In practice, this means that other technology domains have experienced a more accelerated increase in their absolute relevance/occurrence value.

To better understand how the scientific community's interest in the various domains of technologies has evolved, the 9-year period between 2013 and 2021 was divided into three-year periods. For each of these periods, the average value of relevance/occurrence for each domain was calculated. The calculated average value allows mitigating the effects of annual variations that tend to occur in the relevance/occurrence values. Fig. 13 shows the variation on the average value of relevance/ocurrence.

Analysing the information, it is possible to verify the existence of technology domains with a significant increase in their relevance/ocurrence in the 2019 to 2021 period compared to the 2016 to 2018 period. In the case of relevance (TIM), combined photovoltaic/thermal generation, electrification, hydrogen, agricultural sector, geothermal energy and nuclear power stand out. Concerning occurrence (VOSviewer), electrification, hydrogen, hybrid generation system, geothermal energy, synthetic fuel and solar power/steam generation emerge as the six domains with higher increase. Note that electrification, hydrogen and geothermal energy domains materialise in both the TIM and the VOSviewer. As previously anticipated, energy efficiency and management, biofuel and low carbon transports present a significantly lower increase in their relevance/occurrence values in the last three years period.

V. Conclusion

In this paper, an improved methodology based on the use of two text-mining tools is introduced and used to identify domains of decarbonisation technologies that emerge from the scientific literature, research projects and patents. The application of the developed methodology made it possible to define 38 domains of decarbonisation technologies that aggregate 426 subdomains. The relevance (TIM) and occurrence (VOSviewer) values associated with each domain were assessed. The result showed increased values for most domains from 2011 to 2021. Once the relevance and occurrence values are not directly comparable, a rank procedure was used to obtain an indicator that internalises each domain's relevance and occurrence. Energy efficiency/management, hydrogen technology, generic renewable energy, carbon capture, energy storage and biofuels appear as the six better-ranked domains. To realise the domains of technologies that have experienced a more accelerated increase in their absolute relevance/occurrence value in 2019 to 2021, the absolute values of relevance/occurrence were normalised. The hydrogen, carbon capture and energy storage domains stood out in this context. On the other hand, energy efficiency/management, biofuel, and low-carbon transport presented a lower increase. However, the paper explains that a lower relative relevance/occurrence value does not necessarily mean reducing the absolute relevance/occurrence values. Another interesting result is obtained when comparing the relevance/occurrence values between 2019 to 2021 and the 2016 to 2018 periods. This procedure reveals that electrification, hydrogen and geothermal energy materialise as the domains with a more significant increase in the relevance and occurrence values. In the last three years, energy efficiency and management, biofuel and low carbon transports present a significantly lower increase in their relevance/occurrence values. Despite all care taken in defining the search string and the semantic dictionary, there are limitations in the proposed methodology. One of these limitations is that the results can be influenced by the databases used. However, this effect is mitigated using two different bibliometric analysis tools, which use diverse databases. Another limitation results from the aggregation process, which is not straightforward since some technologies can be classified into multiple domains.

In future work, we intend to characterise the domains found regarding their technological readiness level, costs, and potential for decarbonisation. This characterisation will support the definition and comparison of decarbonisation strategies.

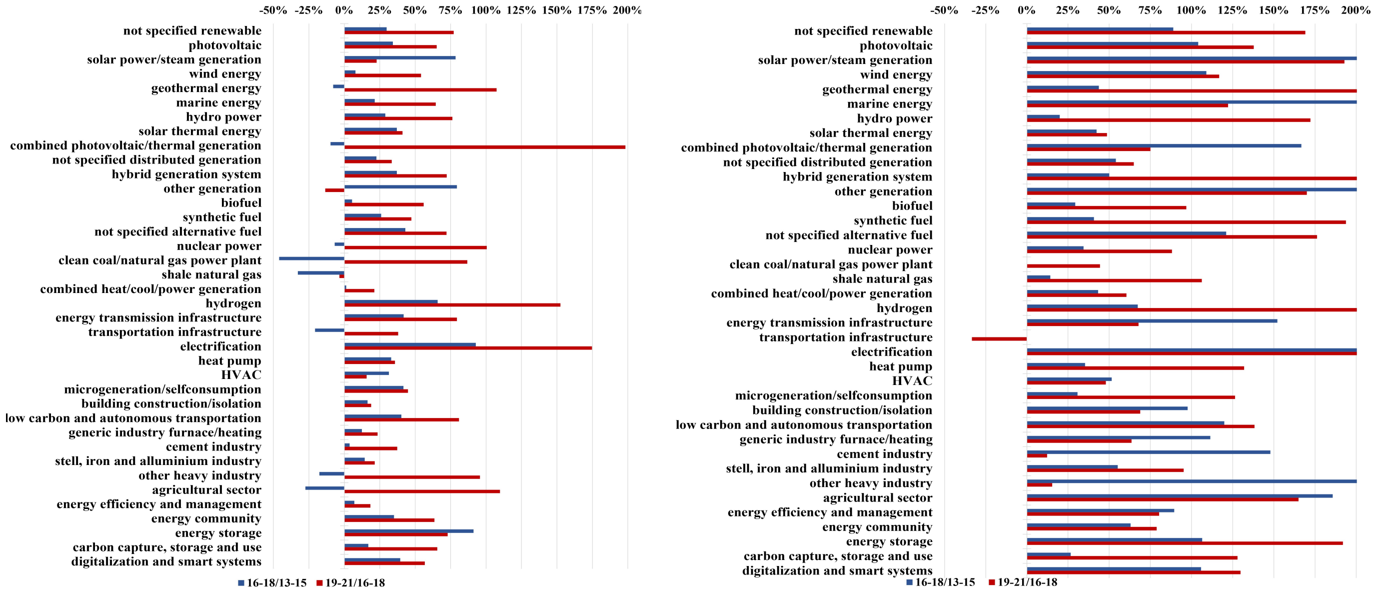

Fig. 13. Variation of the average value of relevance (left) and occurrence (right)

References

[1] Switzerland, "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report Addresses the Most Up-To-Date Physical Understanding of the Climate System and Climate Change, Bringing Together the Latest Advances in Climate Science. Geneva," 2021. [Online]. Available: https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/

[2] M. M. Miralles-Quirós and J. L. Miralles-Quirós, "Decarbonization and the Benefits of Tackling Climate Change," *International Journal of Environmental Research and Public Health*, vol. 19, no. 13, p. 7776, 2022.

[3] R. Zhang and T. Hanaoka, "Cross-cutting scenarios and strategies for designing decarbonization pathways in the transport sector toward carbon neutrality," *Nature Communications*, vol. 13, no. 1, p. 3629, 2022.

[4] Q. Zhu, B. D. Leibowicz, J. W. Busby, S. Shidore, D. E. Adelman, and S. M. Olmstead, "Enhancing policy realism in energy system optimization models: Politically feasible decarbonization pathways for the United States," *Energy Policy*, vol. 161, p. 112754, 2022.

[5] A. Dirnaichner, M. Rottoli, R. Sacchi, S. Rauner, B. Cox, C. Mutel, C. Bauer, and G. Luderer, "Life-cycle impacts from different decarbonization pathways for the European car fleet," *Environmental Research Letters*, vol. 17, no. 4, p. 044009, 2022.

[6] F. Plazas-Niño, N. Ortiz-Pimiento, and E. Montes-Páez, "National energy system optimization modelling for decarbonization pathways analysis: A systematic literature review," *Renewable and Sustainable Energy Reviews*, vol. 162, p. 112406, 2022.

[7] A. Ehsan and R. Preece, "Quantifying the impacts of heat decarbonisation pathways on the future electricity and gas demand," *Energy*, vol. 254, p. 124229, 2022.

[8] S. Garimella, K. Lockyear, D. Pharis, O. E. Chawa, M. T. Hughes, and G. Kini, "Realistic pathways to decarbonization of building energy systems," *Joule*, vol. 6, no. 5, pp. 956–971, 2022.

[9] H. C. Lau, "Decarbonization roadmaps for ASEAN and their implications," *Energy Reports*, vol. 8, pp. 6000–6022, 2022.

[10] ——, "Evaluation of Decarbonization Technologies for ASEAN Countries via an Integrated Assessment Tool," *Sustainability*, vol. 14, no. 10, p. 5827, 2022.

[11] H. Haggi, P. Brooker, W. Sun, and J. M. Fenton, "Hydrogen and Battery Storage Technologies for Low Cost Energy Decarbonization in Distribution Networks," *arXiv*, 2022.

[12] K. J. Holmes, E. Zeitler, M. Kerxhalli-Kleinfield, and R. DeBoer, "Scaling Deep Decarbonization Technologies," *Earth's Future*, vol. 9, no. 11, 2021.

[13] M. Ozkan, A.-A. Akhavi, W. C. Coley, R. Shang, and Y. Ma, "Progress in carbon dioxide capture materials for deep decarbonization," *Chem*, vol. 8, no. 1, pp. 141–173, 2022.

[14] W. He, M. King, X. Luo, M. Dooner, D. Li, and J. Wang, "Technologies and economics of electric energy storages in power systems: Review and perspective," *Advances in Applied Energy*, vol. 4, p. 100060, 2021.

[15] P. M. Costa, P. Tomé, B. Almeida, and N. Bento, "A Bibliometric based Analysis to Identify Promising Domains of Decarbonisation Technologies," *2022 7th International Conference on Smart and Sustainable Technologies (SpliTech)*, vol. 00, pp. 1–6, 2022.

[16] M. Reza, M. Mannan, S. B. Wali, M. Hannan, K. P. Jern, S. Rahman, K. Muttaqi, and T. M. I. Mahlia, "Energy storage integration towards achieving grid decarbonization: A bibliometric analysis and future directions," *Journal of Energy Storage*, vol. 41, p. 102855, 2021.

[17] C. Camarasa, C. Nägeli, Y. Ostermeyer, M. Klippel, and S. Botzler, "Diffusion of energy efficiency technologies in European residential buildings: A bibliometric analysis," *Energy and Buildings*, vol. 202, p. 109339, 2019.

[18] K. Jiang and P. Ashworth, "The development of Carbon Capture Utilization and Storage (CCUS) research in China: A bibliometric perspective," *Renewable and Sustainable Energy Reviews*, vol. 138, p. 110521, 2021.

[19] R. Raman, V. K. Nair, V. Prakash, A. Patwardhan, and P. Nedungadi, "Green-hydrogen research: What have we achieved, and where are we going? Bibliometrics analysis," *Energy Reports*, vol. 8, pp. 9242–9260, 2022.

[20] R. W. Wimbadi and R. Djalante, "From decarbonization to low carbon development and transition: A systematic literature review of the conceptualization of moving toward net-zero carbon dioxide emission (1995–2019)," *Journal of Cleaner Production*, vol. 256, p. 120307, 2020.

[21] X.-C. Meng, Y.-H. Seong, and M.-K. Lee, "Research Characteristics and Development Trend of Global Low-Carbon Power—Based on Bibliometric Analysis of 1983–2021," *Energies*, vol. 14, no. 16, p. 4983, 2021.

[22] L. Wang, Y.-M. Wei, and M. A. Brown, "Global transition to low-carbon electricity: A bibliometric analysis," *Applied Energy*, vol. 205, pp. 57–68, 2017.

[23] N. Desai and M. Narvekar, "Normalization of noisy text data," *Procedia Computer Science*, vol. 45, pp. 127–132, 2015, international Conference on Advanced Computing Technologies and Applications (ICACTA). [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050915003403

[24] Q. Deng, M. Hine, S. Ji, and S. Sur, "Inside the black box of dictionary building for text analytics: A design science approach," *Journal of International Technology and Information Management*, vol. 27, 01 2019.