energies

Article

The New Method for Analyzing Technology Trends of Smart Energy Asset Performance Management

Nguyen Thanh Viet [1,2,*] and Alla G. Kravets [1,3,*]

[1] CAD&RD Department, Volgograd State Technical University, 400005 Volgograd, Russia
[2] Training Department, Pham Van Dong University, Quang Ngai 570000, Vietnam
[3] Institute of System Analysis and Management, Dubna State University, Moscow Region, 141982 Dubna, Russia
* Correspondence: vietqn1987@gmail.com (N.T.V.); agk@gde.ru (A.G.K.)

Abstract: The development of emerging technologies not only has recently affected current industrial production but also has generated promising manufacturing opportunities that impact significantly on social and economic factors. Exploring upcoming renovation tendencies of technologies prematurely is essential for governments, research and development institutes, and industrial companies in managing strategies to achieve dominant advantages in business competitiveness. Additionally, the prospective changes, the scientific research directions, and the focus of technologies are crucial factors in predicting promising technologies. On the other hand, Industry 4.0 revolutionizes standards and models by accompanying significant technology developments in numerous sectors, including the sector of Smart energy. Moreover, asset performance management is always a prominent topic that has attained prevalence over the last decade because numerous challenges force all industrial companies to optimize their asset usability. However, to the best of our knowledge, no study reported an analysis of technology trends of asset performance management in the Smart energy sector by using proper data mining methods. Hence, this paper aims to fill in this gap and provide an analysis of technology trends of asset performance management in the Smart energy sector by structuring and exploring research subjects, considering problems, and solving methods with numerous experiments on scientific papers and patent data.

Keywords: energy; asset performance management; data mining; technology analysis; patent clustering

check for updates

Citation: Viet, N.T.; Kravets, A.G. The New Method for Analyzing Technology Trends of Smart Energy Asset Performance Management. *Energies* **2022**, *15*, 6613. https://doi.org/10.3390/en15186613

Academic Editor: Wen-Hsien Tsai

Received: 3 August 2022
Accepted: 7 September 2022
Published: 9 September 2022

1. Introduction

The evolution of topics in science and technology is always of great interest to scientists, firms, and governments for many decades because emerging technologies can radically change the condition both in the world markets and in a separate enterprises [1,2]. These may be cloud solutions, big data management, industrial Internet of Things, complex system modeling, and analytics. Individually, they provide businesses with the ability to strategically plan, predict, and optimize operations. Together, these technologies form a powerful toolkit that allows us to reduce costs and increase productivity.

Accordingly, practitioners and managers pertaining to technological innovations are continuously searching for means and approaches to make appropriate strategic decisions related to research and development administration, new process technologies, novel product development, commercialization, services, optimization, and investment in emerging technologies. Innovation analysis and forecast can be helpful in the process of making these decisions, such as seeking knowledge about: estimation of technology life cycle by specifying development advancement level, growth rate, and status of other related technologies; (b) acceptance of innovations in the economic and social context with numerous influences; (c) market expectations and product value chains [3].

If companies lack methods to successfully specify the potential industrialization and dissemination of emerging technologies to establish more valuable products and services, then the power of companies in scientific research and development will not be converted into technical commercialization advantages. Hence, with numerous existing emerging technologies, the method of how to select and evaluate precisely and prematurely is absolutely important for management departments and businesses [4].

At the same time, one of the urgent problems is asset performance evaluation in the process of managing industrial enterprises. In nowadays economy, assets are an expensive commodity that forms the magnitude of the enterprise market value as a property complex. In comparison with asset operations, maintenance is a crucial part of leveling the energy cost, given the practical restrictions. Operation and Maintenance (O&M) is admitted as one of the substantial proportions in the total energy cost, usually adding up to 30% of the lifecycle costs [5]. Hence, the selection of maintenance strategies impacts the safety, general efficiency, profit, and sustainability of assets. Currently, the most common approach to ensure the continued operation of equipment is preventive maintenance. It is carried out at regular intervals to reduce the likelihood of failures. However, in most cases, this practice results in either over or under maintenance due to unpredictable situations, differences in equipment age, and operating conditions.

Subsequently, a remarkable number of techniques and methods have been adopted to support decision-making in asset management and strategic planning. Condition monitoring tools are frequently applied due to fault identification ability, performance improvement, and evaluation. Asset Performance Management (APM) is an approach that prioritizes business objectives in addition to traditional asset reliability and availability objectives. APM refers to the decisions that the asset owners have to make to enhance asset availability, minimize risks related to asset operation, and optimize the total expense of asset maintenance. APM has become the main tool for the digital transformation of industrial companies. Risk management tools, appropriate data processing, and advanced analytics via proper technology adoption are needed to form effective maintenance.

The Smart energy sector is the center of the modern economy, wherein an increasing amount of electronic equipment demands more energy [6]. The noticeable increase in energy prices impacts not only organizations and countries but also international policy [7]. The advancement of various strategies and digital technologies in the Smart energy sector can also support the increasing energy demand [8] and can provide the transformation towards more sustainable and renewable, or nuclear technologies [9]. Concurrently, the significance of Smart energy technologies is more and more increasing. Digital technologies are providing new income chances, shifting cost efficiency, and driving business paradigms in Smart energy.

For this reason, the comprehension ability to exist technology trends, especially in APM of the Smart energy sector, is totally crucial for small and medium enterprises, because it helps intensify their competitiveness. While understanding and identifying the market technology tendencies become more and more essential, existing studies have examined those trends by analyzing scientific papers and patent data for many years. Nevertheless, to the best of our knowledge, no study has reported an analysis of technology trends of APM in the Smart energy sector by using proper data mining methods. Thus, the goal of our paper is to provide a method for identifying the popular investigating problems, core research directions, and understanding the technology trends on APM in the Smart energy sector by analyzing scientific papers and patent data sources from famous platforms.

2. Materials and Methods

The technology trend is assessed as a steadily growing practical technology area with a specific pattern, which has been existing for a particular time period [10]. Numerous approaches have been proposed recently to explore patterns and analyze and predict technology trends. One of these approaches that various scholars adopted is investigating the pattern of relationships between science and technology, which has been existing for many decades and continues to be the topic of lengthy discussion within the academic world. A large amount of empirical results has proved that science and technology interact in various ways; moreover, the interdependencies between them have been increasing in recent years [11].

Specifically, basic investigation in science provides the foundation for technological advancement because science is regarded as the seeds of technologies in the innovation model. Hence, substantial and advanced knowledge about technology developments can be attained by analyzing scientific publications and patents [12]. One of the popular methods for analyzing innovations and technologies is bibliometrics, which is the statistical analysis of text, data, and information, usually in the form of publications such as scientific journals, books, articles, conference proceedings, and patent applications. When analyzing technologies, the sources of publications are usually scientific journals and patents [1,13].

Additionally, patents provide a source of modern and reliable information for the disclosure of technological information and development. By analyzing the technological information available in patents, one can better identify and understand the path of technological evolution and be able to identify technology development trends with the help of experts in the field. Thus, researchers have begun to use patent data to analyze and study these technological trends [14].

For this reason, we propose a new method to analyze the technology trends of APM in the Smart energy sector by combining these steps (Figure 1).

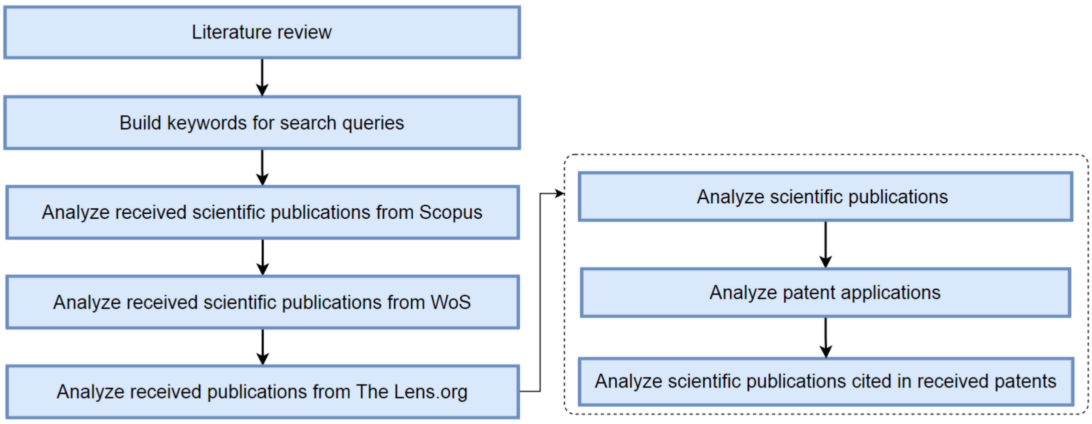

Figure 1. The overall procedure of the proposed method.

2.1. Literature Review

Firstly, the authors started by searching for relevant papers, which are also investigating the problems of analyzing APM technologies in the Smart energy sector from the famous bibliographic database Scopus. The search query was: TITLE-ABS-KEY (energy AND asset AND performance AND management AND review), which returned 113 documents. Then, after filtering by some meaningful keywords: "Asset Management", "Energy Efficiency", "Maintenance", "Operation And Maintenance", "Condition-based Maintenance", "Assets Management"; we obtained 34 papers. Then, the authors examined the title and abstract of the papers in this collection; there were only a few relevant papers. In addition, none of these studies utilized data mining methods for the analysis of technology trends for APM in the Smart Energy sector; however, some noteworthy results of these studies are presented in the following paragraphs.

In the study [15], Bermejo et al. presented a comprehensive review of six intelligent optimization modeling techniques, which were applied to condition-based maintenance for photovoltaic renewable energy plants, utilizing energy production prediction as the target variable. In the area of soft computing, intelligent optimization modeling methods consist of many significant techniques in machine learning that strive to create helpful business knowledge by processing and analyzing raw data. Additionally, due to the fact that data accumulation is becoming very large and is difficult to be processed by normal database management tools, hence modern tools for managing big data have to be under consideration. Their results showed that data mining methods, i.e., support vector machines, artificial neural networks, and decision trees (random forest, boosting), are becoming more and more pertinent since they appear in 61% of the total studies.

Likewise, in the article [16], Wan described the challenges, business drivers, and innovations in the industry to enhance the reliability of the power grid due to APM. The author highlighted the importance of digitization in providing data-driven and accurate decisions via gathering information from diverse organization data repositories. Digital technologies such as the industrial internet, mobile tools, predictive analysis, and big data are some of the most impactful technologies.

Similarly, Biard and Nour [17] explained how industry 4.0 could enhance the APM of electrical networks from a worldwide viewpoint. After reviewing the literature, the authors stated that once appropriately managed and structured, big data are the fundamental implementation of technologies and state-of-the-art tools in electrical networks. The data originated from smart grids in electrical networks and accumulated for many years holding the properties of big data. Concurrently, there are great possibilities for the Internet of Things (IoT) combined with comprehensive analysis, modeling, simulation, and highly relevant information. They optimize the intervention frequency, maximize asset performance, and facilitate the control of reliability-related risks, hence minimizing the usage of resources by supporting decision-making processes.

Moreover, Rinaldi et al. [18] reviewed the existing studies and modern approaches for condition monitoring and O&M planning of offshore renewable energy farms, with attention to the offshore wind sector. The offshore wind sector has been growing notably in recent years, amounting to over 10% of worldwide wind installations. Prior experience with onshore wind turbines, in combination with remarkable investments, has made the offshore wind sector become one of the most profitable and feasible ways of producing electricity. The important meaning of this study is to enable the integration of modern O&M techniques in the premature period of offshore wind installations, adopting fault prognosis/diagnosis techniques, industrial condition-based maintenance deterioration models, and also the latest signs of progress in artificial intelligence, data mining, and robotics. Apart from reviewing existing papers, the present study also explored upcoming technology tendencies and referred to other industrial sectors to examine the transferability of the considered methods.

Further, as stated in the article [19], wind energy is forming much more portions of the global energy sector; however, the main obstacle preventing the wind energy industry from receiving greater investment is the high failure rate of large wind turbines; therefore, within the last two decades, compelling and significant investigation results concerning fault diagnosis, prognosis, and resilient control techniques for wind turbine systems, were analyzed comprehensively.

O&M of offshore wind turbines (OWT) is an essential part of the growth of offshore wind farms. The safety and efficiency of onsite maintenance, including complicated marine operations, depend on practical factors with negative environmental impacts. In order to solve these problems, the article [20] reviewed the latest studies about OWT maintenance, including selecting a strategy, optimizing schedules, onsite repair and operations, evaluation standards, recycling, and environmental issues. Specifically, the optimal schedule should investigate various topics consisting of decreasing downtime, enhancing system reliability, increasing revenue, and perceiving the collaboration among maintenance groups. The initial stage quantifies the matter, and the process is described by numerical deterministic and probabilistic models. Then, the scheduling issue can be transformed into an optimization problem with constraints and several cost functions.

At the same time, the study [21] showed that a sufficient monitoring strategy for wind turbines needs both the structural health monitoring of their load-bearing structural components (blades, tower, and foundations) and the condition monitoring of their mechanical elements. Thus, it includes both mechanical engineering subjects and civil fields. Some traditional and advanced non-destructive techniques (NDT) have been suggested for both application fields in the last few years. These include acoustic emissions, visual inspection, infrared thermography, ultrasonic testing, electromagnetic testing, oil monitoring, radiographic testing, and numerous other methods. These NDT even can be implemented by robots or unmanned aerial vehicles and can also be used for both isolated wind turbines of onshore/offshore wind farms. In this study, these NDT were reviewed extensively with more than 300 scientific articles, technical reports, and other documents in the last two decades (2000–2021).

The work [22] presented a review and classification of numerous structural health monitoring methods as strain measurement utilizing Fiber Bragg Gratings and optical fiber sensors, vibration-based method, active/passive acoustic emission method, ultrasonic methods, and thermal imaging method, relying on the latest studies and positive modern techniques. Since cost-effectiveness, comprehensiveness, and accuracy are the radical features in choosing the structural health monitoring method, the authors provided an extensive summarized review, including methods limitations/capabilities and sensor types. Regarding various types of sensors related to several structural health monitoring methods, the capabilities and limitations of represented methods were analyzed and nominated.

On the other hand, the transformation process of digitalization has already impacted various industry sectors and society in recent years. Especially in the Smart energy sector, numerous digital applications have been carried out recently. Nevertheless, scholars and practitioners are expecting a more radical and extensive change during the next decade. Proper comprehension of feasible digital applications, the associated benefits together with risks from the various perspectives of the affected stakeholders is critically important. Moreover, it is an essential part of the information for enterprises to develop and perform digital applications sustainably.

To achieve this, Lyu and Liu [23] stated that the growing usage of innovative digital and data-driven techniques is a significant element of “energy digitalization”, especially in the renewables sector. Three principal sections of digitalization are forming the Smart energy sector with innovative systems by integrating visualization and software solutions to facilitate the operators. By investigating the data of worldwide online job postings collected in the period 2010–2019, the authors showed that among the promising digital technologies (e.g., big data, artificial intelligence, robotics, blockchain technology, cloud computing, IoT, etc.), artificial intelligence was the most extensively utilized in the Smart energy sector.

Additionally, Weigel et al. [24] suggested a methodology to analyze digital applications holistically and also investigated prospective digitalization applications in the German electricity sector, encompassing the impacted stakeholders and the correlated benefits based on a structured literature review. Subsequently, three impactful subject areas containing various individual digital applications were recognized as categories of digital applications, i.e., “Customer orientation”, “Process optimization”, and “System balance”. The “System balance” applications mostly include applications in the fields of “smart market” and “smart grid”, which control actively and balance both production and consumption based on data-driven monitoring and forecast tools. The authors explored that these applications were the most examined among the collected documents. Furthermore, “Process optimization” applications either automate processes due to robotics or optimize processes relying on data analytics. The major benefits detected in the analyzed documents are cost reduction based on more effective processes and efficient impact on the system stability based on enhancing balance among production, grid capacity, and consumption.

2.2. Extracted Meaningful Keywords

After studying the above papers, the authors extracted the following keywords for further searching relevant documents and generating data analysis about technology trends of APM in the Smart energy sector. These keyword groups are:

- Maintenance: strategic planning, asset management, maintenance planning, condition-based maintenance, o&m planning, preventive maintenance, condition monitoring, fault diagnosis, prognosis, asset life cycle management, and risk management. The keywords in this group will be used for building the search query in bibliographic databases.

- Energy sector: renewable energy, photovoltaic plant, offshore renewable energy, marine renewable, electrical network, and smart grid.
- Algorithm, model: data, digital sensor, predictive model, digitization, digital twin, data mining, genetic algorithm, neural network, and evolutionary algorithm.
- Industry 4.0: artificial intelligence (AI), big data, machine learning, SCADA (supervisory control and data acquisition), cloud, Internet of Things (IoT), cyber-physical system, blockchain, augmented reality, and additive manufacturing.
- Hardware, robot: robotics, inspect-and-repair robot, crawler robot, inspection drone, autonomous vessel, sensor, and soft sensor.

3. Experiment Results

In this section, the authors reveal the results of the technology trends of APM in the Smart energy sector by analyzing data from various famous bibliometric databases, i.e., Scopus, Web of Science (WoS), and The Lens.

3.1. Searching and Analysis Results from the Scopus Database

By utilizing keywords in keyword groups of maintenance, the authors devised the following search query in order to retrieve articles in period 2010–2021 from the bibliometric database Scopus: TITLE-ABS-KEY ({asset performance management} OR {asset management} OR {o&m planning} OR {maintenance planning} OR {preventive maintenance}) AND TITLE-ABS-KEY ({risk management} OR {strategic planning} OR {condition-based maintenance} OR {condition monitoring} OR {fault diagnosis} OR {fault prognosis}) AND PUBYEAR > 2009 AND PUBYEAR < 2022 AND (TITLE-ABS-KEY (energy OR power) OR SUBJAREA (ener)) AND (LIMIT-TO (LANGUAGE, "English")). At the time of writing this paper (May 2022), the result collection includes 691 documents.

3.1.1. Cluster Analysis with VOSviewer Software

Furthermore, the downloaded scopus.csv file containing these above 691 documents is input for the VOSviewer software [25]. VOSviewer is a popular software instrument for visualizing and constructing bibliometric networks. For instance, these networks may include researchers, journals, publications, or extracted keywords, and they can be generated via co-authorship, direct citation, bibliographic coupling, or co-citation relations. Moreover, VOSviewer also provides the text mining ability to visualize and construct a co-occurrence network of the most relevant keywords extracted from a corpus of scientific articles.

More specifically, in the network visualization, keyword items are displayed by their label shaped in a circle (by default) or rectangle. In addition, the weight of a keyword item identifies the size of the label item and its circle area. In other words, the higher the weight of a keyword item is, the larger the label and its circle shape are. For some keyword items, their labels may not be shown in order to avoid overlapping problems. The color of a keyword item is identified by the cluster to which this item belongs. Moreover, by default, the top 1000 lines (links between keyword items) are shown, which represent the 1000 strongest links between items.

A result of the network visualization rendered from VOSViewer software is displayed in Figure 2. The distance between two keywords in the visualization displays approximately the relationship of these keywords due to co-occurrence links. In this manner, the closer two keywords are located to each other, the higher their correlation is. Likewise, the highest co-occurrence links between keywords are also displayed by the lines. The following configuration options were applied to generate the network: "Create a map based on bibliographic data", "Read data from bibliographic database files", "Type of analysis (Co-occurrence)", "Counting method (Full counting)", "Unit of analysis (All keywords)", "Minimum number of occurrences of a keyword (10)".

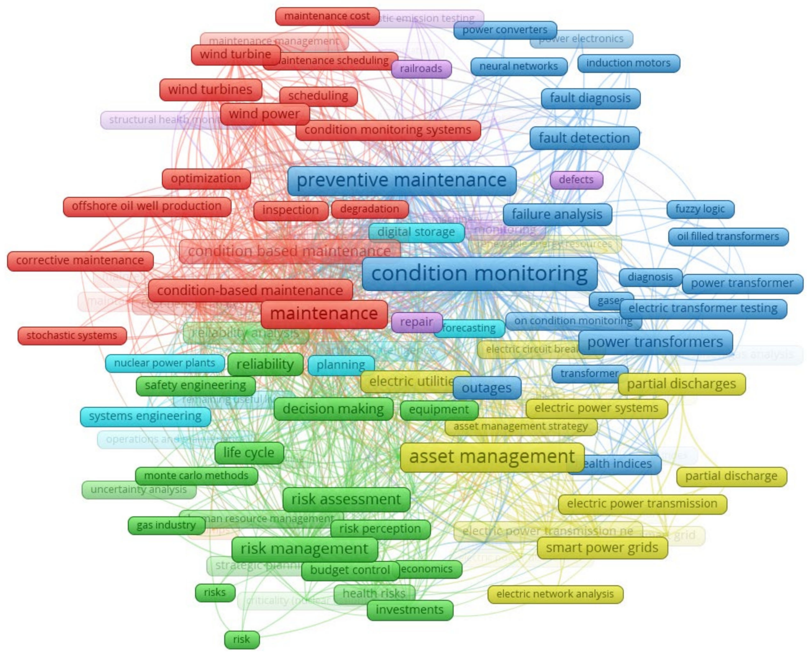

Figure 2. Network visualization of keywords from Scopus papers.

From the above figure, some main trends are extracted as follows:

- **Electricity**: electric utilities, partial discharges, smart power grids, electric power transmission networks, smart grid, power transformer, electric power systems, electric transformer testing, electric power distribution, electric switchgear, electric circuit breakers, induction motors, power electronics, and electric network analysis;
- **Wind**: wind power, wind turbines, turbine components, and offshore wind farms;
- **Oil, gas**: offshore oil well production, oil-filled transformers, gasoline, gases, gas industry, and dissolved gas analysis;
- **Nuclear power**: nuclear power plants, nuclear fuels, criticality (nuclear fission);
- **Algorithm, method**: artificial intelligence, decision support systems, digital storage, Markov processes, machine learning, neural networks, fuzzy logic, data acquisition, big data, and Monte Carlo methods;
- **Others**: renewable energy resources, vibration analysis, signal processing, railroads, railroad transportation, acoustic emission testing, and remaining useful lives.

3.1.2. New Method of Clustering Keyword Co-Occurrence Network

Firstly, we devise an algorithm for extracting keywords from text, which is described in more detail in Section 3.3.3. Herein, the keyword of the noun phrase is identified by the rule: the longest possible noun phrases sequence of many consecutive words within a sentence such that the last word in the sequence is a noun and each of the other words is either a noun or an adjective.

Secondly, the authors build a keyword co-occurrence network by adopting the above-mentioned keyword extraction algorithm, which extracts noun phrases from Scopus publications. Various co-occurrence analysis approaches have been developed recently to generate networks via adopting different aspects of analysis. For instance, word co-occurrence analysis employs the most significant keywords of publications to examine the conceptual structure of a field of study or a large collection of numerous documents. This is the only approach that adopts the text content of publications to generate a similarity measure; the others link publications indirectly through authorship, citations, or affiliations. Thus, the purpose of the analysis of keyword co-occurrence is to build a conceptual structure using a network of keywords to match and cluster them, extracted from the Titles, Abstracts, Author keywords, and Keywords plus in the publication dataset.

In particular, the matrix (network) of keyword co-occurrence can be obtained by the general formula:

$$W = A^T \times A \tag{1}$$

where $\boldsymbol{A}$ is a [Publication × Word] matrix, Word is the keywords extracted from Titles, Abstracts, Author keywords, and Keywords plus in Scopus publication dataset by the abovementioned algorithm for keyword extraction. Matrix $\boldsymbol{A}$ is a rectangular binary matrix where each row is a publication, and each column refers to an extracted keyword from the vocabulary of the entire publication dataset. The common element A_{ij} takes on a value according to the following logic:

$$A_{ij} = \begin{cases} 1, \ if \ publication \ i \ contains \ keyword \ j \\ 0, \ otherwise \end{cases} \tag{2}$$

Moreover, the sum of the jth column $A{+}j$ is the number of publications containing the keyword j. The sum of the ith row $Ai{+}$ is the number of unique keywords that appeared in publication i.

Therefore, the element W_{ij} indicates how many co-occurrences exist between the keywords i and j. The diagonal element W_{ii} is the number of publications containing the keyword i. The general element W_{ij} can be calculated using the following formula (where n is the number of publications):

$$W_{ij} = \sum_{k=1}^{n} A^T{}_{ik} A_{kj} \tag{3}$$

After extracting keywords from each publication, we obtain the result (with a minimum keyword occurrence threshold of 5) with 506 keywords from 691 publications. Then the matrix $\boldsymbol{A}$ ([Publication × Word]) has dimension [691 × 506]; thus, the co-occurrence matrix $\boldsymbol{W}$ has dimension [506 × 506], which is presented in Figure 3.

	1986-2012 ieee	1994-2012 ieee	accident prevention	acoustic emission	acoustic emission testing	advanced condition monitoring	anomaly detection	artificial intelligence	artificial neural network	asset condition	...	wide range	wind energy	wind energy sector	wind farm	wind farm operator	wind industry	wind power	wind turbine	wind turbine component	wireless sensor network
0	8	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	1	0	0	0
1	0	5	0	0	0	0	0	1	1	0	...	0	0	0	0	0	0	0	0	0	0
2	0	0	8	1	1	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	0
3	0	0	1	11	9	0	0	0	0	0	...	1	0	0	0	0	0	1	2	0	0
4	0	0	1	9	12	0	0	0	0	0	...	1	0	0	1	0	0	2	3	1	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
501	0	0	0	0	0	0	0	1	0	0	...	0	2	1	2	0	6	4	6	0	0
502	1	0	0	1	2	3	2	4	1	0	...	0	19	6	12	3	4	55	39	5	0
503	0	0	0	2	3	3	2	6	5	0	...	0	15	3	14	4	6	39	63	9	0
504	0	0	0	0	1	0	0	0	2	0	...	0	1	0	4	0	0	5	9	9	0
505	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	9

506 rows × 506 columns

Figure 3. The keyword co-occurrence matrix $\boldsymbol{W}$.

Then the co-occurrence matrix $\boldsymbol{W}$ is converted to a sparse matrix type to reduce the memory occupied when generating the keyword network analysis (network clustering), which has been used recently for identifying major research trends [1,26]. Network clustering analysis, also known as community detection, is an approach to specifying clusters of nodes that are densely connected to each other on the network, and revealing underlying hidden relationships between them.

In the paper [27], Lee et al. executed a comparative analysis of various network clustering algorithms to analyze their effectiveness, performances, and influence on derived clusters. The authors showed that the Louvain algorithm presented the best performance by processing time and modularity metrics. The modularity computes the density (value between −1 and 1) of edges inside clusters as compared to edges between various clusters. The modularity value is determined as follows:

$$Q = \frac{1}{2m} \sum\nolimits_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \tag{4}$$

where A_{ij} represents the number of connections between node i and j, c_i indicates the cluster to which node i is assigned, and

$$\delta(c_i, c_j) = \begin{cases} 1, \ if \ c_i = c_j \\ 0, if \ c_i \neq c_j \end{cases} \tag{5}$$

$$k_i = \sum\nolimits_j A_{ij} \ , \ m = \frac{1}{2} \sum\nolimits_{i,j} A_{ij} \tag{6}$$

Herein, we performed network clustering by the igraph package [28] in R. This package supports numerous functions for graph construction, centrality calculations, and network clustering. Some of the centrality measures are defined as follows:

Degree centrality C_D is the number of connections a node holds with others in a network.

$$C_D(p_k) = \sum\nolimits_{i=1}^{n} c(p_i, p_k) \tag{7}$$

where p_k is regarded as node p; n is the overall number of nodes; further:

$$c(p_i, p_k) = \begin{cases} 1, \ if \ p_i \ and \ p_k \ are \ linked \\ 0, \ otherwise \end{cases} \tag{8}$$

Betweenness centrality C_B computes how often a node p_k stays on the closest path among others in a network. A node with higher betweenness centrality shows that this keyword plays an essential role in the network in terms of the dissemination and linkage with other keyword nodes.

$$C_B(p_k) = \sum\nolimits_{i<j} \frac{s_{ij}(p_k)}{s_{ij}} \tag{9}$$

where s_{ij} is the number of closest paths from node i to node j; and $s_{ij}(p_k)$ is the number of closest paths from node i to node j which goes through the node p_k.

Eigenvector centrality (also known as prestige score or eigencentrality) is a measurement of the node impact in a network. The relative eigencentrality score $x(p_k)$ of node p_k can be defined as (where λ is a constant):

$$x(p_k) = \frac{1}{\lambda} \sum_{i=1}^{n} c(p_i, p_k) x(p_i) \tag{10}$$

For this reason, Table 1 represents the top-15 keywords by Betweenness centrality, and Eigenvector centrality. Herein some general keywords by the thematic terms of this paper were excluded, e.g., maintenance planning, maintenance strategy, reliability analysis, etc. Among the top-15 keywords, the most important ones are: power transformer, wind power, wind turbine, partial discharge, cost–benefit analysis, artificial intelligence, power system, electric power transmission network, smart power grid, offshore oil well production, etc.

Table 1. Top-15 keywords by Betweenness centrality and Eigenvector centrality values.

№	Keyword	Betweenness	Keyword	Eigenvector
1	power transformer	2144.5	electric utility	0.651
2	electric utility	1530.3	power transformer	0.634
3	wind turbine	1111.2	wind power	0.549
4	partial discharge	1057.4	wind turbine	0.548
5	wind power	1050.0	information management	0.543
6	cost–benefit analysis	836.0	cost–benefit analysis	0.521
7	artificial intelligence	819.9	artificial intelligence	0.502
8	power system	807.2	partial discharge	0.495
9	electric power transmission network	730.7	power system	0.483
10	smart power grid	688.1	electric power transmission network	0.482
11	smart grid	651.4	cost effectiveness	0.477
12	offshore oil well production	491.1	smart power grid	0.474
13	health index	463.4	smart grid	0.466
14	gas analysis	437.5	decision support system	0.443
15	digital storage	432.7	offshore oil well production	0.425

Then, employing the Louivan clustering algorithm in igraph divided 506 keyword nodes into 7 clusters, which can be examined in more detail in the Supplementary Spreadsheet S1 Keyword_clusters_with_centraility_values.xlsx (sheet tab Centrality_df). After that, the Gephi software is adopted to visualize these clusters in Figure 4. For convenient investigation, only 120 nodes with a minimum degree centrality of 86 are displayed. Specifically, the color of a keyword node displays the cluster (group) to which the node belongs, while the node size demonstrates its relative eigenvector centrality. In this manner, the bigger the keyword node is, the more impact this keyword has.

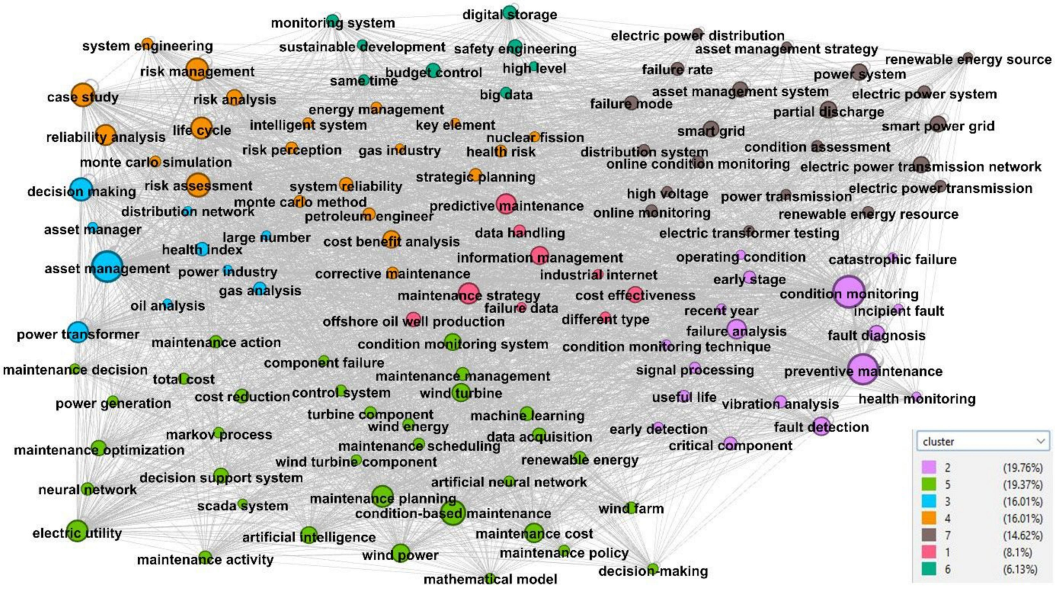

Figure 4. Cluster visualization of keyword co-occurrence network.

Basic description and thematic inference from keywords in each cluster are following:

Cluster 1. Keywords: offshore oil well production, industrial internet, monitoring technology, offshore technology, data analytics, petroleum industry, on-line partial discharge, offshore oil well, digital twin, pattern recognition, etc. This cluster covers that digital twin and data analytics technologies are mainly adopted for maintenance strategy and predictive maintenance in offshore oil well production (petroleum industry). Digital twins are considered as a paradigm integrating multi-physics modeling with data-driven analytics, which are adopted to simulate the life of their respective twin object [29].

Cluster 2. Keywords: vibration analysis, acoustic emission testing, power electronics, support vector machine, structural health monitoring, railroad transportation, power converter, real-time condition monitoring, smart sensor, wireless sensor network, roller bearing, etc. This cluster mentions numerous non-destructive testing techniques, which are adopted in real-time condition monitoring (CM) to monitor both environmental and operational parameters, especially in railroad transportation, power converter, induction motor, solar power generation, hydroelectric power plant, and thermal power plant. Some of these NDT include [21,30,31]:

- Vibration analysis. Utilizes accelerometers and micro-electromechanical or piezoelectric systems to monitor the gearbox, roller bearings, and drivetrain;
- Acoustic emissions testing. Detects potential defects and flaws during operation, such as impacts, excessive deformations, or cracks;
- Temperature measurement and infrared thermography (infrared imaging, image processing, thermal imaging). Detects anomalies or thermal fluctuations in both electrical and mechanical components;
- Partial discharge detection. The technique is employed to assess electrical insulation health. Partial discharge is best described as a part failure of an insulation system to resist the electrical field affecting it. This can be a consequence of poor workmanship, poor design, contamination, defective materials, or aging. Furthermore, we can achieve partial discharge detection by the combination of algorithms and monitoring that utilize big data received from network instrumentation sensors [32].

Moreover, this cluster also covers structural health monitoring using a wireless smart sensor network. A transition from NDT employing ultrasonic testing or transducers to permanent structural health monitoring systems with AI tools and a more sophisticated wireless smart sensor network has been prominent in recent years [33]. Structural health monitoring is a section of CM, which monitors the integrity of constitutive components in terms of Smart, functional sensors [22]. It even helps prevent catastrophic failures, with serious influence on the safety of workers and the environment [18].

Cluster 3. Keywords: power transformer, gas analysis, oil analysis, fuzzy logic, insulation system, electrical power system, transformer failure, plasma diagnostics, electric network analysis, chemical analysis, etc. This cluster mentions power transformer asset management by applying various testing methodologies: chemical methodology (keywords: chemical analysis, gas analysis, oil analysis, Duval triangle, dga result), electrical methodology (keywords: electric network analysis, frequency response analysis), Markov model [34,35], and fuzzy inference-based approach (fuzzy logic). Power transformers are the most critical and indispensable component in an electrical power grid system. Hence, the study [36] dedicates a comprehensive review of established electrical and chemical tests for detecting miscellaneous incipient faults in power transformers.

Cluster 4. Keywords: petroleum engineer, Monte Carlo method, nuclear fission, gas industry, probability distribution, environmental regulation, Bayesian network, Weibull distribution, electricity distribution network, hydroelectric power, etc. This cluster mentions reliability analysis, risk management, and strategic planning in the oil and gas industry (petroleum engineer) and Hydroelectric power plant by applying Monte Carlo simulation, Bayesian network, or Weibull distribution [37]. Moreover, Monte Carlo simulation and Bayesian network methods are also employed regularly in environmental risk assessment and CM of a nuclear power plant.

Cluster 5. Keywords: wind power, wind turbine, artificial intelligence, machine learning, renewable energy, neural network, wind energy, wind farm, Markov process, offshore wind farm, SCADA system, etc. This cluster mentions condition-based maintenance and maintenance planning in offshore wind farm (keywords: wind power, wind turbine, wind energy, wind farm, wind turbine component, offshore wind turbine, wind energy sector, wind farm operator, marine engineering) and nuclear power plant (keywords: nuclear fuel, nuclear energy, nuclear industry) by applying technologies of artificial intelligence, machine learning, neural network, Markov process, SCADA system (SCADA data), and data mining.

Cluster 6. Keywords: digital storage, big data, predictive analytics, fossil fuel power plant, condition-monitoring data, communication technology, electric power system protection, electric load flow, power distribution, data visualization, steam turbine, etc. This cluster mentions predictive analytics, condition-monitoring data, and electric power system protection in a fossil fuel power plant (steam turbine) by applying digital storage, big data, communication technology, and data visualization.

Cluster 7. Keywords: electric power transmission network, smart power grid, smart grid, online condition monitoring, renewable energy resource, electric transformer testing, electric power distribution, electric circuit breaker, HVDC power transmission, electric network parameter, etc. This cluster mentions asset management systems and condition assessment in Smart grid and electric power transmission network (high-voltage direct current—HVDC power transmission) by utilizing online condition monitoring technologies (keywords: on-line monitoring system, data collection, online system, real-time monitoring, online monitoring system, on-line monitoring).

3.2. Searching and Analysis Results from the Web of Science Core Collection Database

Similarly, after applying the search query, which can be found in the Supplementary Spreadsheet S1 Keyword_clusters_with_centraility_values.xlsx (sheet tab Search_queries), at the time of writing this paper, we received 328 results in the period 2010–2021 from bibliometric Web of Science Core Collection database. With the downloaded data, the top 30 most frequent categories in papers, which were assigned by WoS, are disclosed in Figure 5.

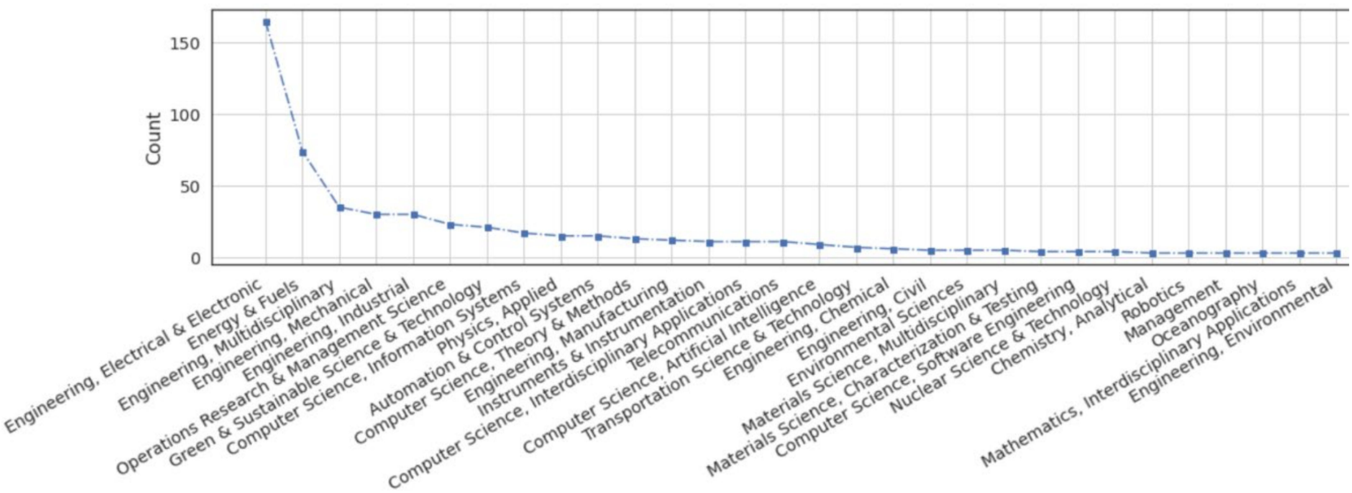

Figure 5. Top 30 most frequent categories in WoS papers.

Among these categories, the most prominent ones include "Engineering, Electrical & Electronic", "Energy & Fuels", "Operations Research & Management Science", "Green & Sustainable Science & Technology", "Computer Science, Information Systems", "Physics, Applied", "Automation & Control Systems", "Computer Science, Theory & Methods", etc.

Furthermore, the downloaded WoS.txt file containing these above 328 documents is input into VOSviewer in order to construct a density visualization network. In this density network, each point is assigned a color that demonstrates the density of items at that point, where the colors range from green to yellow to red. The higher the weights of the neighboring items and the higher the number of items in the neighborhood of a point are, the closer the color of the point is to red. In contrast, the smaller the amount of items in the neighborhood of a point and the lower the weights of the neighboring items are, the closer the color of the point is to green. In this manner, the item density visualization is shown in Figure 6. The following configuration options were applied to generate the network: "Create a map based on bibliographic data", "Read data from bibliographic database files", "Type of analysis (Co-occurrence)", "Unit of analysis (All keywords)", "Counting method (Full counting)", "Minimum number of occurrences of a keyword (3)".

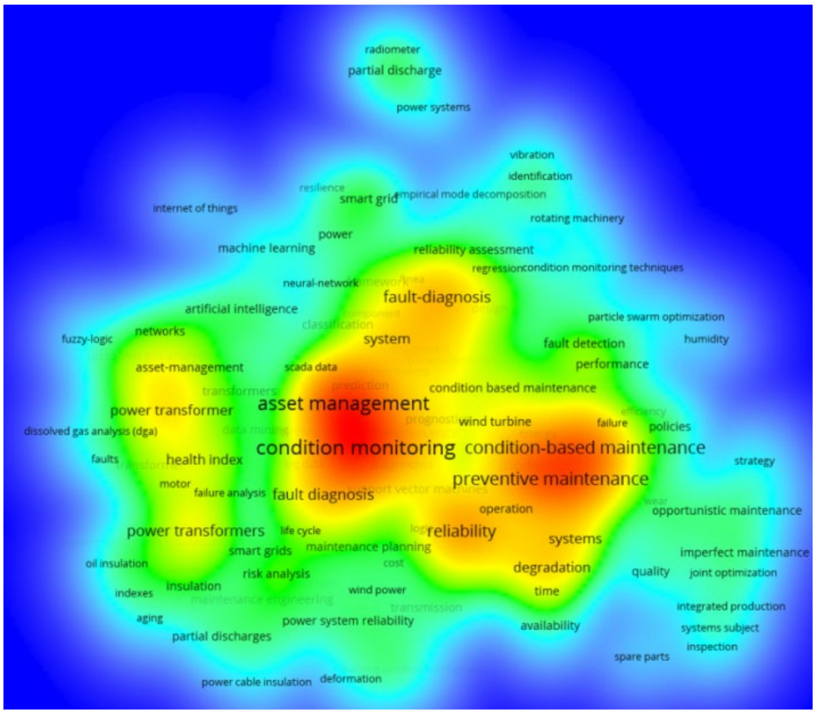

Figure 6. Density visualization of keywords from WoS papers.

From the above figure, some main trends are extracted as follows:

- **Electricity**: power transformers, smart grid, partial discharge, power transformer insulation, high-voltage techniques, power distribution, and power cable insulation;
- **Wind**: wind turbine, wind energy, wind power, and offshore wind energy;
- **Oil, gas**: dissolved gas analysis, oil, and oil insulation;
- **Algorithm, method:** health index, support vector machines, artificial intelligence, machine learning, remaining useful life, artificial neural network, genetic algorithm, data mining, SCADA data, fuzzy logic, big data, decision making, particle swarm optimization, FMEA (Failure Modes and Effects Analysis), data-driven, regression, decision tree, principal component analysis, and internet of things;
- **Others**: gearbox, induction motors, rotating machinery, dielectric response, radiometer, energy storage, infrared thermography, vibration analysis, frequency response analysis, sensitivity analysis, and humidity.

3.3. Searching and Analysis Results from the Lens

The Lens is the flagship project of the social enterprise Cambia, which looks for various sources, and also links and merges diverse open access knowledge collections, encompassing scientific papers and patents, to provide analysis, discovery, and decision making on an open web platform—Lens.org. For over two decades of growth, supported by famous humanitarian organizations, The Lens has collected, processed, normalized, aggregated, and served over 225 million scientific papers, and 127 million worldwide patent records, with large metadata generated from multiple data sources [38].

The Lens is a metadata aggregator that brings together three unique sets of content: Microsoft Academic, Crossref, PubMed, and one basic management tool. This tool supports the main features of The Lens, which are able to detect, analyze, and manage the following objects:

- Scientific papers: exploration and analytics tools supplying access to a universe corpus of metadata of scientific documents with citation indexing,
- Patents: exploration and analytics tools on an extensive collection of patent records with citation indexing,
- PatSeq: a tool for searching and analyzing biological sequences disclosed in patents,
- Collections: a management tool for dynamic or static tracking, monitoring, and analyzing collections of scientific papers or patents.

These basic tools provide global open access, i.e., they are available to any users with or without registration. They are absolutely free, open, and allow for secure and private access to diverse sources of information.

Among them, scientific papers are discovered and analyzed by tools that provide access to a global corpus of scientific literature metadata with citation indexing, mainly from Microsoft Academic. This platform uses machine learning, semantic inference, and knowledge extraction to enable the efficient exploration of scientific information [39]. Specifically, Microsoft Academic used Wikipedia content and image link analysis to broaden the scope of Fields of study (FoS). Starting with a few thousand high-quality FoS seeds, repeating multiple cycles of graph link analysis, and entity, filtering helps reveal more FoS. Hundreds of millions of academic publication metadata such as title, keywords, and abstracts were then scanned to confirm the existence of the new FoS [40]. Based on the subjects and FoS identified by Microsoft Academic in each scientific work, it is possible to summarize the most frequently encountered FoS and Subjects of APM technologies in the Smart energy sector.

3.3.1. Scientific Papers Analysis

Similarly, after applying the search query, which can be found in the Supplementary Spreadsheet S1 Keyword_clusters_with_centraility_values.xlsx (sheet tab Search_queries), at the time of writing this paper, we received 193 results in the period 2010–2021 from the platform The Lens.

By analyzing received data, the heat map of the top 20 most frequent Fields of study is presented in Figure 7.

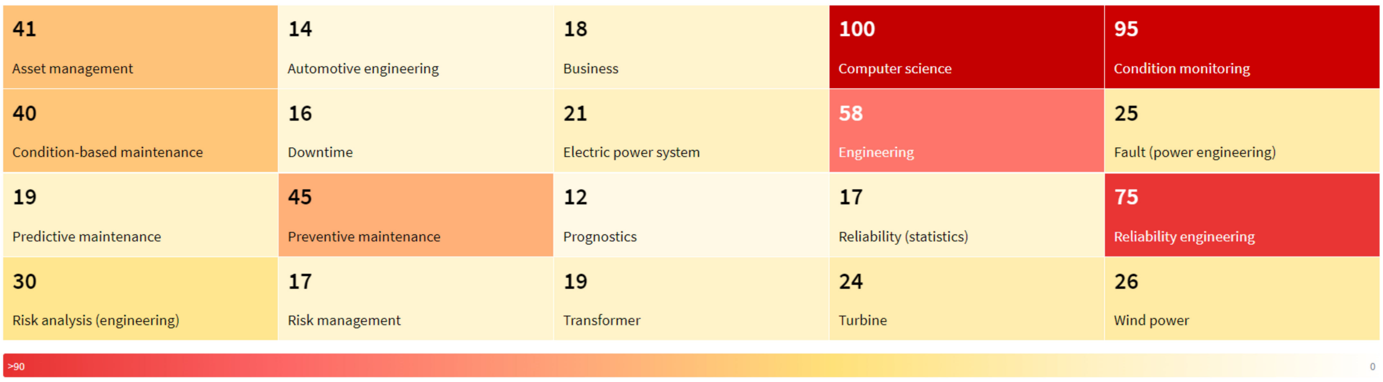

Figure 7. Top 20 most frequent Fields of study in papers of The Lens.

Among them, the most prominent FoS are: "Computer science"; "Wind power"; "Turbine"; "Predictive maintenance"; "Transformer"; "Electric power system"; etc.

Likewise, the heat map of the top 20 most-frequent subjects is presented in Figure 8.

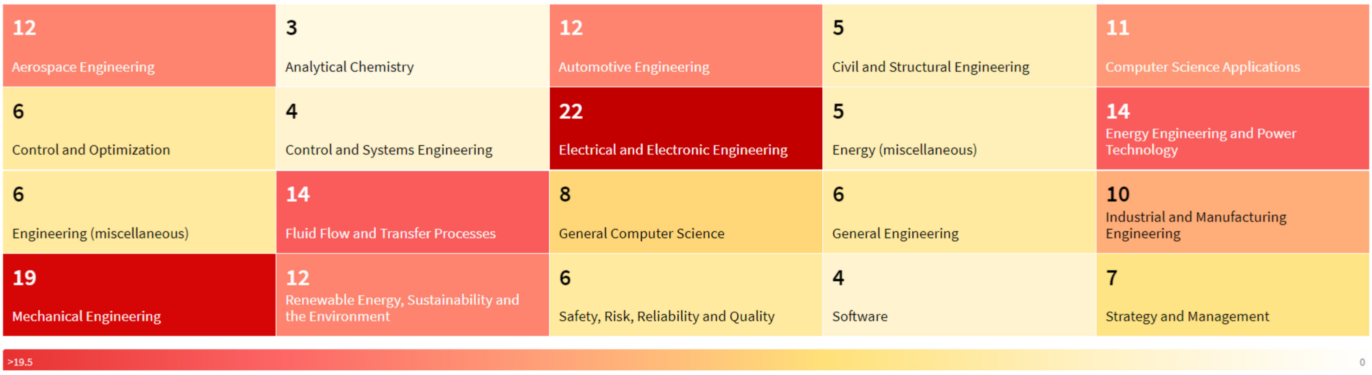

Figure 8. Top 20 most frequent subjects in papers of The Lens.

Among them, the most prominent Subjects are: "Fluid Flow and Transfer Processes"; "Aerospace Engineering"; "Automotive Engineering"; "Renewable Energy, Sustainability and the Environment"; "Computer Science Applications"; "Civil and Structural Engineering"; "Software"; "Analytical Chemistry", etc.

3.3.2. Patent Analysis

After applying the search query, which can be found in the Supplementary Spreadsheet S1 Keyword_clusters_with_centraility_values.xlsx (sheet tab Search_queries), at the time of writing this paper, we received 151 results in the period 2010–2021 from the platform The Lens.

Figure 9 displays the number of patents over time. From the figure, we can see that while the number of Published and Granted patents is increasing, the number of patent applications (Early priority) has had a downward trend over the last few years.

Figure 9. The number of patents over time.

Furthermore, Figure 10 shows the top 20 CPC classification codes extracted from received patents.

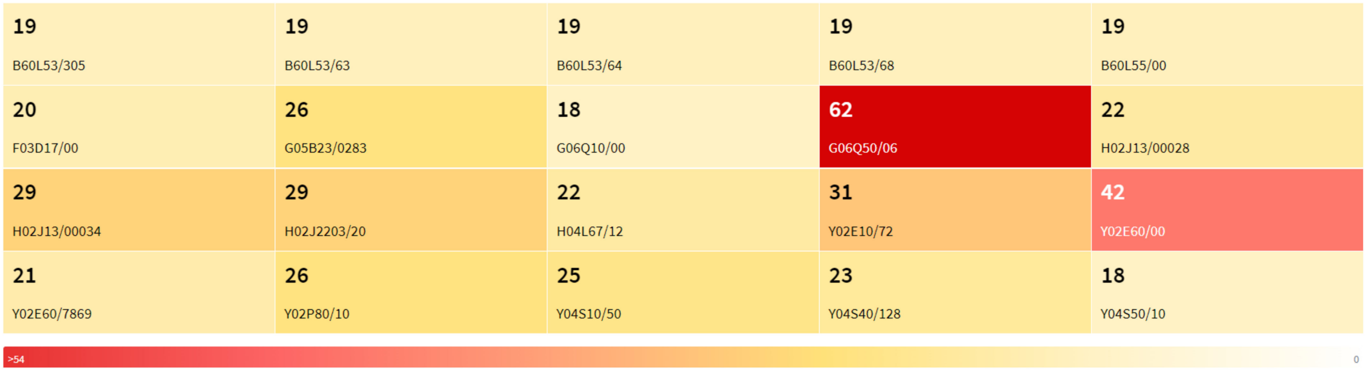

Figure 10. Top 20 CPC classification codes of patents.

Among them, the top 5 CPC classification codes are [41]:

- G06Q50/06: Methods or data processing systems, specially adapted for financial, commercial, administrative, managerial, or predicting intentions, e.g., gas, water, or electricity supply;
- Y02E60/00: Technologies with indirect or potential capability of mitigating greenhouse gas emissions;
- Y02E10/72: Mitigation of greenhouse gas emissions, relating to energy generation, distribution, or transmission due to renewable energy sources;
- H02J13/00034: Circuit arrangements for supplying remote performance of network conditions or providing distant control of switching means in a power distribution network;
- H02J2203/20: Systems or circuit arrangements for distributing or providing electric power; systems for storing electricity.

3.3.3. Clustering Extracted Keywords from Patents in the Period 2017–2021

After filtering patents by the earliest priority date in the last 5 years (2017–2021), we retrieved 61 patents. In order to understand technologies of APM developed in the Smart energy sector, the authors propose a novel patent analysis method with word clustering (Figure 11).

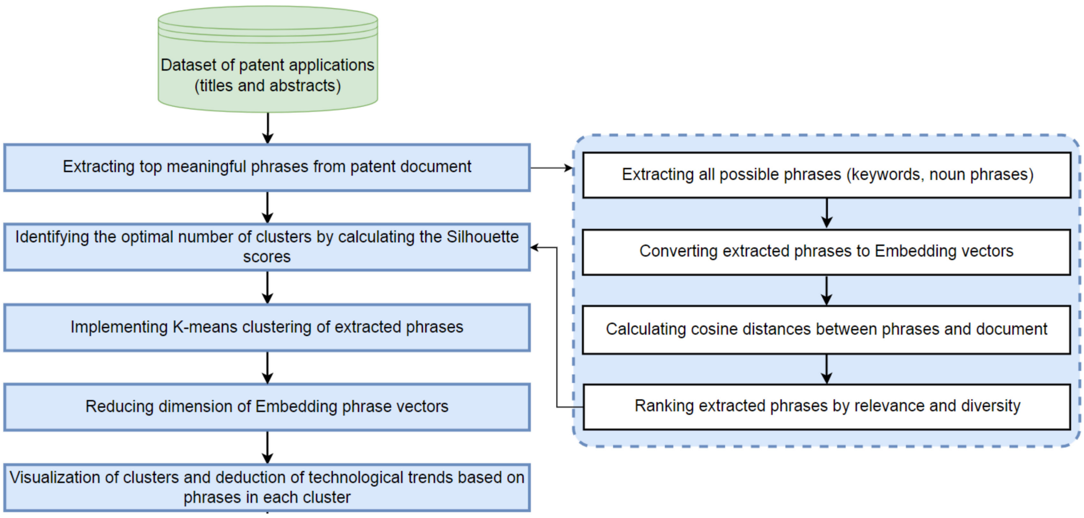

Figure 11. Method for analyzing patents by clustering extracted phrases.

Herein the stage "Extracting all possible phrases" consists of three steps:

- Part-of-speech tagging by utilizing spaCy library [42];

- Noun phrase identification: the longest possible noun phrases sequence of many consecutive words within a sentence such that the last word in the sequence is a noun and each of the other words is either a noun or an adjective;
- Eliminating plural forms by utilizing Natural Language Toolkit (NLTK) platform [43].

The pseudo-code of the noun phrases extraction algorithm is presented below Algorithm 1.

Algorithm 1. Extracting all noun phrases from the text

Input: Text and allowed part-of-speech tags allowed_postags: ADJ (adjective), NOUN (noun), PROPN (pronoun)
Output: List of all noun phrases in singular form

```
masked_words ←[]
for each word in document do
    pos_tag ← get POS tag of word
    if pos_tag is in allowed_postags or word == '-' do
        append word to masked_words
    else if previous word or next word == '-' do
        append word to masked_words
    else do
        append '.' to masked_words
    end
end
extracted_phrases ←[]
while index ≤ length of masked_words do
    get the word at index
    start_index ← index
    if word ≠ '.' do
        start_index ← last index that word not in {'.', 'ADJ'}
        if start_index > index + 1 do
            get phrase between (index and start_index)
            append phrase to extracted_phrases
        end
    end
    index ← start_index + 1
end
lemmatized_phrases ← []
for each phrase in extracted_phrases do
    for each token in phrase do
        get lemmatized word
    end
    join lemmatized words into new phrase
    append new phrase to lemmatized_phrases
end
```

Then, the authors adopt the SentenceTransformers Python framework [44] in the stage “Converting extracted phrases to Embedding vectors” to convert extracted phrases into embedding vectors of 768 dimensions.

In the stage “Ranking extracted phrases by relevance and diversity” we follow the method of maximal marginal relevance, proposed by Maarten [45]. The idea of this method is to minimize redundancy (appearance of similar phrases) and maximize the diversity of results in text summarization tasks (reducing data while increasing knowledge). Concretely, the method starts by selecting the phrase that is the most similar to the document. Then, it iteratively selects new promising phrases that are both similar to the document and different from the previously selected phrases with a predefined diversity parameter 0.3.

Herein, the authors select the most 10 meaningful phrases from each patent document, which consists of Title and Abstract, at the output of stage “Extracting top meaningful keywords/phrases from patent document”. In the final result set, after eliminating duplicates, we obtain 550 unique noun phrases from 61 patent applications in the period 2017–2021.

Then the Silhouette score [46] is used to identify the optimal number of clusters. Figure 12 shows that 60 clusters return the best Silhouette score while using the K-means algorithm.

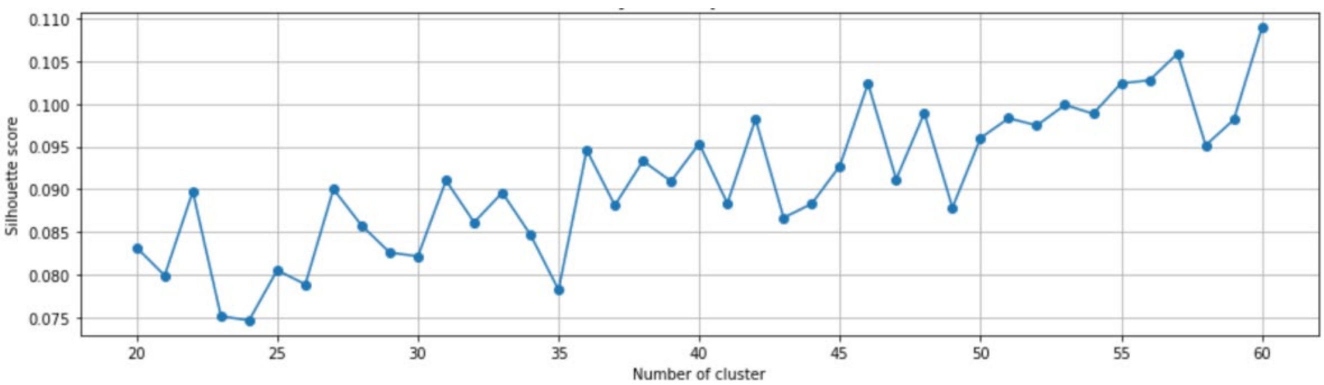

Figure 12. Silhouette analysis for selecting optimal cluster number.

For this reason, by using the K-means algorithm, we divided 550 extracted phrases into 60 clusters, which are available in the Supplementary Spreadsheet S1 Keyword_clusters_with_centraility_values.xlsx (sheet tab Clusters_of_keywords). After scanning these clusters, 13 prominent clusters are listed and demonstrated in Figure 13 below (only some prominent phrases of each cluster are shown). Herein, the principal component analysis (PCA) method is adopted to convert these phrase vectors (768 dimensions) into 2D vectors for visualization:

- Cluster 1: reduction gearbox assembly, hydraulic end assembly, torsional vibration monitoring, self-power-generated bearing module, regenerative burner combustion system, regenerative combustion system, deflection arc, acceleration sensor, rectifier circuit, resistive leakage current, etc.
- Cluster 2: power plant, power transformer, MMC (modular multilevel converter) power device, pumped storage power plant, power capacitor, power grid, power distribution network operation, auxiliary power unit, etc.
- Cluster 6: wireless communication, wireless transmission module, public GPRS (General Packet Radio Services) wireless network, wireless technology, wireless communication interface, wireless gateway device, wireless communication connection, wireless tracking system, RFID-RF (Radio frequency identification-Radio frequency) wireless gateway, RFID tag, RFID-RF tag, RFID circuit, short-range wireless interface, etc.
- Cluster 9: electric signal connection, electric energy quality, electrical machine, electrical energy storage unit, electrical vehicle battery, electrical control system, dc-link capacitor, output voltage reference, dc supply voltage, electric motor, ac main power, electrical transmission system, electrical energy storage, etc.

- Cluster 14: hybrid MMC capacitor reliability evaluation model, power grid risk, electricity theft prevention management, electricity larceny prevention management, power distribution network fault, power supply risk equipment information, power grid operation risk level, EV (electric vehicle) battery monitoring subsystem, valve body leakage detection, local power grid fault detection, power source fault prediction, etc.
- Cluster 16: full cycle data base, big data platform, big data, big data analysis.
- Cluster 18: nuclear power emergency diesel generating set, nuclear power plant, nuclear-grade digital instrument control system, water reactor nuclear power station, reactor protection system, nuclear process control system.
- Cluster 19: machine learning model, machine learning, artificial intelligence circuit, self-learning artificial intelligence, self-learning AI classification, artificial intelligence.
- Cluster 35: entire plunger pump, plunger pump, pump turbines, pump-turbine, valve pipeline system, piston-plunger assembly, quad-quint pump system, pump system, cylinder chamber.
- Cluster 36: geo-location specific solar PV (photovoltaic) power plant, as-operated solar PV power plant, solar photovoltaic power plant, solar PV power plant, photovoltaic plant, PV power generation, PV plant.
- Cluster 37: wind power transmission system, historical wind turbine failure data, wind energy installation, first wind energy installation, wind speed, wind turbine generator.
- Cluster 48: renewable energy transmission grid connection, energy efficiency information gateway, renewable asset management, renewable energy asset, energy utilization rate, energy guide chain, energy chain, energy emission, etc.
- Cluster 57: electrical digital twin-engine, electrical digital twin, AI-driven automatic configuration.

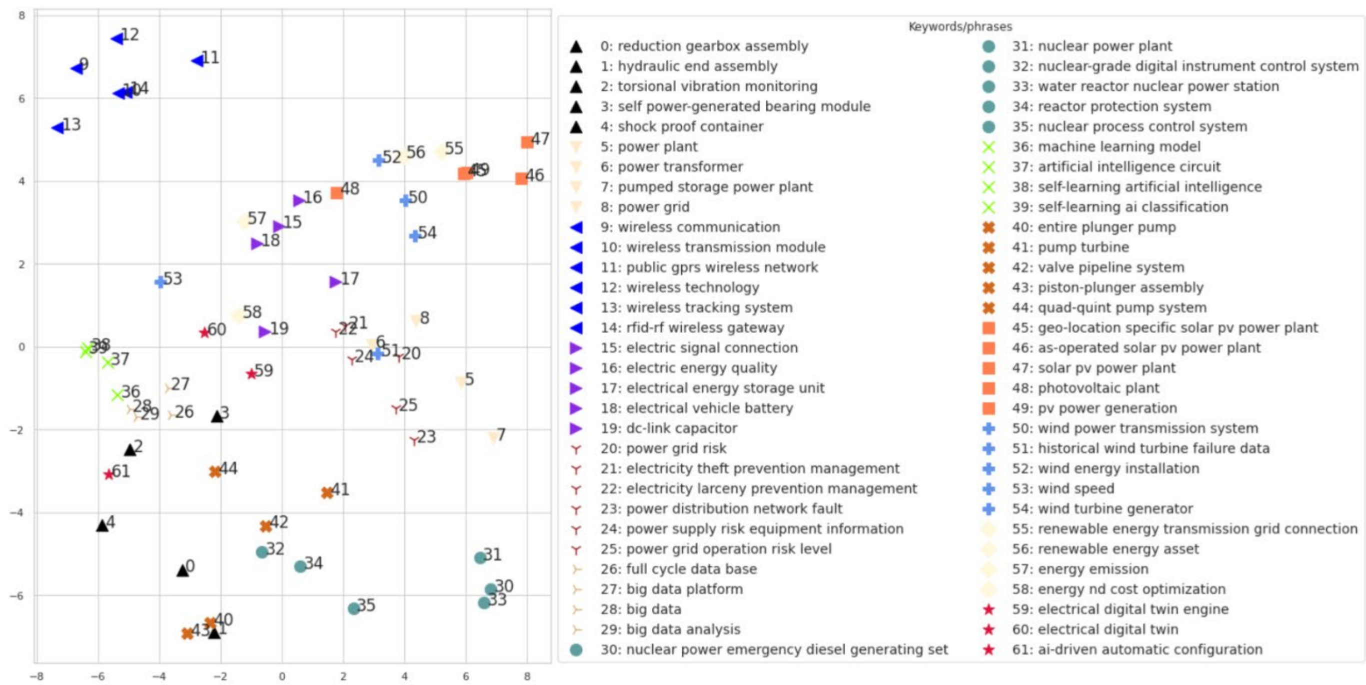

Figure 13. Visualization of 13 phrase clusters extracted from retrieved patents.

3.3.4. Analyzing Scientific Papers Cited in Retrieved Patents

In the past 20 years, the amount of patent citations to scientific papers has increased rapidly. This indicates an ever-stronger flow of knowledge from academic science to industrial innovation. Many studies showed that the propensity of patents to cite academic publications has risen recently, even with changes in the volume and distribution of patents across fields [47].

At the same time, the links between science and technology have been widely studied using references to non-patent literature or comparisons of authors and inventors. The article [48] adopted latent Dirichlet distribution (LDA) to create thematic links between publications and patents based on the semantic content of documents. This approach makes it possible to detect thematic overlaps between patent and scientific publications, highlighting thematic areas used for research and technology.

Additionally, in the study [13], Li et al. put forward a framework that utilized scientific papers and patents for data sources and integrated citation analysis with text mining to analyze the evolutionary way of nanogenerator technology and then predict its trend. The authors started by applying citation analysis to explore technical knowledge covered in scientific papers and to supervise the evolutionary way of nanogenerator technology. Furthermore, the Hierarchical Dirichlet Process (HDP) method of the topic model was applied to detect the technical topics covered in these academic papers. Similarly, the HDP method was also adopted to explore the technical topics in the collected patents. Finally, the authors analyzed the gaps (delay) between science and technology and then combined them with expert knowledge and the evolutionary way of nanogenerator technology in order to predict development tendencies.

Thus, it is very important to analyze the scientific papers referenced by the APM patents. At the time of writing this paper, there are 111 scholarly works cited by the above-mentioned patents. Thus, Figure 14 shows a cloud of the top 50 FoS in cited scientific papers. The most significant ones include: "Computer science", "Risk analysis (engineering)", "Data mining", "Fault (power engineering)", "Artificial intelligence", "Nonlinear system", "Vibration", "Electric power system", "Mathematical optimization", "Support vector machine", "Transformer", "Artificial neural network", "Cluster analysis", etc.

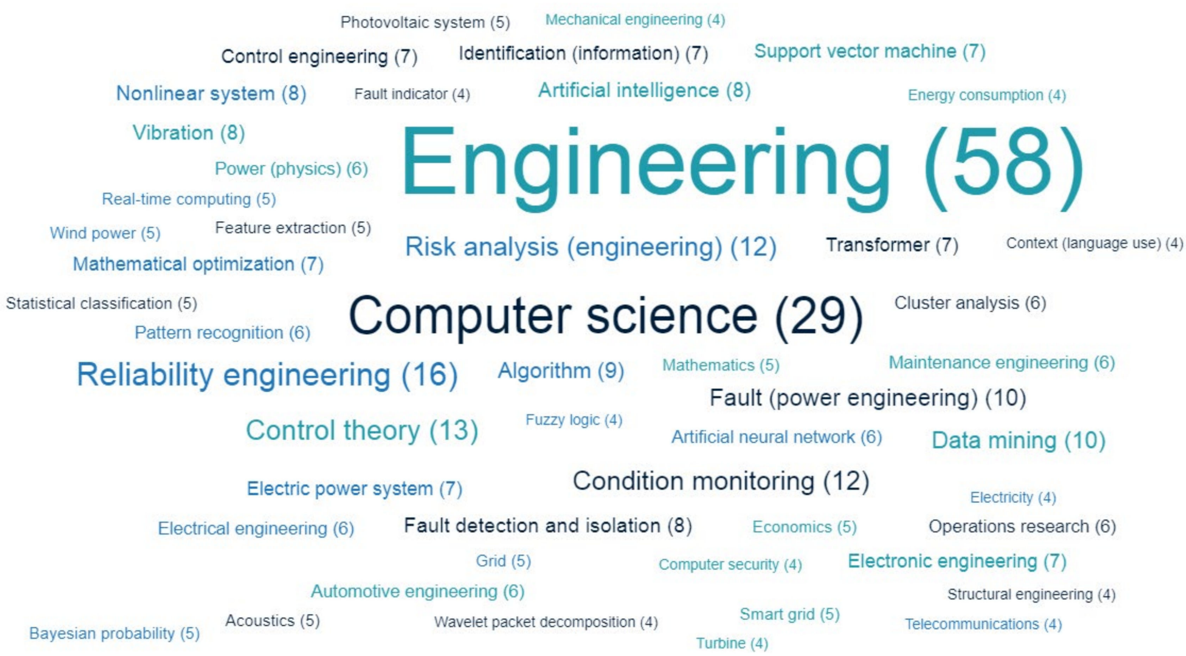

Figure 14. Cloud of top 50 Fields of study in cited scientific papers.

Furthermore, the authors also analyze extracted keywords from cited scientific papers in the period 2010–2020, which include 49 documents. For each paper (Titles and Abstract), all possible keywords were extracted; thus, at the output, 888 unique keywords were retrieved. Then, we make statistics of the keyword frequency. The top 35 most-frequent keywords are displayed in Figure 15 below. These keywords are also considered the most impactful thematic research areas in recent patents, which were filed in the time period 2017–2021.

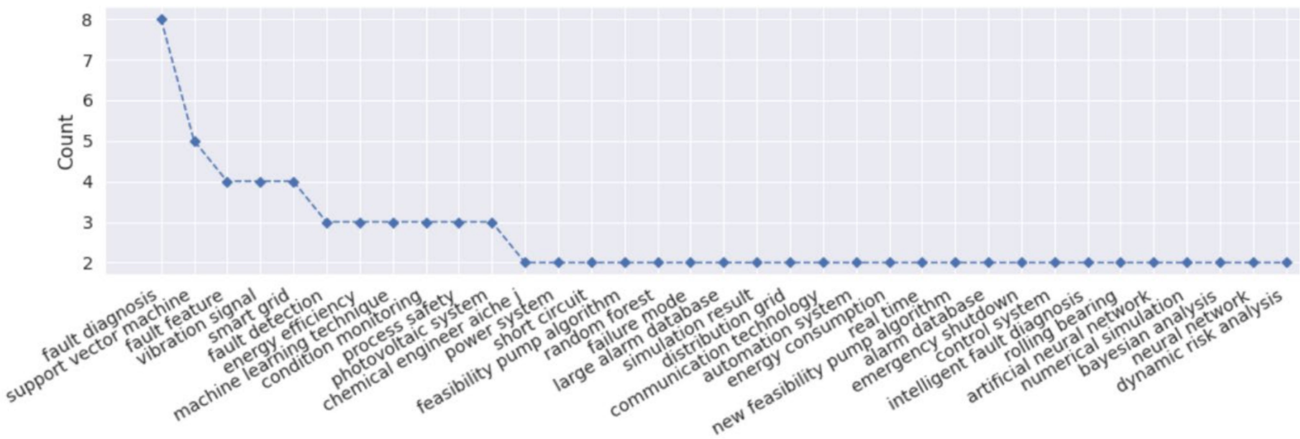

Figure 15. Top 35 most frequent keywords in scientific papers, cited in patents 2017–2021.

Among them, the most significant and impactful keywords include: "support vector machine", "vibration signal", "smart grid", "machine learning technique", "photovoltaic system", etc. These keywords can be divided into some clusters:

- **Problems**: fault diagnosis, fault detection, fault feature, vibration signal, energy efficiency, short circuit, distribution grid, communication technology, automation system, energy consumption, emergency shutdown, control system, rolling bearing, and dynamic risk analysis;
- **Energy sector**: smart grid and photovoltaic system;
- **Algorithm, method**: support vector machine, machine learning technique, feasibility pump algorithm, random forest, large alarm database, real-time, artificial neural network, numerical simulation, and Bayesian analysis.

4. Conclusions

Technological developments have a significant impact on strategic decision-making. Early awareness of possible future or emerging technological trends can lead to increased competitiveness and market positioning of enterprises; however, if innovative companies ignore emerging technological developments, they will not be able to realize the full potential of their own products or technologies. At the same time, the fast growth of emerging technologies has a major impact on people's lives and is considered to transform the current energy industry. Tracking the emergence of impactful technologies and understanding their evolutionary trends at early periods becomes more and more crucial. This information is essential not only for academic research and government investigation but also for strategic planning, firm practices, and social investment.

In particular, Industry 4.0 revolutionizes standards and models by accompanying significant technology developments in numerous sectors, including the sector of energy. Moreover, aging assets and changing requirements always challenge industry energy enterprises. Constraining by limited resources, these organizations have to optimize their current model and method of asset management. Asset performance management is a prominent topic that has attained prevalence in recent years because numerous challenges force all industrial companies to optimize their asset usability. Moreover, climate change issues are still impacting the accelerated aging of active assets and equipment reliability.

Concurrently, the data-driven approach is becoming a positive and feasible tool that enhances asset efficiency, which generates supplementary economic income for enterprises and administrators. At the same time, asset performance management has become the main tool for the digital transformation of industrial companies; therefore, this paper proposes and implements a method for analyzing the trend of technology development in the field of APM in the Smart energy sector based on scientific papers and patents. To achieve this, the most influential factors, significant fields of study, thematic subjects, technologies, and main trends in this subject area are identified. In conclusion, it should be noted that the main trends of the APM technologies in the Smart energy sector over the past 12 years (2010–2021) are as follows:

Industry 4.0, which consists of cyber-physical systems, IoT, cloud, big data, digital twin, machine learning, augmented reality, additive manufacturing, Smart factory, and blockchain, is revolutionizing the traditional industry by creating radical shifts in both business processes and models and also operational activities. Via adopting advanced analytics, it can support APM activities, enable predictive maintenance, improve potential failure detection, and optimize resource assignment.

More specifically, based on the digitalization of Industry 4.0, innovative systems are integrated with visualization and software solutions to facilitate the duty of managers and administrators. A dynamic and simulation representation of an asset management system can be attained due to digital twin technology, while digitalization is an increasingly essential strategy of energy innovation. Various existing results showed that among the promising digital technologies (e.g., big data, artificial intelligence, robotics, blockchain technology, cloud computing, IoT), artificial intelligence is the most extensively utilized in the Smart energy sector. In particular, prominent algorithms of machine learning are becoming completely extended and powerful in controlling renewable energy systems. These sets of facilities are specified by the utilization of a large number of sensors feeding to the SCADA (supervisory control and data acquisition systems). Moreover, numerous authors stated how data mining algorithms, e.g., decision trees (random forest, boosting), support vector machines, and artificial neural networks, are becoming increasingly essential since they represent the majority of the studies.

On the other hand, by analyzing data from scientific papers in the time period 2010–2021, we can conclude that the area of APM in the Smart energy sector is still in the stage of active research development. Concretely, the authors proposed efficient methods and algorithms for data mining, which can help extract rapidly useful knowledge of scientific research and technology trends from a large corpus of publications, hence reducing human efforts in reviewing the existing literature. More specifically, the algorithm for keyword extraction, the method for keyword co-occurrence network analysis (for scientific papers), and keyword clustering (for patents) were devised. Then, we conducted experiments on bibliometric data of famous platforms, i.e., Scopus, Web of Science, and The Lens; the most frequent and important categories, fields of study, and research areas were explored and presented. Moreover, some important and prominent technology trends derived from analyzing scientific papers are the following:

(1) Energy sectors:
- Electricity: electrical network, smart grid, power transformer, electric power transmission networks, electric power distribution, electric network analysis, partial discharges, electric switchgear, electric circuit breakers, etc.;
- Renewable energy: offshore renewable energy, wind power, offshore wind farms, wind turbines, photovoltaic plant, etc.;
- Oil, gas: offshore oil well production, oil-filled transformers, gasoline, gas industry, dissolved gas analysis, etc.;
- Nuclear power: nuclear power plants, nuclear fuels, criticality (nuclear fission).

(2) Algorithm, model, method:
- Digitization, digital twin, digital storage, data-driven, data acquisition, SCADA, FMEA (Failure Modes and Effects Analysis);
- Evolutionary algorithm, genetic algorithm, particle swarm optimization;
- Machine learning, data mining, support vector machine, decision support systems, neural network, Markov processes, fuzzy logic, Monte Carlo methods, regression, decision tree, principal component analysis;
- Industry 4.0: artificial intelligence, big data, cloud, IoT, cyber-physical system, blockchain, augmented reality, additive manufacturing;

(3) Hardware, robot:
    - Robotics, inspect-and-repair robot, crawler robot;
    - Inspection drone, autonomous vessel, sensor, soft sensor.
(4) Others: vibration analysis, signal processing, railroads, railroad transportation, acoustic emission testing, remaining useful lives, gearbox, induction motors, rotating machinery, dielectric response, radiometer, energy storage, infrared thermography, vibration analysis, frequency response analysis, sensitivity analysis, humidity.

At the same time, after clustering keyword co-occurrence network from 691 Scopus papers (titles & abstracts & author keywords & keyword plus), basic description and thematic inference from keywords in each cluster are the following scientific research and technology trends:

- Digital twin and data analytics technologies are mainly adopted for maintenance strategy and predictive maintenance in offshore oil well production;
- Non-destructive testing techniques (Vibration analysis, acoustic emissions testing, temperature measurement and infrared thermography, partial discharge detection) are adopted in real-time condition monitoring to monitor railroad transportation, power converter, induction motor, solar power generation, hydroelectric power plant, and thermal power plant. Moreover, structural health monitoring adopts technologies of wireless smart sensor networks;
- Power transformer asset management is performed by applying various testing methodologies: chemical methodology (gas analysis, oil analysis, Duval triangle, dga), electrical methodology (electric network analysis, frequency response analysis), Markov model, and fuzzy inference-based approach (fuzzy logic);
- Reliability analysis, risk management, and strategic planning in the oil and gas industry and hydroelectric power plant are performed by applying methods of Monte Carlo simulation, Bayesian network, or Weibull distribution. Moreover, Monte Carlo simulation and Bayesian network methods are also employed regularly in environmental risk assessment and CM of nuclear power plants;
- Condition-based maintenance and maintenance planning in offshore wind farms and nuclear power plants are performed by applying technologies of artificial intelligence, machine learning, neural network, Markov process, SCADA system, and data mining;
- Predictive analytics, condition-monitoring data, and electric power system protection in fossil fuel power plants are performed by applying technologies of digital storage, big data, communication, and data visualization;
- Asset management systems and condition assessment in Smart grids and HVDC power transmission networks are performed by utilizing online condition monitoring technologies.

Moreover, the most important keywords among the top-15 by Betweenness centrality, and Eigenvector centrality are: power transformer, wind power, wind turbine, partial discharge, cost–benefit analysis, artificial intelligence, power system, electric power transmission network, smart power grid, offshore oil well production, etc. All of these abovementioned clusters and keywords can also be considered as future trends in the next few years (1–3 years).

Furthermore, by analyzing data of patents in the period 2010–2021, the top 20 CPC classification codes were revealed. Then, by using the K-means algorithm, we divided extracted keywords from patents in the period 2017–2021 into 60 clusters, wherein 13 prominent clusters were presented; however, meaningful clusters could be discovered more with the support of domain experts. These 13 clusters from recent patents are the following:

- Reduction gearbox assembly (hydraulic end assembly, torsional vibration monitoring, regenerative burner combustion system, resistive leakage current, etc.);
- Power plant (power transformer, MMC power device, pumped storage power plant, power capacitor, power grid, auxiliary power unit, etc.);

- Wireless communication (wireless transmission module, public GPRS wireless network, wireless tracking system, RFID-RF wireless gateway, RFID-RF tag, RFID circuit, short-range wireless interface, etc.);
- Electric signal connection (electric energy quality, electrical machine, electrical energy storage unit, electrical vehicle battery, electrical control system, dc-link capacitor, output voltage reference, dc supply voltage, electric motor, etc.);
- Hybrid MMC capacitor reliability evaluation model (power grid risk, electricity theft prevention management, electricity larceny prevention management, power distribution network fault, power supply risk equipment information, etc.);
- Full cycle database (big data platform, big data, big data analysis).
- Nuclear power plant (nuclear power emergency diesel generating set, nuclear-grade digital instrument control system, reactor protection system, etc.);
- Machine learning model (machine learning, artificial intelligence circuit, artificial intelligence, etc.);
- Entire plunger pump (plunger pump, pump turbine, valve pipeline system, piston-plunger assembly, quad-quint pump system, etc.);
- Geo-location specific solar PV power plant (as-operated solar PV power plant, solar photovoltaic power plant, solar PV power plant, etc.);
- Wind power transmission system (historical wind turbine failure data, wind energy installation, wind turbine generator, etc.);
- Renewable energy transmission grid connection (energy efficiency information gateway, renewable asset management, renewable energy asset, etc.);
- Electrical digital twin engine (electrical digital twin, ai-driven automatic configuration).

Moreover, by analyzing scholarly works, which were published in the period 2010–2020 and cited by the abovementioned patents, we discovered the top-50 frequent fields of study; also, by extracting keywords, we detected the most important and impactful problems, methods (algorithms), and energy fields.

Furthermore, we cannot ignore the Industry 5.0 concept, which is the upcoming event horizon of the manufacturing world. The Industry 5.0 concept is mentioned as people working together with smart machines and robots. It is about robots helping humans work faster and better by utilizing modern technologies such as big data and IoT. Because the dissemination of Industry 5.0 is still in its early infancy period, considering numerous challenges and opportunities of Industry 5.0, manufacturers and practitioners should be aware of strategies that integrate human and machine workers in order to maximize the attained benefits. It may be a problem of how we can best leverage the advanced technologies stated in this study to gain optimal outcomes from human and machine interactions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/en15186613/s1, Spreadsheet S1: Keyword_clusters_with_centraility_values.xlsx.

Author Contributions: N.T.V.: methodology, software, validation, formal analysis, investigation, data curation, visualization, writing—original draft preparation; A.G.K.: conceptualization, writing—review and editing, supervision, project administration. All authors have read and agreed to the published version of the manuscript.

Funding: The reported study was funded by RFBR, project No. 20-37-90092.

Data Availability Statement: Data sharing not applicable.

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

References

1. Viet, N.T.; Kravets, A.; Duong, Q.H.T. Data Mining Methods for Analysis and Forecast of an Emerging Technology Trend: A Systematic Mapping Study from SCOPUS Papers. In *Artificial Intelligence*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2021; pp. 81–101. [CrossRef]
2. Viet, N.T.; Kravets, A.G. Analyzing Recent Research Trends of Computer Science from Academic Open-access Digital Library. In Proceedings of the 2019 8th International Conference System Modeling and Advancement in Research Trends (SMART), Moradabad, India, 22–23 November 2019; pp. 31–36. [CrossRef]
3. Sheikh, N.J.; Sheikh, O. Bibliometrics and patents: Case of forecasting biosensor technologies for emerging point-of-care and medical iot applications. In *Innovation Discovery*; World Scientific: Singapore, 2018.
4. Viet, N.T.; Gneushev, V. Analyzing and Forecasting Emerging Technology Trends by Mining Web News. In *Communications in Computer and Information Science*; Springer: Cham, Switzerland, 2021; pp. 55–69. [CrossRef]
5. Röckmann, C.; Lagerveld, S.; Stavenuiter, J. Operation and Maintenance Costs of Offshore Wind Farms and Potential Multi-use Platforms in the Dutch North Sea. In *Aquaculture Perspective of Multi-Use Sites in the Open Ocean*; Springer International Publishing: Cham, Switzerland, 2017; pp. 97–113. [CrossRef]
6. Trzaska, R.; Sulich, A.; Organa, M.; Niemczyk, J.; Jasiński, B. Digitalization Business Strategies in Energy Sector: Solving Problems with Uncertainty under Industry 4.0 Conditions. *Energies* **2021**, *14*, 7997. [CrossRef]
7. Tutak, M.; Brodny, J.; Bindzár, P. Assessing the Level of Energy and Climate Sustainability in the European Union Countries in the Context of the European Green Deal Strategy and Agenda 2030. *Energies* **2021**, *14*, 1767. [CrossRef]
8. Borowski, P. Digitization, Digital Twins, Blockchain, and Industry 4.0 as Elements of Management Process in Enterprises in the Energy Sector. *Energies* **2021**, *14*, 1885. [CrossRef]
9. Sulich, A.; Sołoducho-Pelc, L.; Ferasso, M. Management Styles and Decision-Making: Pro-Ecological Strategy Approach. *Sustainability* **2021**, *13*, 1604. [CrossRef]
10. Ena, O.; Mikova, N.; Saritas, O.; Sokolova, A. A methodology for technology trend monitoring: The case of semantic technologies. *Scientometrics* **2016**, *108*, 1013–1041. [CrossRef]
11. Li, X.; Xie, Q.; Daim, T.; Huang, L. Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology. *Technol. Forecast. Soc. Change* **2019**, *146*, 432–449. [CrossRef]
12. Wang, M.-Y.; Fang, S.-C.; Chang, Y.-H. Exploring technological opportunities by mining the gaps between science and technology: Microalgal biofuels. *Technol. Forecast. Soc. Change* **2015**, *92*, 182–195. [CrossRef]
13. Li, X.; Fan, M.; Zhou, Y.; Fu, J.; Yuan, F.; Huang, L. Monitoring and forecasting the development trends of nanogenerator technology using citation analysis and text mining. *Nano Energy* **2020**, *71*, 104636. [CrossRef]
14. Lee, C.; Kwon, O.; Kim, M.; Kwon, D. Early identification of emerging technologies: A machine learning approach using multiple patent indicators. *Technol. Forecast. Soc. Change* **2018**, *127*, 291–303. [CrossRef]
15. Ferrero Bermejo, J.; Gómez Fernández, J.F.; Pino, R.; Crespo Márquez, A.; Guillén López, A.J. Review and Comparison of Intelligent Optimization Modelling Techniques for Energy Forecasting and Condition-Based Maintenance in PV Plants. *Energies* **2019**, *12*, 4163. [CrossRef]
16. Wan, S. Asset Performance Management for Power Grids. *Energy Procedia* **2017**, *143*, 611–616. [CrossRef]
17. Biard, G.; Nour, G.A. Industry 4.0 Contribution to Asset Management in the Electrical Industry. *Sustainability* **2021**, *13*, 10369. [CrossRef]
18. Rinaldi, G.; Thies, P.R.; Johanning, L. Current Status and Future Trends in the Operation and Maintenance of Offshore Wind Turbines: A Review. *Energies* **2021**, *14*, 2484. [CrossRef]
19. Gao, Z.; Liu, X. An Overview on Fault Diagnosis, Prognosis and Resilient Control for Wind Turbine Systems. *Processes* **2021**, *9*, 300. [CrossRef]
20. Ren, Z.; Verma, A.S.; Li, Y.; Teuwen, J.J.E.; Jiang, Z. Offshore wind turbine operations and maintenance: A state-of-the-art review. *Renew. Sustain. Energy Rev.* **2021**, *144*, 110886. [CrossRef]
21. Civera, M.; Surace, C. Non-Destructive Techniques for the Condition and Structural Health Monitoring of Wind Turbines: A Literature Review of the Last 20 Years. *Sensors* **2022**, *22*, 1627. [CrossRef]
22. Malekimoghadam, R.; Krause, S.; Czichon, S. A Critical Review on the Structural Health Monitoring Methods of the Composite Wind Turbine Blades. In Proceedings of the 1st International Conference on Structural Damage Modelling and Assessment, Ghent, Belgium, 4–5 August 2020; Lecture Notes in Civil Engineering. Springer: Cham, Switzerland, 2021; pp. 409–438. [CrossRef]
23. Lyu, W.; Liu, J. Artificial Intelligence and emerging digital technologies in the energy sector. *Appl. Energy* **2021**, *303*, 117615. [CrossRef]
24. Weigel, P.; Fischedick, M. Review and Categorization of Digital Applications in the Energy Sector. *Appl. Sci.* **2019**, *9*, 5350. [CrossRef]
25. Van Eck, N.J.; Waltman, L. Visualizing Bibliometric Networks. In *Measuring Scholarly Impact*; Springer International Publishing: Cham, Switzerland, 2014; pp. 285–320. [CrossRef]
26. Chang, S.-H. Revealing Development Trends and Key 5G Photonic Technologies Using Patent Analysis. *Appl. Sci.* **2019**, *9*, 2525. [CrossRef]

27. Lee, Y.; Lee, Y.; Seong, J.; Stanescu, A.; Hwang, C.S. A comparison of network clustering algorithms in keyword network analysis: A case study with geography conference presentations. *Int. J. Geospat. Environ. Res.* **2020**, *7*, 1–16.
28. Csárdi, G.; Nepusz, T.; Airoldi, E.M. *Statistical Network Analysis with Igraph*; Springer: New York, NY, USA, 2016.
29. Wanasinghe, T.R.; Wroblewski, L.; Petersen, B.K.; Gosine, R.G.; James, L.A.; De Silva, O.; Mann, G.K.I.; Warrian, P.J. Digital Twin for the Oil and Gas Industry: Overview, Research Trends, Opportunities, and Challenges. *IEEE Access* **2020**, *8*, 104175–104197. [CrossRef]
30. Yang, W. Condition monitoring of offshore wind turbines. In *Offshore Wind Farms*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 543–572. [CrossRef]
31. Martinez-Luengo, M.; Kolios, A.; Wang, L. Structural health monitoring of offshore wind turbines: A review through the Statistical Pattern Recognition Paradigm. *Renew. Sustain. Energy Rev.* **2016**, *64*, 91–105. [CrossRef]
32. Montanari, G.C.; Hebner, R.; Seri, P.; Ghosh, R. Self-Assessment of Health Conditions of Electrical Assets and Grid Components: A Contribution to Smart Grids. *IEEE Trans. Smart Grid* **2021**, *12*, 1206–1214. [CrossRef]
33. Sofi, A.; Jane Regita, J.; Rane, B.; Lau, H.H. Structural health monitoring using wireless smart sensor network—An overview. *Mech. Syst. Signal Process.* **2022**, *163*, 108113. [CrossRef]
34. Liang, Z.; Parlikad, A. A Markovian model for power transformer maintenance. *Int. J. Electr. Power Energy Syst.* **2018**, *99*, 175–182. [CrossRef]
35. Yahaya, M.S.; Azis, N.; Abidin Ab Kadir, M.Z.; Jasni, J.; Hairi, M.H.; Talib, M.A. Estimation of transformers health index based on the markov chain. *Energies* **2017**, *10*, 1824. [CrossRef]
36. Soni, R.; Mehta, B. Review on asset management of power transformer by diagnosing incipient faults and faults identification using various testing methodologies. *Eng. Fail. Anal.* **2021**, *128*, 105634. [CrossRef]
37. Ozonuwe, S.N.; Onyekachi, D.; Oside, C.U. Application of the Two-parameter Weibull Distribution Method to Assess the Reliability of Gas Turbine Compressors. *J. Eng. Res. Rep.* **2020**, *18*, 12–20. [CrossRef]
38. About The Lens. Available online: https://about.lens.org/ (accessed on 19 June 2022).
39. Sinha, A.; Shen, Z.; Song, Y.; Ma, H.; Eide, D.; Hsu, B.-J.; Wang, K. An Overview of Microsoft Academic Service (MAS) and Applications. In Proceedings of the 24th International Conference on World Wide Web, Florence, Italy, 18–22 May 2015; ACM: New York, NY, USA, 2015; pp. 243–246. [CrossRef]
40. Wang, K.; Shen, Z.; Huang, C.; Wu, C.H.; Eide, D.; Dong, Y.; Qian, J.; Kanakia, A.; Chen, A.; Rogahn, R. A Review of Microsoft Academic Services for Science of Science Studies. *Front. Big Data* **2019**, *2*, 45. [CrossRef]
41. Cooperative Patent Classification. Available online: https://worldwide.espacenet.com/classification?locale=en_EP (accessed on 2 July 2022).
42. Industrial-Strength Natural Language Processing. Available online: https://spacy.io/ (accessed on 26 February 2022).
43. Natural Language Toolkit. Available online: https://www.nltk.org/index.html (accessed on 27 February 2022).
44. SentenceTransformers Documentation. Available online: https://www.sbert.net/ (accessed on 26 February 2022).
45. KeyBERT—Quickstart. Available online: https://maartengr.github.io/KeyBERT/guides/quickstart.html (accessed on 26 February 2022).
46. Sklearn.Metrics.Silhouette_Score. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html (accessed on 26 February 2022).
47. Wei, F.; Zhang, G.; Zhang, L.; Liang, Y.; Wu, J. Decreasing the noise of scientific citations in patents to measure knowledge flow. In *17th International Conference on Scientometrics and Informetrics, ISSI 2019–Proceedings*; International Society for Scientometrics and Informetrics: Leuven, Belgium, 2019; pp. 1662–1669.
48. Suominen, A.; Ranaei, S.; Dedehayir, O. Exploration of Science and Technology Interaction: A Case Study on Taxol. *IEEE Trans. Eng. Manag.* **2021**, *68*, 1786–1801. [CrossRef]