Review

Healthcare Trust Evolution with Explainable Artificial Intelligence: Bibliometric Analysis

Pummy Dhiman 1, Anupam Bonkra 2,*, Amandeep Kaur 1,*, Yonis Gulzar 3,*, Yasir Hamid 4,*, Mohammad Shuaib Mir 3,*, Arjumand Bano Soomro 3,5 and Osman Elwasila 3

1 Institute of Engineering and Technology, Chitkara University, Rajpura 140417, Punjab, India
2 Information Technology Department, Chandigarh Group of Colleges, Landran 140307, Punjab, India
3 Department of Management Information Systems, College of Business Administration, King Faisal University, Al-Ahsa 31982, Saudi Arabia
4 Information Security and Engineering Technology, Abu Dhabi Polytechnic College, Abu Dhabi 111499, United Arab Emirates
5 Department of Software Engineering, Faculty of Engineering and Technology, University of Sindh, Jamshoro 76080, Pakistan

* Correspondence: anupam.4829@cgc.edu.in (A.B.); ygulzar@kfup.edu.sa (Y.G.); Tel.: +966-545719118 (Y.G.)

Abstract: Recent developments in IoT, big data, fog and edge networks, and AI technologies have had a profound impact on a number of industries, including medical. The use of AI for therapeutic purposes has been hampered by its inexplicability. Explainable Artificial Intelligence (XAI), a revolutionary movement, has arisen to solve this constraint. By using decision-making and prediction outputs, XAI seeks to improve the explicability of standard AI models. In this study, we examined global developments in empirical XAI research in the medical field. The bibliometric analysis tools VOSviewer and Biblioshiny were used to examine 171 open access publications from the Scopus database (2019–2022). Our findings point to several prospects for growth in this area, notably in areas of medicine like diagnostic imaging. With 109 research articles using XAI for healthcare classification, prediction, and diagnosis, the USA leads the world in research output. With 88 citations, IEEE Access has the greatest number of publications of all the journals. Our extensive survey covers a range of XAI applications in healthcare, such as diagnosis, therapy, prevention, and palliation, and offers helpful insights for researchers who are interested in this field. This report provides a direction for future healthcare industry research endeavors.

Keywords: Explainable Artificial Intelligence; XAI; healthcare; neural network; bibliometric analysis

1. Introduction

New applications for artificial intelligence (AI) have been generated by recent developments in machine learning (ML), the Internet of Things (IoT) [1], big data, and assisted fog and edge networks, which offer several benefits to many different sectors. However, many of these systems struggle to justify their own decisions and actions to those who are not computers. The emphasis on explanation, according to some AI researchers, is incorrect, unrealistic, and perhaps unnecessary for all applications of AI [2]. The authors of [3] proposed the phrase “explainable AI” to highlight a training system developed for the US Army’s capacity to justify its automation choices. The Explainable Artificial Intelligence (XAI) program was started in 2017 via the Defense Advanced Research Projects Agency (DARPA) [3] to construct methods for comprehending intelligent systems. DARPA refers to a collection of methods as XAI to describe how they develop explainable models that, when combined with successful explanation procedures, allow end-users to grasp, correctly trust, and efficiently manage the next generation of AI systems.

In keeping with the perception of keeping humans in the loop, XAI aims to make it simpler for people to comprehend opaque AI systems so they may use these tools to...
help with their work more successfully. Recent applications of XAI include those in the military, healthcare, law, and transportation. In addition to software engineering, socially sensitive industries, including edification, law enforcement and forensics, healthcare, and agriculture, are also seeing an increase in the usage of ML and deep learning feature extraction and segmentation techniques [4,5]. This makes using them considerably more difficult, especially given that many people who are dubious about the future of these technologies just do not know how they operate.

AI has the potential to help with a number of critical issues in the medical industry. The fields of computerized diagnosis, prospects, drug development, and testing have made significant strides in recent years.

Within this particular framework, the importance of medical intervention and the extensive pool of information obtained from diverse origins, including electronic health records, biosensors, molecular data, and medical imaging, assume crucial functions in propelling healthcare forward and tackling pressing concerns within the medical sector. Establishing treatments, decisions, and medical procedures specifically for individual patients is one of the objectives of AI in medicine. The current status of artificial intelligence in medicine, however, has been described as heavy on promise and fairly light on evidence and proof. Multiple AI-based methods have succeeded in real-world contexts for the diagnosis of forearm sprains, histopathological prostate cancer lesions [4], very small gastrointestinal abnormalities, and neonatal cataracts. But in actual clinical situations, a variety of the systems that encompass them have been demonstrated to be on par with or even better than those used by specialists in experimental studies and have large false-positive rates. By improving the transparency and interpretability of AI-driven medical applications, Explainable Artificial Intelligence has the potential to completely transform the healthcare system. Healthcare practitioners must comprehend how AI models make judgments in key areas, including diagnosis, therapy suggestions, and patient care.

Clinical decision making is more informed and confident thanks to XAI, which gives physicians insights into the thinking underlying AI forecasts. Doctors may ensure patient safety by identifying potential biases, confirming the model’s correctness, and offering interpretable explanations. Additionally, XAI promotes the acceptance of AI technology in the healthcare industry, allaying worries about the “black box” nature of AI models. By clearly communicating diagnoses and treatment plans, transparent AI systems can improve regulatory compliance, resolve ethical concerns, and increase patient participation.

Healthcare professionals may fully utilize AI with XAI while still maintaining human supervision and responsibility. In the end, this collaboration between AI and human knowledge promises to provide more individualized and accurate healthcare services, enhance patient outcomes, and influence the course of medical research.

There are some important taxonomies of XAI that exist to show the antithesis of some AI, ML, and particularly DL models’ black-box characteristics. The following terms are distinguished in Figure 1.

![Figure 1. Terms of XAI.](image-url)
Transparency: A sculpture is said to be translucent if it has the capacity to make sense on its own. Thus, lucidity is the contradiction of a black box [5].

Interpretability: The term “interpretability” describes the capacity to comprehend and articulate how a complicated system, such as a machine learning model or an algorithm, makes decisions. It entails obtaining an understanding of the variables that affect the system’s outputs and how it generates its conclusions [6]. Explainability is an area within the realm of interpretability, and it is closely linked to the notion that explanations serve as a means of connecting human users with artificial intelligence systems. The process encompasses the categorization of artificial intelligence that is both accurate and comprehensible to human beings [6].

According to the authors of [7], XAI is required within any of the following scenarios:

- Where in the interest of fairness and to help customers make an informed decision, an explanation is necessary.
- Where the consequences of a wrong AI decision can be very far-reaching (such as recommending surgery that is unnecessary).
- In cases where a mistake results in unnecessary financial costs, health risks, and trauma, such as malignant tumor misclassification.
- Where domain experts or subject matter experts must validate a novel hypothesis generated by the AI.
- The EU’s General Data Protection Regulation (GDPR) [8] gives consumers the right to explanations when data are accessed through an automated mechanism.

1.1. Taxonomy of XAI

1.1.1. Translucent Model

The authors of [5] provide a list of a few well-known transparent models, including Fuzzy systems, decision trees, principal learning, and K-nearest neighbors (KNN). Typically, these models yield decisions that are unambiguous; however, it should be noted that mere transparency does not guarantee that a given concept will be easily comprehensible, as illustrated in Figure 2.

**Figure 2.** Taxonomy of XAI.
1.1.2. Opaque Models

Black-box or opaque models are those used in machine learning or artificial intelligence that lack transparency and are challenging for humans to understand. These models are difficult to comprehend or describe in a way that is intelligible to humans since they base their choices on intricate relationships between the input characteristics. Transparent models encompass various types, such as rule-based models, decision trees, and linear regression. The comprehension of predictions made by opaque models can be achieved through the utilization of transparent models or procedures, such as post hoc explanations like LIME (Local Interpretable Model-Agnostic Explanations) or SHAP (Shapley Additive Explanations). These approaches provide a balance between accuracy and interpretability.

1.1.3. Model-Agnostic Techniques

Explainable Artificial Intelligence techniques that are model-agnostic are meant to make machine learning model decisions more understandable and interpretable without relying on the specifics of the model’s internal architecture. These methods attempt to be applicable to many models and may be used in a variety of contexts, making them extremely adaptive and versatile. Establishing a clear and understandable link between the input characteristics (data) and the model’s output (predictions) is the main goal of model-agnostic XAI. These methods do not need access to the model’s internal parameters, intermediate representations, or procedures. They are “agnostic” to the model’s underlying complexity since they only focus on the input–output connection [9].

1.1.4. Model-Specific Techniques

The purpose of model-specific XAI techniques is to enable interpretability and transparency for a particular model or a subset of models by leveraging the internal architecture, knowledge, and features of a given model. Model-specific XAI focuses on maximizing interpretability for a specified model, as opposed to model-agnostic techniques, which are flexible across multiple models. These methods make use of the model’s built-in structure, such as neural network attention processes or decision tree decision rules, to produce explanations that are consistent with the model’s knowledge. Model-specific XAI strives to give more precise and educational insights for that particular model or a specific group of related models by customizing explanations to the model’s complexities [10].

1.1.5. Simplification of Enlightenment

By using a rough framework, it is possible to find different ways in which computer versions could be made that help explain the prophecy being looked into. To show a more complicated structure, a regression analysis or decision tree can be used as it is based on the model’s predictions [11].

1.1.6. Relevance of Explanation by Feature

This concept is comparable to generalization. After considering all potential combinations, this kind of XAI approach evaluates an attribute according to its predicted marginal contribution to the model’s decision [10,11].

1.1.7. Graphic Explanation

This particular XAI strategy is built around visualization. In light of this, it is possible to use the family of data visualization techniques to construe the prophecy or decision taken in light of the input data [12].

1.1.8. Narrow Explanation

By focusing on inputs similar to the one we are seeking to explain, narrow explanations shed light on the behavior of the model. They carry out the model’s decision-making process in a constrained space centered on an interesting case [13]
Healthcare has advanced significantly, with more innovative research and a shift towards healthcare 5.0. In order to transition into the new era of smart disease control and detection, virtual care, smart health management, smart monitoring, and decision making and decision explanation, the healthcare industry is in the midst of a paradigm shift. The purpose of XAI is to offer machine learning and deep learning algorithms that perform better while being understandable, making it easier for users to trust, comprehend, accept, and use the system [8]. Several studies give insight into how XAI is utilized in healthcare [14].

Exploring XAI in the healthcare industry is essential for ensuring ethical, transparent, and responsible AI use, which will improve patient care and boost confidence in AI-driven medical decisions. Furthermore, it aids in reducing the risk of diagnostic errors and ensures compliance with healthcare standards. With this in mind, the objective of this study is to evaluate a prior research-based analysis in order to obtain insight into the work performed and opportunities given by AI advancement, and the explainability feature of AI in the healthcare sector. To accomplish this task, a comprehensive analysis of the existing literature was conducted in order to address the following research inquiries.

RQ1: What are the local and most pertinent sources in the healthcare industry that worked with XAI between 2019 and 2022?
RQ2: Who is the most relevant author with affiliations to the healthcare domain who worked with XAI between 2019–2022?
RQ3: What are the most productive geographical regions in terms of social research collaboration in the field of healthcare that have worked with XAI?

Section 2.1 outlines the methods employed to address the aforementioned questions. Answering these questions will give us comprehensive knowledge of the current studies employed in this domain. The following describes the foremost contributions of this work:

- It discusses the latest papers investigating the intermingling of XAI with the healthcare domain.
- Based on the various research published in the past year, it elaborates on various publishing patterns.
- It shows how much different nations or areas have contributed to this area of study.
- It talks about the importance of academic writers who have contributed considerably to the integration of XAI in the healthcare industry.
- It talks about a lot of places where publishing patterns are dependent on relationships (colleges/organizations).
- It displays the number of citations a publication received for each contribution connected to the impact of XAI on clinical health practices and increases transparency for predictive analysis, which is essential in the healthcare industry.

2. Literature Review

This section discusses the critical role of XAI in increasing the trust in and overall acceptance of AI systems in healthcare. Explainable AI (XAI) provides users with an explanation of why a method produces a particular result. The outcome can then be understood in a particular context. A crucial application of XAI is in clinical decision support systems (CDSSs) [15]. These methods aid doctors in making judgments in the clinic but, because of their complexity, may lead to issues with under- or overreliance. Practitioners will be better able to make decisions that, in some circumstances, could save lives as the practitioners are given explanations for the processes used to arrive at recommendations. The demand for XAI in CDSS and therapeutic industries in general has arisen as a result of the necessity for principled and equitable decision making as well as the actuality that AI skilled in chronological data might perpetuate pre-existing behaviors and prejudices that should be exposed [15].
Medical Imaging and Diagnosis

Medical imaging and diagnosis often benefit from the use of these techniques. XAI can provide valuable insights into the decision-making process and model behavior, setting it apart from other artificial intelligence methods, such as deep learning. Current advancements have placed significant emphasis on the utilization of Explainable Artificial Intelligence (XAI) in the domains of surgical procedures and medical diagnoses. For instance, Explainable Artificial Intelligence (XAI) has the potential to improve the comprehensibility and transparency of medical image analysis [16], specifically in the context of breast cancer screening. This statement pertains to the issue presented by the lack of transparency in AI systems [17,18].

Chronic Disease Detection

Chronic disease management poses a continuous healthcare burden, particularly in places such as India, where diseases like diabetes and asthma prevail. Artificial intelligence (AI), including explainable AI (XAI), assumes a crucial role in facilitating the coordination of therapies for chronic illnesses. It offers valuable insights into the physical and mental well-being of individuals, thus assisting patients in efficiently managing their health [19,20].

COVID-19 Diagnosis

During the COVID-19 pandemic, AI, including XAI, has significantly improved diagnostic accuracy. For instance, chest radiography, a critical screening tool, is employed to identify COVID-19 cases, particularly when traditional methods like polymerase chain reaction fall short. XAI contributes by elucidating the factors influencing COVID-19 detection, thereby enhancing the screening process [21].

Global Health Goals

Global health objectives can be effectively pursued through the utilization of digital health technologies, which encompass artificial intelligence (AI). These technologies are in line with several initiatives proposed by the United Nations and play a significant role in advancing global health goals. These technologies utilize patient data, environmental information, and connectivity to enhance healthcare delivery, demonstrating significant value during times of emergencies and disease epidemics [22].

Pain Assessment

The assessment of pain in patients has been significantly enhanced by advancements in Artificial Intelligence (AI), particularly with a focus on Explainable AI (XAI). Artificial intelligence (AI) and machine learning (ML) models are capable of analyzing facial expressions, which can serve as reliable indicators of pain and suffering. This technical application exhibits potential in the field of healthcare, specifically in the assessment of pain levels among patients [23].

Biometric Signal Analysis

The study presented in [24] employed a modified bidirectional LSTM network with Bayesian optimization for the automated detection and classification of ECG signals. This system has the capability to improve accuracy and has practical implications for effectively addressing issues associated with categorizing biometric data. A Bayesian optimization-modified bidirectional LSTM (BiLSTM) network is employed in this inquiry to provide an automated ECG detection and classification approach. The two hyperparameters of the BiLSTM network that are optimized using Bayesian techniques are the initial learning rate and the total number of hidden layers. When categorizing five ECG signals in the MIT-BIH arrhythmia database, the improved network’s accuracy rises to 99.00%, an increase of 0.86% from its pre-optimization level. The potential practical relevance of the presented
approach to further quasi-periodic biometric signal-based categorization problems might be considered in future research [24].

Stroke Recognition

Stroke, a deadly medical ailment that occurs when the brain’s blood supply is cut off, is the subject of much investigation. Brain cells die if blood flow is abruptly interrupted. Kaggle provided a dataset for our research experiment comparing stroke recognition machine learning models. The data were pre-processed cheaply because XAI contributes little overhead time during training. Explainable artificial intelligence, or “XAI”, is a cutting-edge machine learning technology that adds interpretation [25]. In [25], the author’s study of numerous research methods begins with artificial intelligence’s interpretability and explainability. Two explainable AI-based cardiac disease prediction experiments are compared. This comparison can help AI beginners choose the best techniques [26]. In another study, deep learning models in electronic health records (EHRs) are examined, along with interpretability in medical AI systems [27].

Overall, medical professionals employ AI to speed up and improve several healthcare processes. These include forecasting, risk assessment, diagnosis, and decision making. They finish by carefully studying medical images to find hidden anomalies and patterns that humans cannot see. Many healthcare professionals have already integrated AI into their workflows, but doctors and patients often become frustrated with its operations, especially when making important judgments. This industry’s demand for explainable AI (XAI) drives its adoption. Complex AI suggestions like surgical procedures or hospital hospitalizations need explanations for patients and physicians [28]. Examining the revised contributions revealed substantial implications for academics and professionals, prompting an exploration of methodological aspects to promote medical AI applications. The philosophical foundations and contemporary uses of 17 Explainable Artificial Intelligence (XAI) techniques in healthcare were examined. Finally, legislators were given goals and directions for building electronic healthcare systems that emphasize authenticity, ethics, and resilience. The study examined healthcare information fusion methodologies, including data fusion, feature aggregation, image analysis, decision coordination, multimodal synthesis, hybrid methods, and temporal considerations [29]. AI has been used to manage healthcare services, forecast medical outcomes, enhance professional judgment, and analyze patient data and diseases. Despite their success, AI models are still ignored since they are seen as “black boxes”. Lack of trust is the largest barrier to their widespread usage, especially in healthcare. To ease this concern, Explainable Artificial Intelligence (XAI) has evolved. XAI improves AI model predictions by revealing their logic. XAI helps healthcare providers embrace and integrate AI systems by explaining the model’s inner workings and prediction approach [30]. They thoroughly analyze healthcare scenarios involving explanation interfaces. Our systematic search of leading research databases for healthcare studies shows the adaptability of intelligent systems and the variety of academic approaches to answering questions. Saliency maps, natural language writing, parameter effect evaluations, and data pattern graphs are examples of these explanations [31].

As is generally known, artificial intelligence is crucial for the diagnosis, detection, and prevention of diseases. The black-box characteristic of ML and DL is overcome by XAI. In light of this, the goal of this research is to perform a bibliometric study on the impact of XAI on the growth of trust in the outputs of AI black boxes [5].

2.1. Methodology

The following research methodology (Figure 3) was utilized in order to accomplish the above-mentioned goal. To conduct this bibliometric study, some guidelines should be followed; with this in mind, this study followed guidelines given in [32].
Figure 3. Five-step methodology adopted for bibliometric study.

2.1.1. Planning

This section outlines the design, methodology, and execution of the investigation. The plan for carrying out the entire research process is described here, and it is elaborated upon in the following subsections. The goal of this study is to examine trends or patterns in the use of Explainable Artificial Intelligence in the field of healthcare, as well as knowledge organization and knowledge synthesis.

2.1.2. Data Collection

Reviewing existing literature and prior research findings is an essential step in the process of conducting research and drafting a research paper. We decided to perform a bibliometric analysis as part of this study [32]. We asked ourselves which method of analyzing academic literature is quantitative? So, the focus was placed on the quantity of scientific and academic literature to analyze its impact. Therefore, academic research databases came into play because only trusted resources would be used. There are various academic databases available, including Scopus [33], Web of Science [34], and PubMed [35], to name a few. Scopus, which was launched in 2004 and is owned by Elsevier, was used for this study. Scopus is a multidisciplinary database that is not only the largest abstract and citation database, but also ranks journals and authors. It offers some services for free, but full access to the database requires a subscription. The statistics intended for this work were taken from the Scopus repository for the years 2019 through 2022. Here, articles in the press were also considered to show that XAI is an emerging technique.

2.1.3. Search Strategy

Here, a search string using keywords was formulated, which was then searched throughout in the Scopus database. By utilizing the Boolean OR, AND operator, as shown below, the search was conducted in the article title, abstract, and keywords. Our initial string was as follows:

(“XAI” OR “Explainable Artificial Intelligence”) AND (“Health Care” OR “Diagnosis” OR “Classification”).

Initially, there were 1058 results after processing this string.
2.1.4. Screening

Initially, 1058 search results were obtained using the initial search string, but we refined them based on the following search criteria: Although there are various types of documents available in the Scopus database, for this study, only peer-reviewed articles with open access were considered, which limits this study to articles and conference papers. There were five types of sources available after executing this search string, but as mentioned earlier, we only opted for peer-reviewed articles; therefore, only journal articles and conference proceedings were selected. There were 28 papers with undefined authors, which we omitted. Apart from English, there were articles in other languages including Korean (3), Chinese (2), Russian (1), and Turkish (1). For this study, we limited our documents to those in the English language. Following refinement, the following search string was created, yielding 190 results.

```
```

After further refinement, such as after reading the content, we found 171 papers out of the above extracted 190 to be relevant to our study, which were exported in CSV format [36]. Every publication contained details like the author, the country, citations, papers, references, etc.

2.1.5. Performance Scrutiny

The impact of publications on the scientific community may be ascertained via bibliometric analysis [37,38]. This involves the statistical analysis of books, papers, or book chapters that have been published in science. The usage of specialized software has been made necessary since this study adheres to a scientific computer-assisted review technique. Biblioshiny, the shiny app for the Bibliometrix R package version 4.2.3 [37–39], and the Visualizing Scientific Landscape software VOSviewer version 1.6.18 [40,41] were used to examine the exported CSV file from the Scopus database and find the most influential articles, authors, and their connections.

3. Data Analysis and Results

3.1. Overview of the Data Collected and Annual Scientific Production

Statistical analysis and network analysis are two distinct types of analysis used to delve into data [42]. The important information gathered from Scopus is detailed in Table 1, which we then integrated into the bibliometric analysis tools.

The number of publications published every year is displayed in Table 2, which demonstrates that the connection between artificial intelligence and healthcare has attracted increasing attention in the literature.
Table 1. Scopus’s primary statistics and document type information.

<table>
<thead>
<tr>
<th>Description</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAIN INFORMATION ABOUT DATA</td>
<td></td>
</tr>
<tr>
<td>Timespan</td>
<td>2019:2022</td>
</tr>
<tr>
<td>Sources (journals, books, etc.)</td>
<td>104</td>
</tr>
<tr>
<td>Documents</td>
<td>171</td>
</tr>
<tr>
<td>Document average age</td>
<td>0.725</td>
</tr>
<tr>
<td>Average citations per doc</td>
<td>8.947</td>
</tr>
<tr>
<td>References</td>
<td>8631</td>
</tr>
<tr>
<td>DOCUMENT CONTENTS</td>
<td></td>
</tr>
<tr>
<td>Keywords Plus (ID)</td>
<td>1767</td>
</tr>
<tr>
<td>Authors’ keywords (DE)</td>
<td>551</td>
</tr>
<tr>
<td>AUTHORS</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>863</td>
</tr>
<tr>
<td>Authors of single-authored docs</td>
<td>4</td>
</tr>
<tr>
<td>AUTHOR COLLABORATION</td>
<td></td>
</tr>
<tr>
<td>Single-authored docs</td>
<td>4</td>
</tr>
<tr>
<td>Co-authors per doc</td>
<td>5.23</td>
</tr>
<tr>
<td>International co-authorships %</td>
<td>30.41</td>
</tr>
<tr>
<td>DOCUMENT TYPES</td>
<td></td>
</tr>
<tr>
<td>Articles</td>
<td>134</td>
</tr>
<tr>
<td>Conference papers</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 2. Annual scientific publications.

<table>
<thead>
<tr>
<th>Year</th>
<th>Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>10</td>
</tr>
<tr>
<td>2020</td>
<td>25</td>
</tr>
<tr>
<td>2021</td>
<td>44</td>
</tr>
<tr>
<td>2022</td>
<td>92</td>
</tr>
</tbody>
</table>

We see a clear upward trend beginning in 2019, and continuous growth afterwards. Additionally, following COVID-19 [43–46], the overall tendency grew. As a result, we may infer from worldwide publishing patterns that the industry is currently going through a phase of steady expansion. Table 3 presents the mean number of citations per publication and per annum. It is noteworthy that the computation of the average total citations per year for 2022 remains incomplete due to the ongoing nature of the year. Consequently, the determination of the mean total citations per year for 2022 is yet to be ascertained.

Table 3. Citation details of scientific publications.

<table>
<thead>
<tr>
<th>Year</th>
<th>MeanTCperArt</th>
<th>MeanTCperYear</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>60.50</td>
<td>20.17</td>
</tr>
<tr>
<td>2020</td>
<td>14.88</td>
<td>7.44</td>
</tr>
<tr>
<td>2021</td>
<td>6.75</td>
<td>6.75</td>
</tr>
<tr>
<td>2022</td>
<td>2.78</td>
<td></td>
</tr>
</tbody>
</table>
3.2. Most Relevant Sources

In the Scopus database, a wide range of document types can be found, including books, conference papers, book chapters, editorials, notes, short surveys, letters, data papers, and journals. However, for the purpose of this study, we specifically selected journals and conference proceedings due to their adherence to a rigorous review process, which ensures a higher standard of article quality. Additionally, we exclusively selected pieces that had reached the final stage of publication. Upon the application of these specific criteria, a total of 171 articles were obtained from various sources. Figure 4 illustrates the ranking of the top 10 sources based on the quantity of articles, with Procedia Computer Science having published the most papers. This subseries also contains the International Joint Conference Proceedings and BMC Medical Informatics and Decision Making.

Table 3. Citation details of scientific publications.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean TC per Art</th>
<th>Mean TC per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>60.50</td>
<td>20.17</td>
</tr>
<tr>
<td>2020</td>
<td>14.88</td>
<td>7.44</td>
</tr>
<tr>
<td>2021</td>
<td>6.75</td>
<td>6.75</td>
</tr>
<tr>
<td>2022</td>
<td>2.78</td>
<td></td>
</tr>
</tbody>
</table>

3.2.1. Most Locally Cited Sources

The volume of citations a text, author, or publication obtains is a measure of its influence. However, older publications will definitely have a higher number of citations compared to the latest publications. But citations are a good factor to consider when checking the impact of publications. Two types of citations are used to check this impact. The number of times a publication is cited in other works from various sources, demonstrating the extent of its influence, is known as Global Citations. Local Citations are references within a publication to other sources that indicate its integration into specific research contexts [47]. According to our statistics, IEEE ACCESS has earned the most citations so far among local sources, and this trend is anticipated to continue in the future (Figure 5).

3.2.2. Source Dynamics

The top five journals’ source dynamics are shown in Figure 6, for which we utilized LOESS (locally estimated scatterplot smoothing) to demonstrate the volume of publications across the time period. BMC Medical Informatics and Decision Making and Frontiers in Neuroscience exhibit a sharp rise in publications from 2019 onwards, whereas the remaining journals exhibit a progressive increase in publications in recent years, notably from 2021 onwards, as seen in Figure 6.
3.2.2. Source Dynamics

The top five journals’ source dynamics are shown in Figure 6, for which we utilized LOESS (locally estimated scatterplot smoothing) to demonstrate the volume of publications across the time period. BMC Medical Informatics and Decision Making and Frontiers in Neuroscience exhibit a sharp rise in publications from 2019 onwards, whereas the remaining journals exhibit a progressive increase in publications in recent years, notably from 2021 onwards, as seen in Figure 6.

3.2.3. Most Relevant Authors

From the standpoint of potential future studies in this field, these works are significant. In bibliometric analysis, finding the most relevant authors is important because it helps us recognize the experts and leaders in a certain field. These authors have made important contributions that guide research and shape ideas. Knowing them helps researchers understand what is important in their area, collaborate better, and make decisions about
funding and resources [48]. In Biblioshiny, there are three frequency measures used to learn about the most relevant authors, namely, the number of documents, the percentage, and the fractionalized frequency. We can use any method we like. The number of documents per author parameter is used here. With a fractionalized frequency of 0.53, Holzinger A has published four articles. This means that their work accounts for 53% of all articles in the domain. Weitz. K, with three publications, has a fractionalized frequency of 0.57, implying a 57% contribution to the total number of articles. And Ahmed S. has two publications with a fractionalized frequency of 0.33, representing a 33% contribution (Figure 7).

3.3. Analysis of Documents by Affiliation

Several prominent associations have emerged as key contributors to “Health Care Trust Evolution With XAI”. The top ten institutions that influenced the total number of publications in current research are shown in Figure 8. The “University of Edinburgh” leads the pack with 16 articles, followed by “Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi)” and “Ruhr-University Bochum”, with 12 apiece. “Augsburg University” and “Imperial College London” both provided nine articles, while “Stanford University” and “University of Antwerp” provided eight. Furthermore, “Aix Marseille University”, “Graz University of Technology”, and the “Mayo Clinic” each provided seven publications. These collaborations weave a rich tapestry of research, emphasizing the investigation of healthcare trust evolution in conjunction with Explainable Artificial Intelligence.

3.4. Most Relevant Countries

This section looks at articles from different countries to see which countries are doing the most research and which are getting cited the most. The United States leads with 109 articles and a substantial number of citations (280). Germany follows with 105 articles, though fewer citations (63), followed by Italy with 94 articles and 123 citations. The United Kingdom boasts 63 articles and 107 citations, signifying active engagement. India, with 50 articles and 9 citations, has room for growth. Spain presents 47 articles and 4 citations, and China has 46 articles and 23 citations, both indicating potential for further exploration. Unexpectedly, South Korea’s 42 articles receive a high citation count (182), demonstrating significant influence. Japan contributes 37 articles with 13 citations, and France offers 33 articles and 164 citations, showcasing their involvement and impact, as shown in Table 4.
Figure 8. Most relevant affiliations.

Table 4. Citation structure per country.

<table>
<thead>
<tr>
<th>Country</th>
<th>No of Articles</th>
<th>Total Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>109</td>
<td>280</td>
</tr>
<tr>
<td>Germany</td>
<td>105</td>
<td>63</td>
</tr>
<tr>
<td>Italy</td>
<td>94</td>
<td>123</td>
</tr>
<tr>
<td>UK</td>
<td>63</td>
<td>107</td>
</tr>
<tr>
<td>India</td>
<td>50</td>
<td>9</td>
</tr>
<tr>
<td>Spain</td>
<td>47</td>
<td>4</td>
</tr>
<tr>
<td>China</td>
<td>46</td>
<td>23</td>
</tr>
<tr>
<td>South Korea</td>
<td>42</td>
<td>182</td>
</tr>
<tr>
<td>Japan</td>
<td>37</td>
<td>13</td>
</tr>
<tr>
<td>France</td>
<td>33</td>
<td>164</td>
</tr>
</tbody>
</table>

Overall, this shows how different countries are involved in researching healthcare trust and AI. It highlights where more work is needed and where there is strong progress.

3.4.1. Co-Occurrence Research for All Keywords

Keywords are short, simple terms in a certain context that define what a paper’s content is about [47,49]. Keywords Plus are additional keywords added to improve discoverability, particularly in academic literature databases [50]. In Figure 9, the top 10 relevant terms are shown, which were derived using Biblioshiny and visualized using Excel.
These results emphasizes the rise in “Explainable Artificial Intelligence (XAI)” for transparent AI outcomes and the significance of “Trust”. “Artificial Intelligence” refers to machine capabilities that are similar to those of humans, whereas “Machine Learning” and “Deep Learning” describe learning processes. Figure 9 also highlights the impact of AI on healthcare through “Digital Pathology”, “Clinical Decision Support Systems”, and “Computer-Aided Diagnosis”. Outside of medicine, “Image Classification” and “Predictive Models” demonstrate AI’s broad application [51]. By fusing innovation and ethical considerations, it enhances the story of AI by encouraging a balance between technological advancement and human comprehension. A tree map is frequently used as a visualization tool in bibliometric analysis to display and analyze the distribution and relationships of diverse bibliographic data, such as publications, authors, journals, keywords, or citations [48]. Figure 10 depicts a tree map illustrating the distribution of keywords. In a bibliometric study, a hierarchical tree map of these terms could show how often a phrase is used based on the size of the rectangles. “Artificial Intelligence” (250) and “XAI” (Explainable AI) (59) are at the top of the list. As we go deeper into “Artificial Intelligence”, we find “Machine Learning” (14), “Algorithms” (13), “Decision Making” (20), and “Prediction” (20). “Convolutional Neural Network” (17), “Decision Trees” (13), “Support Vector Machine” (14), and “Black Boxes” (14) are all in the “Machine Learning” part. There are 30 “Diagnosis” items, 27 “Learning Systems” items, 20 “Major Clinical Study” items, and 13 “Diagnostic Accuracy” items in the “XAI” group. Lastly, “Algorithm” (19), “Adult” (18), “Nuclear Magnetic Resonance Imaging” (14), and “Forecasting” (15) are included. This visually organized tree-map helps readers understand how important each term is and how it fits into the hierarchy of the bibliometric dataset. This helps us to find the most important topics and trends in the field.

A network map can be shown visually by using scientific mapping techniques that examine text data such as keywords collected from titles and abstracts [52]. Each term in this map is represented as a node, and the connections between them are shown as edges between the nodes. When two nodes are connected, it means that the relevant terms are related. Greater significance is indicated by nodes that are bigger and more closely spaced in relation to one another. To depict the co-occurrence of author keywords, VOSviewer was used, and keywords used by authors in research articles were analyzed. We focused on keywords that appeared at least three times, which gave us 36 keywords. We then
measured how often these keywords were mentioned together in articles to find their strongest connections. The top keywords with the strongest connections were selected. All these keywords were then grouped into nine clusters based on their frequent co-occurrence, helping researchers identify important themes in the research, as seen in Figure 11.

The grouping of keywords like “Brain tumor”, “Convolutional Neural Network”, “LIME [7]”, “SHAP”, “COVID-19 [53]”, “Grad-CAM”, and “Digital Pathology” under the red cluster likely depicts a thematic focus within the field of medical research and image analysis. This could imply research into the application of advanced machine learning techniques for diagnosing brain tumors, understanding COVID-19’s effects on the brain, and improving the interpretability of AI models in medical contexts.
Every cluster suggests different research to be conducted, such as the green cluster, which contains keywords such as Semantic Web, Explainable Artificial Intelligence (XAI), deep learning, magnetic resonance imaging, and image classification. Based on these keywords, here are some research suggestions:

- Can the accuracy and reliability of MRI diagnoses be enhanced by incorporating Explainable Artificial Intelligence (XAI) in conjunction with deep learning methods for image categorization, within the context of the Semantic Web framework?
- When XAI (Explainable Artificial Intelligence) and deep learning techniques are employed for the purpose of classifying MRI (Magnetic Resonance Imaging) images, some ethical concerns arise. Future researchers can dig into the ethical issues and propose potential strategies to mitigate them. In what ways may the application of Semantic Web principles facilitate the efficient organization and retrieval of data, while simultaneously upholding the ideals of patient privacy and informed consent?
- How does implementing XAI within Semantic Web-driven clinical decision support systems affect user trust and acceptance in AI-driven diagnostics, particularly for MRI image classification, and what cross-domain knowledge transfer opportunities exist to improve model performance [54]?
- How might the utilization of XAI approaches, specifically LIME and SHAP, contribute to the improvement of interpretability in Convolutional Neural Networks (CNNs) within the field of digital pathology, ultimately leading to enhanced accuracy in disease detection?
- What are the potential biomarkers for the early diagnosis of Alzheimer’s disease utilizing machine learning (ML) and deep learning (DL) models, and how may XAI techniques enhance their interpretability?
- In what ways may active learning methodologies be utilized to train Artificial Neural Networks (ANNs) for MRI-based diagnoses, with the aim of enhancing user trust and confidence in AI-driven healthcare decisions?

3.4.2. Network for Co-Citation

A method for mapping the body of scientific literature known as co-citation analysis makes the assumption that works that are frequently referenced together have comparable themes [55]. Using this approach, one might learn about a research area’s fundamental themes and other intellectual underpinnings. When two papers are frequently cited in conjunction with one another, this is known as co-citation (Figure 12). When two publications appear in the reference section of another publication, the two publications are connected in a co-citation network.
In this analysis, a source-based co-citation network is presented (Figure 13). In the red cluster, *IEEE Access* has the most significance. *Nature* is the most co-cited journal in the blue cluster.

![Source-based co-citation network](image)

**Figure 13.** Source-based co-citation network.

The author-by-author co-citation network [56] is depicted in Figure 14. All authors are grouped into two clusters, namely, Red and Blue. Wang and Ribeiro were the highest co-cited authors. Therefore, future researchers can benefit from articles authored by these researchers.

![Author-by-author co-citation network](image)

**Figure 14.** Author-by-author co-citation network.
3.5. Conceptual Structure

Academics frequently utilize conceptual structures to understand the issues addressed by academics (so-called research fronts) in order to determine which are the most current and important. Conceptual structure gives insight into the topography of a scientific topic by clustering a binary tree network of terms built from keywords, titles, or abstracts [56].

Thematic Map

Figure 15 illustrates the two-dimensional typological themes, commonly known as a thematic map [56]. The development of themes within a study topic is facilitated through the utilization of co-word analysis, a method that identifies clusters of keywords. The aforementioned themes can be categorized into four quadrants on a graph that encompasses two dimensions, namely, centrality and density. Each subject is visually represented by a bubble on the map, and correspondingly, each subject is graphically depicted as a bubble on the graph. One prominent subject that currently holds significant importance within the field and has garnered considerable attention in recent research is the theme situated in the upper right quadrant. This theme encompasses various aspects such as Convolutional Neural Networks, disease prediction, diagnostic accuracy, AI algorithms, nuclear magnetic resonance imaging, and support vector machines with high density and centrality. The lower left quadrant of the diagram highlights growing issues in the field of artificial intelligence (AI), explainable AI (XAI), diagnosis, forecasting, black-box algorithms, and machine learning.

![Thematic Map Diagram]

Figure 15. Thematic map.

3.6. Social Structure

International academic research collaboration is increasing. As a result, research collaboration can be summed up as researchers cooperating to produce new scientific information [57]. On the most fundamental level, people interact more than institutions. The
primary building block of collaboration is direct cooperation between two or more scholars. However, we frequently discuss cooperation at other levels, including that between research groups within departments, departments within the same institution, institutions, sectors, and geographical regions and countries. Academic research is increasingly conducted abroad, whether it is to locate specialized tools, produce novel ideas, or discover new funding sources [58]. With the threshold set through the use of at least three research papers and three citations each using VOSviewer version 1.6.19, out of 51 countries, 24 really stood out. All these countries were divided into five clusters based on their bibliographic coupling, India, Greece, Germany, Canada, and Austria come under the green cluster, as shown in Figure 16.

![Figure 16. Country-wide collaboration network.](image)

Developed countries made the largest contribution to the literature on this topic, which depicts the contributions of various nations to the discipline. Most of the publications in the region were from the United States and Germany. The total number of citations the United States has earned in this field shows how much research is being conducted there. The fact that India has partnerships with developed nations and is low on the production index could indicate that developing nations are slowly but steadily advancing towards high-quality studies in this field.

4. Conclusions

Research Findings: The contribution of each feature to black-box prediction is captured by XAI approaches like LIME and SHAP, which estimate feature attributions in individual instances. Second, the visualization of bibliometric networks, such as co-citations, couplings between bibliographies, keyword occurrences, and co-authorship networks, will aid researchers in understanding their future work.

Limitations: We need to be aware of the constraints of our efforts. As a starting point, even if we are certain of a single database, future research will use additional databases, such as WoS, to examine more potential papers. The wide variety of papers necessary for
our study were available in Scopus thanks to its sizable database. In order to increase the number of articles relevant to this study, the keyword search might be changed to include additional terms. The results of this study might be reinforced by a qualitative analysis of the articles via careful reading of the complete texts in order to gain a deeper understanding of the subject area, as this study is only of a correlational and quantitative character.

Concluding Remarks: Using the widely known database Scopus, a bibliometric study on explainable AI in the medical field was carried out. This database is thought to have begun in 2019. The AND operator was used in conjunction with the keyword search to search the database. The search resulted in 171 documents being obtained in total. The study of this database took a number of factors into account. It is evident that the majority of documents were written in English. According to the keyword search results, the most publications contained the phrase Explainable Artificial Intelligence. Although the 2022 year is not yet complete, in this study, we considered all articles that were in the final publication stage or in the press to have a look at this emerging technique, even though the present year (2022) had the most documents (92), followed by the year 2021. Nearly 63.74% of the documents fell under the category of computer science. Journal articles accounted for 134 of the papers and conference papers accounted for 37 when it comes to the type of document. The United States had the most documents over the time period, according to the examination of the various nations. Biblioshiny and VOSViewer were used to carry out the network analysis. Many analysis types were carried out. These utilized the same database and included co-authorship analysis, co-occurrence analysis, keyword analysis, and bibliographic coupling. These various network analyses revealed some very important information regarding the many subjects discussed above. Additionally, it is clear that the two years of 2021 and 2022 saw the bulk of the effort in this area related to medical imaging. This analysis also showed current health-related AI research trends, indicating that in recent years, the growth rate of publications on AI in the field of healthcare has increased significantly, and this rate has been rising consistently.


Funding: This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia, under Project GRANT4245.

Data Availability Statement: Not applicable.

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

References


