sustainability

MDPI

Review

Survey on the Biomedical Text Summarization Techniques with an Emphasis on Databases, Techniques, Semantic Approaches, Classification Techniques, and Similarity Measures

Dipti Pawar [1], Shraddha Phansalkar [1,*], Abhishek Sharma [2], Gouri Kumar Sahu [3], Chun Kit Ang [4] and Wei Hong Lim [4,*]

[1] Department of Computer Engineering, MIT Art, Design and Technology University, Pune 412201, India
[2] Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248002, India
[3] Department of Physics, Centurion University of Technology and Management, Bhubaneswar 761211, India
[4] Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia
* Correspondence: shraddhaphansalkar@gmail.com (S.P.); limwh@ucsiuniversity.edu.my (W.H.L.)

Abstract: Biomedical text summarization (BTS) is proving to be an emerging area of work and research with the need for sustainable healthcare applications such as evidence-based medicine practice (EBM) and telemedicine which help effectively support healthcare needs of the society. However, with the rapid growth in the biomedical literature and the diversities in its structure and resources, it is becoming challenging to carry out effective text summarization for better insights. The goal of this work is to conduct a comprehensive systematic literature review of significant and high-impact literary work in BTS with a deep understanding of its major artifacts such as databases, semantic similarity measures, and semantic enrichment approaches. In the systematic literature review conducted, we applied search filters to find high-impact literature in the biomedical text summarization domain from IEEE, SCOPUS, Elsevier, EBSCO, and PubMed databases. The systematic literature review (SLR) yielded 81 works; those were analyzed for qualitative study. The in-depth study of the literature shows the relevance and efficacy of the deep learning (DL) approach, context-aware feature extraction techniques, and their relevance in BTS. Biomedical question answering (BQA) system is one of the most popular applications of text summarizations for building self-sufficient healthcare systems and are pointing to future research directions. The review culminates in realization of a proposed framework for the BQA system MEDIQA with design of better heuristics for content screening, document screening, and relevance ranking. The presented framework provides an evidence-based biomedical question answering model and text summarizer that can lead to real-time evidence-based clinical support system to healthcare practitioners.

Keywords: text summarization; databases; semantic enrichment; text similarity; biomedical question answering

Citation: Pawar, D.; Phansalkar, S.; Sharma, A.; Sahu, G.K.; Ang, C.K.; Lim, W.H. Survey on the Biomedical Text Summarization Techniques with an Emphasis on Databases, Techniques, Semantic Approaches, Classification Techniques, and Similarity Measures. *Sustainability* **2023**, *15*, 4216. https://doi.org/10.3390/su15054216

Academic Editor: Andreas Kanavos

Received: 19 January 2023
Revised: 22 February 2023
Accepted: 23 February 2023
Published: 26 February 2023

1. Introduction

Over the past few decades, there has been an exponential increase in the amount of biomedical information in the scientific literature. This contains syntactically and semantically heterogeneous data which makes different sense in different contexts. Gaining concise information from this heterogeneous data is the biggest challenge in the healthcare community. Hence many works [1–7] by researchers and clinical practitioners establish the need for automated tools to summarize biomedical literature.

There is a method to obtain a condensed form of a given input document by preserving its overall meaning and important content [8,9] called as text summarization (TS). In the area of biomedical effective TS, tools play a crucial role in effectively obtaining the abstract or condensed form in a specific area with a summary from single or multiple text

documents of huge sizes [10]. TS approaches can be mainly classified into extractive and abstractive. Extractive summarization aims to identify and extract the most relevant text from input text whereas abstractive summarizer makes use of natural language processing (NLP) and generating methods to interpret given text to infer and build an abstractive summary [9]. As we can see most biomedical information is represented with structured patient electronic health records (EHR) and represented with semi-structured knowledge sources such as ontologies, dictionaries, etc., hence the extractive approach of text summarization is preferred for the clinical domain [11–13].

As shown in Figure 1 there are multiple application areas of TS in the biomedical area such as scientific literature summarization [4,6,7], treatments information summarization [11], medicine information summarization [14], EBM practice and clinical decision making [15–17], summarization of patient's clinical summary notes [18], summarization of patients electronic medical records (EMR), EHR summarization [12], etc., are addressed. An effective biomedical text summarization method should be able to handle the diversities and generate meaningful insights for clinicians, researchers as well as patients by addressing these challenges in the biomedical domain.

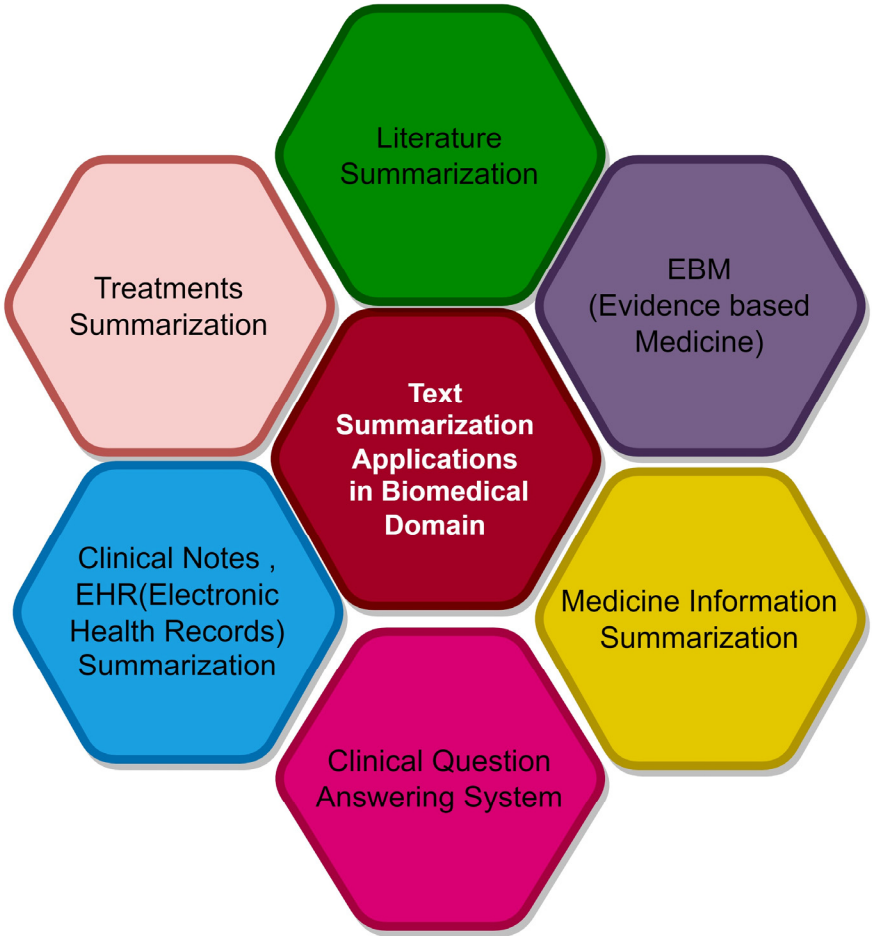

Figure 1. Application areas of text summarization in the biomedical domain.

The literary works [1–5] in recent year shows widely investigated text summarization and proposed systems to help biomedical users and clinical practitioners deal with these hassles in getting crucial information contained in a huge volume of biomedical textual information sources [19–21]. The summarization approaches are also preferred in most of the BQA systems for mining relevant data from the corpus.

The majority of biomedical text information on the Internet is in one of three formats (1) EHR for patient records, (2) EMR with patient clinical trials, and (3) biomedical literature databases with medical vocabularies.

The most important source of biomedical textual information is the biomedical literature [21–24], which provides a huge knowledge base to clinical practitioners and researchers through reviews, statistics, and experiments. Other sources for biomedical text summarization are clinical trials [20], EMR [25–28], and clinical documents.

Biomedical text summarization helps evaluate treatment methods, analyze trends and patterns, compare treatments, and study advancements in a particular study, develop and test new hypotheses with the conduction of experiments, and interpret their results. Although the research in biomedical text representation standards is evolving significantly with ontologies, semantic maps, and dictionaries, the challenge in their common interpretation and understanding persists. The huge volume, variety, and heterogeneity in the biomedical text challenge text summarization techniques in this domain.

The literary work [13,29–32] shows many methods such as NLP and text mining have been developed for the tasks such as information extraction, knowledge discovery, and text processing. Even though advances in information retrieval techniques have proven helpful in assisting clinicians and clinical researchers in managing information overload [33,34], they still have to go through multiple text documents to find meaningful and relevant information that meets their needs. Additionally, the biomedical text is highly diverse and rich concerning vocabulary and lacks standardization in semantics. Moreover, the challenge of understanding the context and retrieving correct information still stands unaddressed [35]. Effective automated text summarization is thus a challenge in a broader area of the biomedical domain as it involves searching relevant, scientifically sound information, and ever-increasing volume of heterogeneous information contained in the biomedical literature and clinical information sources. Our survey provides a thorough review of the recent research concerning biomedical text summarization, presenting their state-of-the-art works respectively to the important artifacts. Additionally, this review also focuses on the most popular application of biomedical text summarization, BQA systems which leverage summarization and classification techniques.

1.1. Significance and Rationale

A huge amount of biomedical data are available for biomedical research and investigations. The data are obtained from multiple heterogeneous and discrete information sources such as scientific literature, data collected through clinical trials, clinical summary notes of patients, and EHR and EMR of patients. EBM practice is an integration of quality research evidence with clinical expertise for clinical decision-making [15,16]. So, as per the EBM [36], practice clinicians need to consider the best evidence to provide optimal care to patients, for which they need to effectively retrieve, interpret, and integrate significant information from various medical knowledge sources [21–23].

However, it is challenging for biomedical users to effectively deal with a huge volume and variety of textual information from growing biomedical resources. Though some advances have been achieved in information retrieval technology to assist biomedical users with their information needs, they need to deal with several documents to obtain abstracts or summaries of the required topic [4,33,35] from the retrieved corpus. Automatic text summarization (ATS) plays a very crucial role in the biomedical domain to assist clinical researchers and practitioners, patients who are seeking information to obtain a summary of information from single or multiple documents. It is the process of condensing input documents to a shorter version or summary which presents meaningful information without redundancy. ATS [8] is a promising approach to help clinical users to retrieve and process pertinent information more effectively and precisely in the biomedical domain. This became a crucial tool to assist clinical researchers, practitioners, and patients with their information extraction and knowledge discovery tasks [1,5]. This systematic literature

review needs to investigate various challenges in ATS in the clinical area, analyze the maturity of its solutions, and study it in a condensed manner with all important artifacts.

1.2. Motivation and Applications

Text summarization [29] has come out as a promising solution to bridge the gap between unstructured text and structured visualization of clinical information [31,37,38], which processes huge text collections using artificial intelligence techniques including NLP, machine learning (ML), and deep learning (DL). These computational techniques facilitate clinical research and have made significant progress in the field of clinical research [30]. NLP-based tasks in the biomedical domain employ many ML and DL algorithms to extract structured data from intricate, diverse clinical reports [31,32].

Biomedical text summarization and BioNLP is an interesting sub-field in the clinical research domain which deals with processing information from scientific journals, clinical health records, patients' clinical notes, and other biomedical documents. There has been an increasing interest in BioNLP for extracting information, relationships, and insights from text data [37].

The biomedical question–answering system that is employed in biomedical text summarization helps present concise, relevant, and summarized answers to the questions with better accuracy and evidence. The choice of appropriate literature, dictionaries, and knowledge corpus is significant for this study. Additionally, choosing appropriate text similarity metrics, and ranking techniques help retrieve the most relevant solutions to the questions. Context-aware semantic analysis plays a very important role in biomedical text summarization. Although there is an easy reach to a variety of biomedical knowledge sources, finding semantically relevant answers to clinical queries is a challenge, hence the study of the context-aware text summarization model is important in BTS. To avoid any misinterpretation, the study of the semantic similarity finding techniques plays a crucial role in the biomedical domain, because it ensures not only obvious clinical context but also implicit relationships between text sequences. To the best of our knowledge, we are presenting a complete review that involves in-depth study of all the major artifacts in the biomedical text summarization.

Moreover, our work outlines various popular applications of automatic BTS in Figure 1. Nowadays BTS has been greatly used in various applications based on biomedical information retrieval and extraction, BQA systems [39], text mining, and analytics. Moreover, the optimization in a search engine can be made broader with BTS in various applications such as EBM practice [15], telemedicine [38], BQA system [40–42], clinical decision support [36], etc.

1.3. Objectives

We aim to unveil a review of the state-of-the-art literary works in the area of BTS concerning important artifacts. Research challenges and discover the future area of work in the BTS domain. The objectives are stated as:

- Study of the significant highly cited biomedical databases, the search filters, and the query strings used in the survey of BTS.
- Study of the popular clinical knowledge sources, ontologies, dictionaries, vocabularies, and their applications in BTS for semantic enrichment of text.
- Listing of the different similarity metrics used in BTS.
- To study recent literary works in the BTS.
- Study of the state-of-the-art literary works in BQA systems to investigate the challenges in the domain.

1.4. Prior Research

As the amount of information available to clinicians and biomedical researchers grows tremendously, research in BTS is becoming increasingly relevant. In the biomedical areas, ATS condenses information in an attempt to assist clinical users in rapidly and accurately

locating and understanding important source materials. In the biomedical arena, there has been a significant increase in the number of studies undertaken to design and assess various text-summarizing algorithms in recent years. The purpose of this research was to conduct a comprehensive analysis of key work on a textual document summarizing the biomedical area.

An early stage review of the literature was conducted by Afantenos et al. [2]. They have summarized ten biomedical published research works from 1999 to 2003. However, the focus of the work was on general text-summarizing techniques, and summarization factors such as input, output, and evaluation methodologies, with less emphasis on the challenges in the BTS.

Rashmi Mishra and Jiantao Bian [1] 2014, studied and summarized over ten biomedical published research works from 2002 to 2013. The study provides systematic literature review (SLR) on techniques for text summarization and a thorough evaluation of the works on text-summarizing techniques in this field. The work aims in identifying various methodologies, wide applications, and assessment processes.

Milad Moradi, Nasser Ghadiri [3] 2019, presented a review of the common tasks that basically make use of TS and recent developments in the area of BTS. The majority of the papers in this review study focused on summarizing biomedical literature and addressing the issues associated with that. Earlier works focused on only general text summarization techniques and did not emphasize EBM practice.

Mengqian Wang, Manhua Wang [7] 2021 presented and analyzed the methodologies, application areas, as well as assessment techniques utilized in each of the most recent BTS studies on biomedical literature and EHR. They followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [43] methodology to identify 58 studies from 1 January 2013 to 8 April 2021.

Andrea Chaves, Cyrille Kesiku [8] 2022 presented and analyzed methodologies, areas of application, and assessment methods used in the most recent works on BTS. They followed the PRISMA methodology and identified 28 studies from January 2014 and March 2022.

Despite the fact that the number of studies on EHR summarization has shown an increase, most works still focus on the literature summarization. Unlike our survey which discusses work on BTS on basis of a database, semantic enrichment approaches, text summarization techniques, and similarity metrics, this work analyses the articles collected by considering dimensions such as input, aim, output, technique, and assessment. The presented survey reviews recent works on BTS with more emphasis on

- Databases: For appropriate text summarization of biomedical text documents, it is necessary to investigate and explore databases with their applications and structures.
- Semantic Enrichment Approaches: As semantic enrichment plays a very crucial role to obtain contextual relations between text sequences our survey focuses on various approaches for semantic enrichment.
- Text Similarity Metrics: survey focus on commonly used textual similarity metrics in the biomedical domain.
- Text Summarization Techniques and Applications: a comparative analysis of various text summarization systems with an enhanced emphasis on biomedical Question answering systems.

A comprehensive examination of current developments in the field of BTS is important to examine the newfangled because researchers in this area are constantly identifying new difficulties and addressing them with unique approaches. This can support familiarizing with new challenges and problems, the most effective solutions, and the most important outcomes produced through evaluation methods. Overall, this comprehensive review gives an overview of contemporary research that is pushing the frontiers of BTS and posing some new issues that have yet to be uncovered and solved. As a result, it can be a very useful place to start for people who want to explore BTS and its research trends.

Since then, there have been substantial advancements in the biomedical domain's summarization tools and approaches. A trivial contribution towards SLRs on the topic of ATS in the biomedical area has been published to our knowledge.

This SLR paves the way for more research in this field in the future. This paper tries to provide a comprehensive list of databases, methodological trends, validation methodologies, evaluation criteria, and publicly available biomedical text-summarizing systems. A systematic review will aid in the better knowledge of literary works, determine gaps, and point researchers on the right path for future research. We systematically reviewed TS methods used in medical literature and EHR models in the current work. The SLR aims to: (1) Identify distinct methodology, application areas, and assessment methods during the previous period; (2) recognize study trends; (3) identify study shortcomings; and (4) make endorsements to lead the upcoming studies.

1.5. Research Goal

This study focuses on artifacts such as biomedical databases, semantic enrichment approaches, different techniques of BTS, BQA [44] systems, and textual similarity measures to analyze existing knowledge about, guidelines, tools, and methods in realizing the BTS framework. The research questions proposed are stated here in Table 1.

Table 1. Research Questions for literature review.

RQ No.	Research Question (RQ)	Objective/Discussion
RQ1	What are the various biomedical databases available online for automatic biomedical text summarization?	For appropriate text summarization of biomedical documents, it is necessary to investigate and explore databases, their application, structure, and query techniques.
RQ2	What are the different Semantic Enrichment Approaches used in biomedical text summarization and their comparative evaluation?	This area must be in order to determine the significance of their application in summarization techniques.
RQ3	What are the different similarity metrics used in biomedical text summarization?	A comprehensive analysis of existing textual similarity measures that can be used in biomedical text-summarizing systems is carried out
RQ4	What are the various approaches for automatically summarizing biomedical text, and how are they compared?	A systematic review was conducted with a comparative study of existing systems, taking into account the techniques, feature extraction methods employed, and performance in the form of accuracy.
RQ5	What are the different approaches used for automatic biomedical QA systems and their comparative analysis?	The BQA system is one of the most prevalent and significant applications of the BTS system. As a result, a comparative analysis of various biomedical Question answering systems was necessary, taking into account significant factors

1.6. Contribution of Work

We lay down contributions of the work as:

1. This study examines the underlying theories and evolution of automatic biomedical text-summarizing systems by conducting a systematic literature review.
2. The analysis of current databases, feature extraction techniques employed, semantic enrichment approaches, text summarization approaches and algorithms, assessment metrics, and challenges are part of the survey.
3. Based on the current approaches, question processing, and formulation techniques, passage retrieval and answer processing methods, and datasets, the review compares various existing BQA systems. Furthermore, the limitations of such systems are explained in this work.

4. The study concludes with the identification of present issues and challenges in biomedical ATS architectures, as well as future research goals.
5. The work culminates in proposing a framework of a biomedical question answering system using the potential of text summarization on the biomedical corpus. The study of research gaps in the discussion section shows the scope for the design of automated BQA with unique features such as heuristics for sentence extraction, Document Screening, and Context-Aware Semantic Enrichment technique.
6. As shown in Figure 2 the rest of this paper is organized into distinct segments.

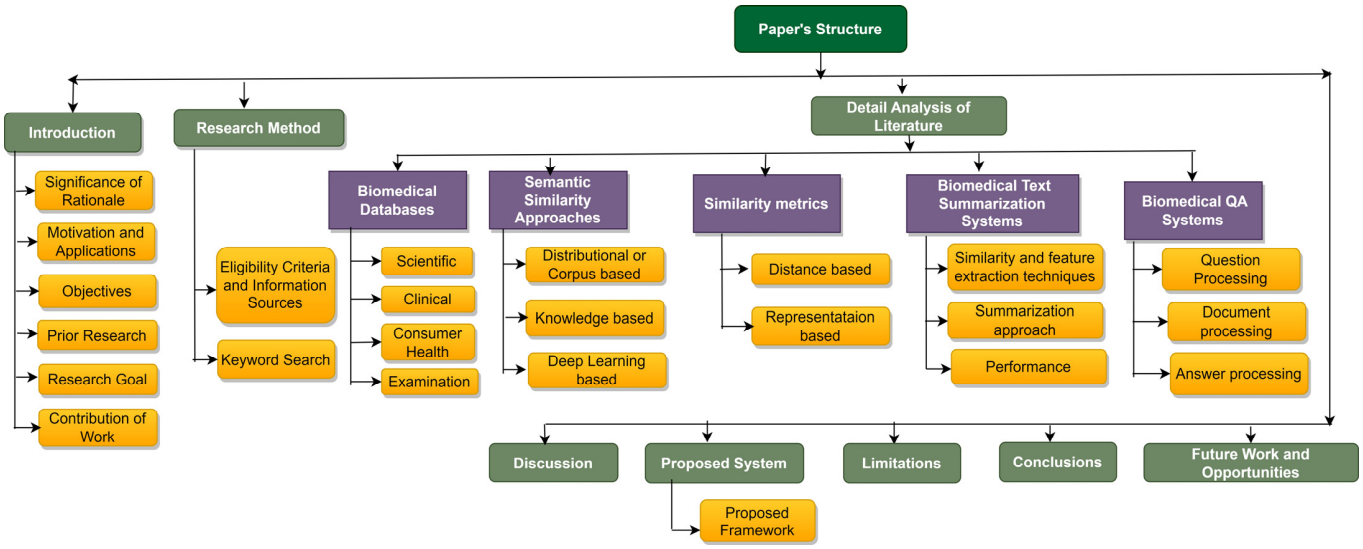

Figure 2. Organization of systematic literature review.

2. Research Method

As we have investigated biomedical ATS, an in-depth and efficient literature study of the existing literary works was conducted to address the RQs identified. Here, the study of the needed academic collection between 1998 to 2022 is done. We followed an SLR approach for the conduction of our survey. The objective of this research was to look at the discoveries of a few basic inquiries about disciplines. The PRISMA workflow chart is utilized to put together the fundamental materials for this consideration. Figure 3 delineates the PRISMA [43] strategy for this study.

2.1. Eligibility Criteria and Information Sources

Existing systems on text-summarizing applications for biomedical domains were surveyed because our focus was on biomedical text summarization. Articles for biomedical text-summarizing were retrieved from multiple biomedical databases. Table 2 presents the mentioned norms of inclusion and exclusion for different categories of searches. For searches, we used the Pub-Med interface and the Google search engine.

2.2. Keyword Search

The keywords for our study were BTS, semantic enrichment techniques, biomedical question-answering systems, and textual similarity measures. Initial queries on the PubMed database and other extra sites yielded a number of studies relevant to our concerns. After removing duplicates and screening, we used inclusion/exclusion criteria to narrow down the studies, as shown in Figure To retrieve the research articles from the databases, a specific query was created. biomedical", "text", "summarization", "clinical", "medical", "passage", "question responding", and "summary" were utilized to build the query. As stated in Table 3, a multiple database search strategy was used.

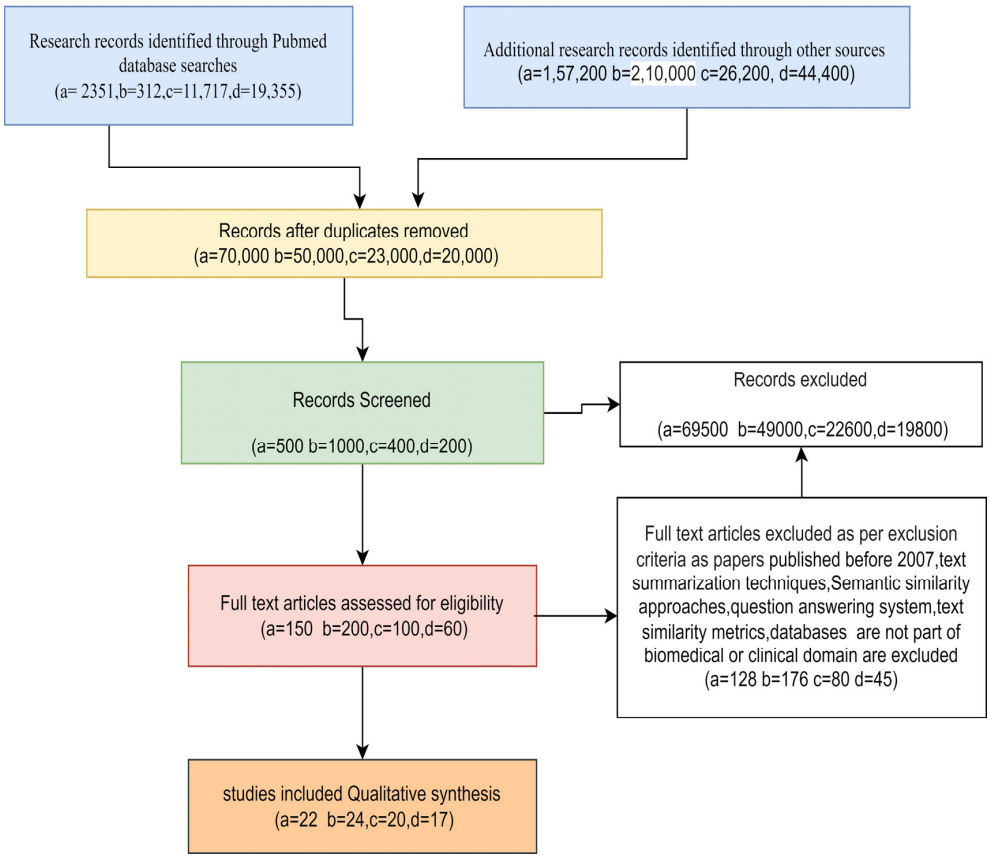

Figure 3. The PRISMA stream chart for precise audit specifying look, the number of modified works/records screened, and the quantity of full-text articles recovered where a = biomedical text summarization, b = semantic similarity, c = biomedical question answering, d = similarity metric.

Table 2. Inclusion and exclusion criteria for literature review.

Topic of Study	Inclusion Criteria	Exclusion Criteria
Biomedical Databases	The work must refer to highly cited biomedical or clinical databases	Papers that used databases other than biomedical or clinical domains.
Semantic Enrichment	The work must focus on highly cited semantic similarity approaches to compute the similarity between biomedical terms or text using available biomedical knowledge sources.	The papers which are presented applied semantic similarity approaches for non-biomedical domains are excluded
Text Similarity Metric	The work must focus on the most popularly used text similarity metrics in the biomedical domain	The papers focus on metrics used in domains other than the biomedical or clinical domain
Biomedical Text Summarization	The work must refer to various techniques used for automatic text summarization in the field of biomedical or clinical domains.	The automatic text summarization methods are applied to non-biomedical documents. The work focuses on the automatic summarization of inputs other than text such as video summarization, dialog summarization
Biomedical Question Answering System	The work must focus on question-answering systems in the biomedical or clinical domain with various techniques applied.	The papers presented on non-biomedical QA systems.

Table 3. Keywords Search.

Database	Query Executed
SCOPUS	TITLE-ABS KEY ((“biomedical text” OR “document” OR “Biomedical” OR “clinical notes” OR “biomedical domain” OR “biomedical literature” OR “clinical” OR “medical” OR “medical records” OR “clinical records” OR “semantic similarity”) AND (“summarization” OR “text summarization” OR “summary” OR “patient” OR “EHR”) AND (“passage retrieval” OR “question answering OR “graph-based” OR “machine learning” OR “transformer based” OR “Evidence-based Medicine(EBM)” OR” deep learning” OR “databases” OR “knowledge base” OR “knowledge sources” OR “metric” OR “measure”))
Web of Science	TOPIC ((“biomedical text” OR “document” OR “Biomedical” OR “clinical notes” OR “biomedical domain” OR “biomedical literature” OR “clinical” OR “medical” OR “medical records” OR “clinical records” OR “semantic similarity”) AND (“summarization” OR “text summarization” OR “summary” OR “patient” OR “EHR”) AND (“passage retrieval” OR “question answering OR “graph-based” OR “machine learning” OR “transformer based” OR “deep learning” OR “Evidence-based Medicine(EBM)” OR “databases” OR “knowledge base” OR “knowledge sources” OR “metric” OR “measure”))
IEEE	((“biomedical text” OR “document” OR “Biomedical” OR “clinical notes” OR “biomedical domain” OR “biomedical literature” OR “clinical” OR “medical” OR “medical records” OR “clinical records” OR “semantic similarity”) AND (“summarization” OR “text summarization” OR “summary” OR “patient” OR “EHR”) AND (“passage retrieval” OR “question answering OR “graph-based” OR “machine learning” OR “transformer based” OR “deep learning” OR Evidence-based Medicine(EBM)” OR “databases” OR “knowledge base” OR “knowledge sources” OR “metric” OR “measure”))
Pubmed	(1) Biomedical Text Summarization (((((((biomedical text summarization) OR (clinical summary)) AND (biomedical document)) AND (medical documents)) AND (biomedical literature)) OR (automatic text summarization)) OR (clinical records)) OR (clinical notes)) AND (biomedical) (2) Semantic similarity approaches (((biomedical text similarity) OR (semantic similarity)) OR (biomedical domain)) AND (similarity measures) OR (semantic enrichment)) (3) Biomedical Question answering systems ((((biomedical QA)) OR (question answering)) OR (passage retrieval)) OR (biomedical domain) (4) similarity metrics ((((similarity measures) AND (text)) OR (document)) OR (NLP)) OR (text similarity metrics) OR (text similarity) OR (biomedical domain)

3. Detailed Analysis of Literature

3.1. Biomedical Databases

As shown in Figure 4, there are different forms of biomedical information. Because of the constant advancement in biomedical research and development in the recent decade, the volume of biomedical literature and patient EHR, and EMR articles is getting enormous. As a result, it is extremely difficult for doctors, patients, and clinical researchers in biomedical-related disciplines to construct their own knowledge base and keep it up to date on a daily basis by identifying all the relevant literature that is released. For EBM practice, clinicians should know about the latest clinical proof for diagnosing and treating patients' infirmities [15] in order to give optimal patient care as well as the general public interest in obtaining their own health on the Internet. Biomedical information procurement is a basic errand with regard to data recovery and information, so biomedical experts and the overall population, for instance, need important help in recovering, deciphering, and incorporating applicable data from numerous biomedical information sources. Clinicians must be able to search for and obtain the most recent clinical data, as well as keep up with the recent advancements in their field of expertise, in order to follow EBM. Electrical possessions, for example, online biomedical works databases, EHR, and EMR systems have been developed to assist biomedical researchers, doctors, and patients with their information management needs. Each scientific work is documented in data produced according to guidelines established by a number of scientific associations and institutions [40].

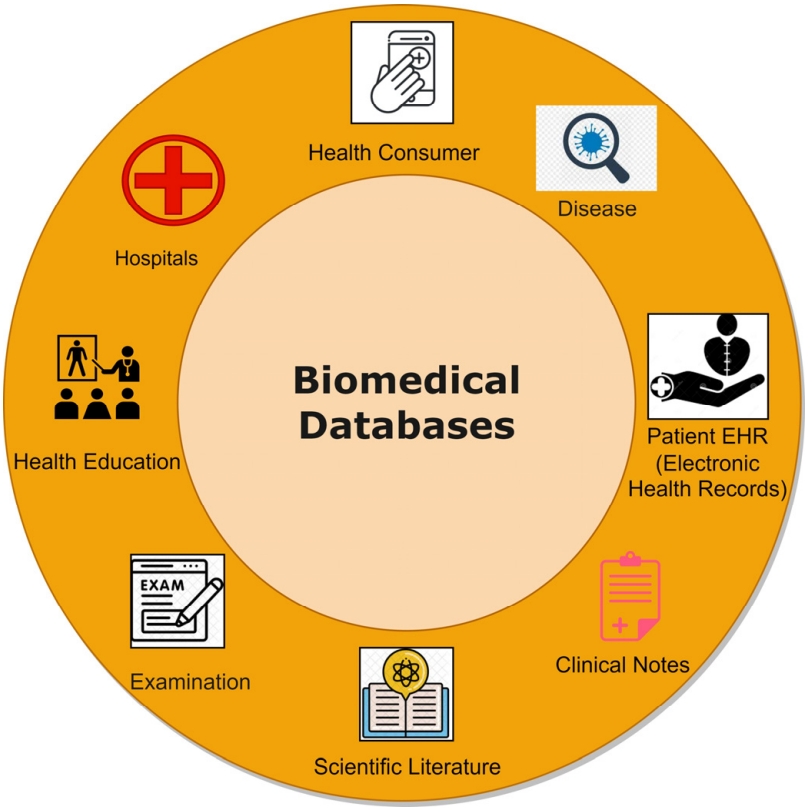

Figure 4. Different forms of biomedical information.

The various information bases are additionally used to gather and handle the best logical and proficient articles, as well as assessments and contextual investigations, from scientific journals and other publications. The producers of databases can depend on the information's exactness and quality. Numerous significant biomedical data sets are housed in notable colleges and scholarly foundations, such as the National Library of Medicine (NLM), the Institute of Medicine (ISI), Elsevier (Amsterdam), and (Ipswitch) (EBSCO), University of Melbourne (BMC), UK (Cochrane Library), Cambridge (Physio-net), USA (SEER), Stanford University (Bio-portal), etc. The majority of them are shown in Table 4, with content type and search queries taken into account. Based on the information provided by the content type, which determines the type of material contained in the database, it can be split into four groups [41] as: A. scientific database, B. clinical database, C. consumer health database, D. examination database. Information regarding biomedical breakthroughs can be found in the scientific database. Patients' EHRs, EMRs, and clinical notes literature are all included in the clinical databases. The consumer health database contains information on a variety of topics related to consumer data. Information about medical license examinations can be found in the examination database [45]. Numerous studies and improvements have been made to the search, classification, and presentation of texts that contain vital information, as shown by search tools included in websites such as PubMed [21], The Cochrane Library [25], and Trip Database [46].

Table 4. Significant databases in the biomedical domain.

Name	Description	Content-Type	Search
PubMed [21]	MED-LINE contains over 33 million citations and abstracts for biomedical literature.	Scientific	Informal keyword searching. Automatic mapping to Mesh terms. Narrow search results with better results. Search for clinical trials, efficient audits, and therapeutic hereditary quality themes.
Pub-Med Central (PMC) [22]	The National Center for Biotechnology Information is in charge of it (NCBI) Full-text digital repository of biomedical and life sciences magazine articles (NCBI)	Scientific	Advanced Search Builder search by keywords, author, journal, etc. Combined search by Boolean operators.
BioMed Central [47]	publishes about 300 peer-reviewed publications that communicate research findings from scientific, technological, engineering, and medical research teams and is part of Springer Nature	Scientific	Allows a number of searches that can only be done using templates.
Ebase [48]	An Elsevier subscription Comparable content as PubMed/MEDLINE Extra consciousness on capsules and pharmacology, clinical devices, scientific medicine, and primary technological know-how applicable to scientific medicine.	Scientific	Quick search by title/abstract/author keywords. Combined search by Boolean operator "OR" keywords from EMBASE (=EMTREE).
The Cochrane Library [25]	Group of databases in medicinal drugs and different healthcare specialties, well-conducted controlled trials	Clinical	Search by Title, Abstract, or Keywords. Keywords are called EMTREE terms, Mesh terms, and other keywords.
CINAHL [26]	A database of nursing and allied health writing is called the Cumulative Index to Nursing and Allied Health Literature (CINAHL). 3604 active indexed and abstracted journals	Scientific	Search by title, abstract, and keywords. Effective search by using subject headings. Search allows several synonyms, divided by OR, and answers with double inverted commas.
Allied and Complementary Medicine Database (AMED) [49]	produced with the help of the British Library's Health Care Information Service. specialized bibliographic database created with doctors, therapists, scientists, and historians in mind.	Scientific	Simple search by keyword or phrase. Search with multiple words can use inverted commas around the phrase. By default search using keywords, author, and subject if the "Select a Field" option is not selected. Combine search by OR AND options.

Table 4. *Cont.*

Name	Description	Content-Type	Search
MedLine [23]	More than 28 million journal articles in the current sciences are cited in the NLM's bibliographic data set, with a focus on biomedicine. It's an interesting feature that the NLM is listed with the data in MEDLINE.	Scientific	Advanced Search provides a guided mapping of keywords to Mesh terms. Narrow search by using subheadings. A clinical search query for EBM clinical Reviews.
ELSEVIER [24]	Specialized in scientific, technical, and medical content.	Scientific	Simple search by keyword, title, and subject area. Combine search by Boolean operators.
PhysioNet [27]	An exploration asset for complex physiological signs. It gives free admittance to huge assortments of physiological and clinical information as well as related open-source programming.	Clinical EHR	Simple keyword search. Narrow the search by selecting relevance and resource type.
PCORnet [28]	A public asset that gives a long-wanted sort of examination biological system: a completely coordinated network with tremendous, profoundly agent well-being information, research skill, and patient experiences working in and open.	Clinical	Simple search by keyword Narrow the search by selecting "category, resource type, network partners, and audience"
Surveillance, Epidemiology, and End Results (SEER) [50]	Division of Cancer Control and Population Sciences of the National Cancer Institute's Surveillance Research Program (SRP) and gives malignant growth insights with an end goal to decrease the disease trouble in the US populace (DCCPS)	Clinical	Simple keyword search. Combine search with Boolean operators. Keyword search for statistical information.
BioPortal [51]	The most important tool provided by NCBO (National Center for Biomedical Ontology) is a Web gateway and Internet-based tools that encourage biomedical specialists to access, audit, and coordinate unique ontological resources in all areas of clinical practice and biomedical examination.	Scientific	Simple search by class name, ontology name. Advanced search by Property values, Obsolete classes, Ontology views. Limit your search using classes definition or exact matches.

3.2. Semantic Enrichment Approaches

According to the literature, a lot of work is done on semantic analysis in biomedical text, which offers great potential for identifying semantic relationships between biomedical entities, terms, and terminologies [52,53]. Semantic analysis in the biomedical domain employs multiple NLP tasks, including WSD [54], clustering [55], ontology learning [56], information retrieval [33], text classification [57], question answering [39,41,42], text Summarization [58], topic detection [58], and many others. Extracting semantic similarity implies determining and quantifying the contextual relationship between concepts based on shared features. The calculation of semantic similarity between text components can

be carried out with data recovery from different biomedical sources [19,46–51], heterogeneous data integration [59] automatic grouping of semantically related clinical terms [60], and clinical model clustering from patient EHR [56] in the domain. To find the similarity between any two texts, traditional frequency-based techniques [61,62] can be employed. However, to avoid misinterpretations in the biomedical sector, it is vital to assure not just explicit relevance between two-word sequences, but as well as the underlying clinical setting. As a result, research into semantic similarity techniques is critical in this field. For instance, because they are both respiratory system ailments, "bronchitis" and "influenza" have semantic similarities.

EBM demands physicians and practitioners review relevant and up-to-date information from multiple biomedical sources on a regular basis. Due to the diverse information resources and contextual models, semantic resemblance for medical questions and a biomedical corpus has been a more prominent research question in recent decades [63–65]. Depending on the techniques and tools employed for finding relatedness in the clinical area, the literature discussed three ways of semantic similarity as follows in Figure 5: A. distributional-based approach or corpus-based approach; B. knowledge-based approach; C. deep learning-based approach.

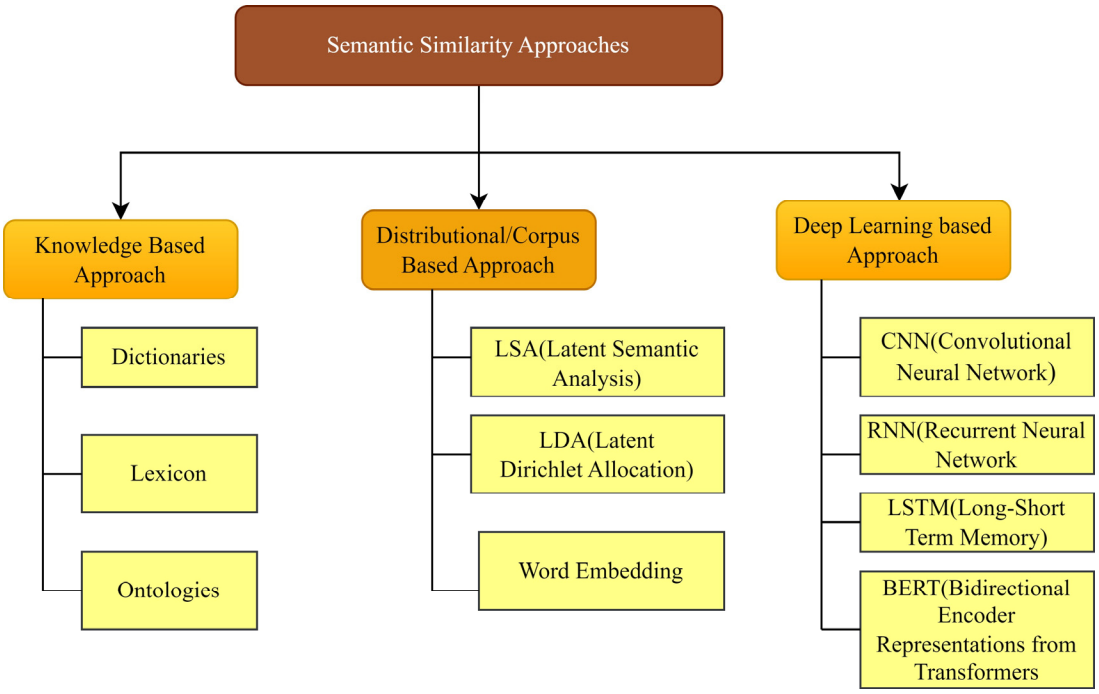

Figure 5. Semantic enrichment approaches.

A. Distributional or Corpus-based Approach

In a distributional-based strategy for computing semantic similarity in the biomedical domain, a domain corpus and a knowledge source are utilized. In the distributional-based approach, semantic similarity is measured by constructing contextual vectors with the notion that words are similar to each other in the same context [66,67]. The steps for finding context vectors are:

1. Initially, word vectors are created from the corpus using word co-occurrence.
2. concept descriptors are retrieved from an information source such as a word reference or thesaurus, and they can be extended to incorporate descriptor terms from related concepts [68].
3. Then in the last step term vectors matched to concept descriptors are aggregated to form context vectors.

Word embedding is a commonly used method for text expressions. Recent research has proven that context embedding generated by word embedding can accurately capture a precise meaning [67]. It is an unsupervised technique that uses related vectors to capture contexts and semantically related terms in a large collection of words [66]. The bag-of-words format is converted to a continuous vector space representation using the word2vec approach [66]. Extensions to the word2vec technique, such as sentence2vec and doc2vec, are developed by embedding sentences and documents at the sentence and document levels [69]. GloVe [32], as well as skip-gram [66], are two popular word-embedding models.

We have compared different word-embedding models that are often used on biomedical corpora in the table below. The parameters in Table 5 are as follows:

1. Pretrained: It indicates whether or not the model has been trained on similar tasks (Y/N). These models converge fast because their weights are already optimized and reduce time and effort.
2. OOV (Out of Vocabulary): OOV models are richer than non OOV which are the terms encountered in NLP that are not part of the usual lexicon (Y/N).
3. Prediction: when building, processing, and validating a model that can be used to predict future occurrences using known results, it is indicated whether it is a predictive model (Y/N).
4. Frequency: Based on how frequently certain terms appear in the text or document, it vectorizes the text(Y/N).
5. Morphological Information: It investigates and describes the structure of words and their relationships (Y/N).
6. Work level: It depicts the various levels at which models, such as embeddings, can be applied to individual words, phrases, paragraphs, or texts.
7. Evaluation: It explains the model's benefits and drawbacks.

Table 5. Comparison of different word-embedding models.

Model	Pretrained	Out of Vocabulary	Prediction	Frequency	Encode Morphological Information	Work Level	Evaluation
One hot coding [66]	-	-	-	-	-	words	Computationally expensive and sparse for a large corpus. Context independent.
Cooccurrence matrix	-	-	-	√	-	words	Faster but requires huge memory.
Word2Vec [66]	√	-	√	-	-	words	It Consumes less space. Good for semantic relation. CBOW and Skip-grams variants.
PV-DM [69]	√	-	√	-	-	Sentences, Paragraphs, and Documents	Softmax weights and word vectors call for extra memory
PV-DBOW [69]	√	-	√	-	-	Sentences, Paragraphs, and Documents	Simple and faster. Less memory is needed because it only stores the word vectors.
Glove [32]	√	-	-	√	-	words	Trained on the global co-occurrence matrix of all words combined. Denser and expressive vector representation.

Table 5. *Cont.*

Model	Pretrained	Out of Vocabulary	Prediction	Frequency	Encode Morphological Information	Work Level	Evaluation
FastText [70]	√	√	√	-	√	Characters N-grams and words	Incorporates sub-word information. Memory and computationally intensive needs rise as the corpus size does.
ELMo [70]	√	√	√	-	√	words	Context-dependent vector representations. Computationally intensive require more training time.

B. Knowledge-based Approach

To compute semantic similarity between biomedical terms, an augmented source of knowledge sources was used in the approach [52,53]. These leveraged sources of biomedical knowledge are listed below.

(1) Dictionaries
(2) Lexicons
(3) Ontologies

1. Dictionaries

In the biomedical field, word references are utilized for a wide scope of purposes, including technique, hereditary qualities, species, living beings, prescriptions, clinical codes, clinical gadgets, phrasing, and methodology. Stakeholders utilize these word references in their medical fields. At times, unique concept IDs are allocated to a similar idiom or terms. Because no single dictionary in the Unified Medical Language System (UMLS) [71] provides complete coverage of the clinical domain, it is difficult to rely on one at a time. The process of selecting dictionaries and integrating the relevant dictionary elements is a vital step in obtaining accurate information.

2. Lexicons

Lexicon is another rich source of biomedical knowledge that helps to construct the vocabulary of diverse biomedical words. The lexicon is made up of lexemes, which are lexical entries. Each entry is a word (lexical item or lexeme) that includes one or more spellings in a certain area of speech and provides the morphologic, orthographic, and syntactic features of a word [72]. The lexical metadata is depicted in Figure 6 below.

Researchers frequently employ domain-specific biomedical lexicons for natural language processing applications [73]. Various approaches for automatically creating a semantic lexicon from biomedical text [74] have been created, according to the literature, which ties words and phrases to certain semantic types. For the generation of the semantic lexicon, these methods employ existing knowledge sources in the biomedical domain, such as medical terminologies and ontologies. The UMLS is a comprehensive knowledge source in the medical domain that includes a semantic lexicon known as the SPECIALIST lexicon. Each lexeme is linked to one or more syntactic types, each of which can be linked to one or more semantic types. It contains Lexemes that are matched to strings in the NLM's UMLS Met thesaurus from 1997 [71].

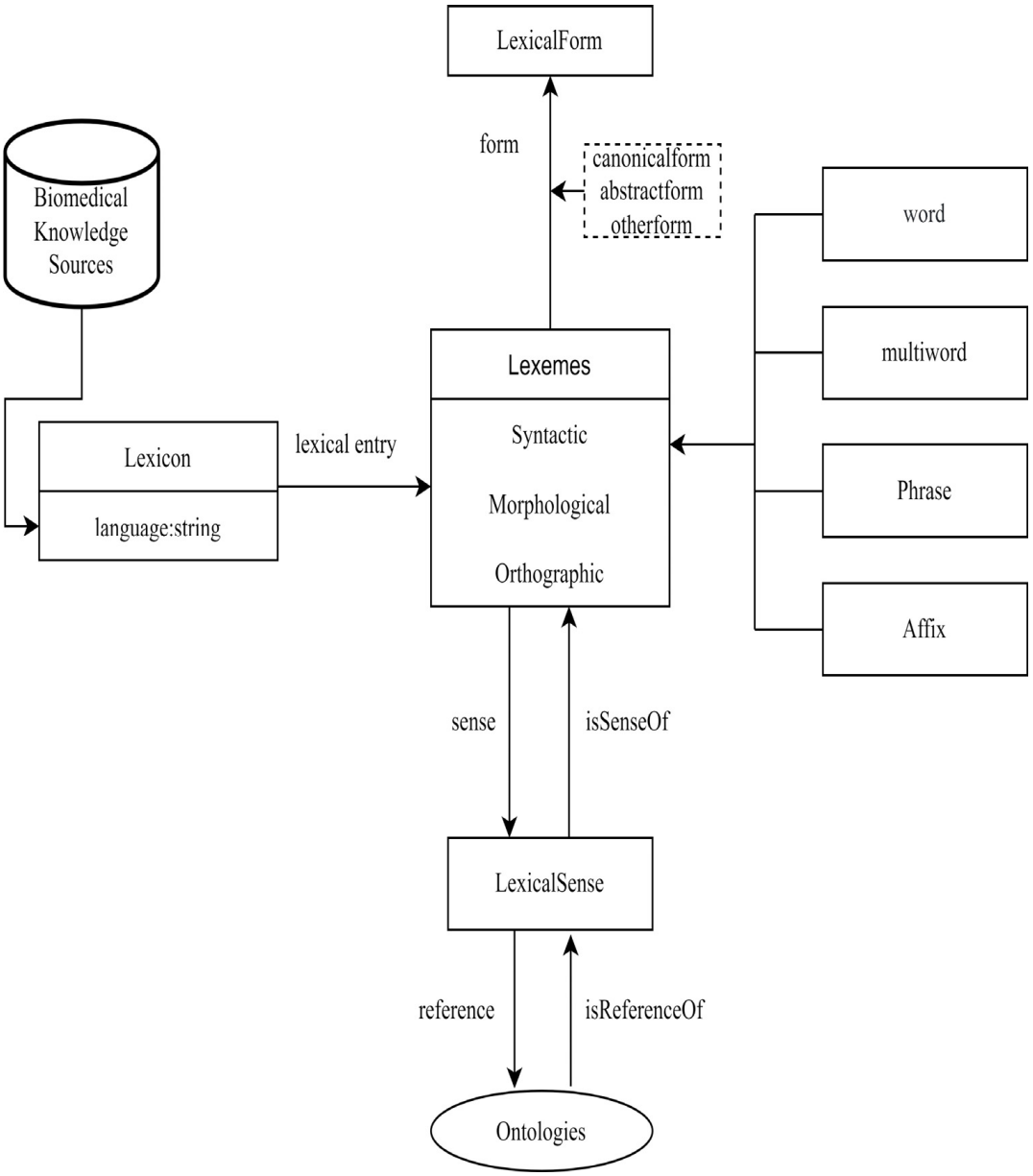

Figure 6. Metadata of lexicon.

3. Ontologies

Among knowledge base approaches substantial work in the clinical area has been devoted to approaches that use taxonomic structure to compute semantic similarity between biomedical texts. So, in the last decade, the ontology-based similarity was the most popular way which brings conceptual similarity using biomedical ontologies which have three basic elements such as (i) a group of concepts that is utilized to address items and relations; (ii) formal adages that compel the semantics and guarantee that those ideas are utilized accurately; and (iii) definitions [75–77]. There can be great variations in Ontologies construction, granularity, content, and different qualities. As displayed in Table 6 beneath, each studied philosophy is assessed utilizing a couple of terms.

Table 6. Significant ontologies in biomedical domain.

Name	Content	Structure	Classes	Maximum Depth
SNOMED-CT [78]	Clinical terms	Collection of medical terms created by the College of American Pathologists. Medical terms provide necessary codes, synonyms, terms, and descriptions required in clinical reports.	361,588	28
RCD [79]	Clinical Terms Version 3 (CTV3) (Read Codes).	Standard jargon for clinicians to record patient discoveries and strategies in well-being and social consideration.	140,065	17
National Drug File Reference Terminology (NDRT) [80]	pharmacy	a formal representation used to depict the components of medicine as well as its chemical makeup, dosage form, physiological effects, mode of action, pharmaceutics, and associated disorders.	36,202	11
International Classification of Diseases (ICD) [81]	Morbidity entities	Provides information about mortality and morbidity in population coded with ICD codes	12,445	4
Medical Subject Headings (MESH) [82]	Medical Subject Headings	The indexing of life sciences books and journal articles serves a purpose. MeSH headings such as anatomy, diseases, chemical drugs, etc.	347,692	15
MedDRA (Medical Dictionary for Regulatory Activities (MedDRA) [83]	International medical terminology	Use for information passage, recovery, investigation, and presentation are underlined. It applies to all phases of medication advancement, except for creature toxicology.	75,741	-

1. Content: It specifies the types of terms or entities it contains, such as clinical terminology, medications, and morbidity entities, among others.
2. Structure: It displays the type of relationship that exists between several terms.
3. Classes: These are groups of different ontology concepts.
4. Maximum Depth: It displays the hierarchy tree's maximum depth or tiers.
5. Citations: It includes citations to articles that use the relevant ontology.

The UMLS was created by the US NLM and is a bunch of records and programming that unites different well-being and biomedical vocabularies to permit interoperability with multiple healthcare systems. In more detail, UMLS comprises three information sources and a bunch of programming apparatuses for getting to them. Metathesaurus, semantic network, and SPECIALIST Lexicon are among the information sources used by UMLS. UMLS likewise houses a few biomedical ontologies and vocabularies. Notable models incorporate MESH and SNOMED-CT ontologies. Only a couple of models incorporate MESH, ICD, and SNOMED-CT.

C. Deep Learning based Approaches

The recent years have seen a rise in the use of DL techniques in the majority of biomedical text mining applications. Additionally, DL-based semantic similarity techniques have been used in conjunction with recent advancements in neural networks to improve performance in the biomedical sector. For paraphrase identification tasks, numerous scientists employed a variety of characteristics such as n-gram features [70,84], and syntactic and

linguistic properties [85]. Researchers' attention has shifted to the semantic representation of text as a result of the adoption of DL algorithms. Biomedical tasks have been successfully handled using DL techniques, including BQA [38,86] and BTS [87] in order to process a huge quantity of data in less time. To represent the textual features biomedical researchers can utilize recurrent neural networks (RNN), long short-term memory (LSTM), bidirectional long short-term memory (Bi-LSTM), and other neural network frameworks to identify the relationship between various elements in text sequences. DL models in light of transformer designs, such as bidirectional encoder representations from transformers (BERT) and robustly enhanced BERT approach (RoBERTa) are popularly employed. Cai et al. [63] proposed an unsupervised clustering technique used to mine the user intent taxonomy, and a CNN-LSTM-based model used to predict user intents. As a clinical setting mindful classifier, Muhammad Afzal [65] fostered a Biomed summarizer utilizing a bidirectional long-transient memory repetitive neural organization. Milad Moradi [87] proposed a biomedical message summarizer that evaluates the enlightening substance of sentences utilizing contextualized installation created by the BERT model, a profound learning model. BioBERT, SciBERT, ClinicalBERT, and different variations [31,78,84,88,89] of pre-prepared BERT are accessible in the biomedical area.

3.3. Text Similarity Metrics

For text-related research and applications, textual similarity measures play a pivotal role. It acts as a basis of performance for the NLP tasks such as information retrieval [33,34], text classification [57], document clustering [56], topic detection [58], question answering [39], text summarization [58], etc. Finding word similarity is the prior step of text similarity which can be utilized to compute sentence relatedness, paragraph, and test articles, etc. Word similarity can be in two ways:

1. Lexical Similarity: It can be measured as a similar character or word sequence using its intersection [90].
2. Semantic Similarity: If words are related by some type, or opposite to each other, mean the same thing or are used in a similar context then they are semantically similar [91].

In the literature, broad work has been done on approaches and measures for textual similarity in biomedical documents [91]. Jiapeng Wang and Yihong Dong [92] followed the advancement of semantic similarity and recognizing them based on strategies utilized in them. It is observed that test distance and text representation are two viewpoints for textual similarity methods. The text distance methods obtain similarity between two words by considering their distance. It is divided into three categories based on the three techniques of measuring the distance.

A. Length distance
B. Distribution distance
C. Semantic distance

Text similarity metrics are broadly classified as shown in Figure 7.

The first method utilizes the numerical characteristics of the test to compute the length of vector text. The second method is utilized to compare whether two text articles come from the same collection or not. It calculates the distance at a semantic level to check whether there are common terms in the text. Text representation speaks to content as a numeric highlight that can be calculated straightforwardly. String-based, corpus-based, single-semantic content, multi-semantic content, and graph-structure-based representations are the several types of this method [90]. A string-based technique uses character composition and string groupings to compare the similarity or differences between two strings. The corpus-based technique employs the data from the corpus to determine the degree of text similarity [90]. This data can be either a literary include or a co-occurrence likelihood. In a method based on the graph structure, the joins between vertices and the edges of the graph are used to calculate text similarity in order to more accurately anticipate the degree of similarity between vertices. Literary works show graph-based representation methods

popularly utilized for BTS [57,93–95]. It has been observed from the existing literature work [68,75,76,94] that there are various measures to compute textual similarity [92] but we are focusing on measures mostly used in the biomedical domain which are analyzed with few parameters as shown in Table 7.

Figure 7. Classification of text similarity metrics.

Table 7. Significant textual similarity metrics.

Paper	Name	Proximity by	Assessment by	Description	Limitation	Range
[96]	Cosine	Distance	Length	Distance computed by cosine angle between two vectors. Used for continuous and categorical variables.	The magnitude and direction of vectors are not considered. Does not work efficiently with nominal data.	0 to 1
[62,97]	Jaccard	Representation and numerical features	Phrase-based	It is calculated by dividing the size of the intersection by the size of the union of two sets. Used for continuous and categorical data.	Does not work efficiently with nominal data. Large datasets can have a big impact on the index.	0% to 100%
[98]	Word-movers	Distance	Semantics	A minimum distance of words in semantic space is computed using earth mover's distance method. Word vectors and linear programming.	High computational cost OOV words	-
[96]	Euclidean	Distance	Length	Euclidean space straight line separation between two points.	Not good with Higher dimensional data	-
[99]	JS (Jensen–Shannon) Divergence	Distance	Distribution	Measure the similarity between two probability distributions. Used with LDA (latent Dirichlet allocation).		[0, 1]
[100]	KL (Kullback–Leibler) Divergence	Distance	Distribution	A comparison of two well-known discrete probability distributions	Triangular inequality is not satisfied and is not symmetrical.	$(0, +\infty)$
[101]	LCS (longest common substring)	Representation and numerical features	Character based	Measures the similarity between two strings	Less accurate	0 to 1
[102]	Dice	Representation and numerical features	Phrase-based	Two sets of data are compared statistically by dividing the total number of elements in each set by the number of elements that are shared by both sets twice.	Does not satisfy the triangle inequality	0 to 1
[66]	Word2vec	Representation and numerical features	Corpus-based Shallow window based	Word vectors distributed numerical representations of word features	Incapacity to deal with unfamiliar or OOV terms. The definition of sub-linear relationships is implicit.	-1 to 1

Table 7. *Cont.*

Paper	Name	Proximity by	Assessment by	Description	Limitation	Range
[32]	Glove	Representation and numerical features	Corpus-based shallow window-based method	Trained on the co-occurrence matrix of words. Limits the use of the word vectors to refer to sub-linear connections in vector space.	Inability to handle unknown or OOV words. A lot of memory for storage.	−1 to 1
[103]	BERT (Bidirectional Encoder Representations from Transformer)	Representation and numerical features	Corpus-based shallow window-based methods	Encodes a huge amount of information into a set of dense vectors. Vectors that are more inline are more semantically alike, and vice-versa.	computationally intensive at inference time. lack of ability to handle long text sequences.	−1 to 1
[104]	LSA (Latent Semantic Analysis)	Representation and numerical features	Corpus-based Matrix Factorization	Extracts the hidden themes that the text or document is trying to convey. Singular value decomposition (SVD).	SVD, which requires a lot of computing, is frequently used. lacks the ability to appropriately handle polysemy (words with many meanings). Not fit well for all types of problems	−1 to 1
[105]	LDA(Latent Dirichlet Allocation)	Representation and numerical features	Corpus-based Matrix Factorization	Probabilistic topic modeling. Better disambiguation of words. More precise assignment of documents to topics.	Additionally, there must be unrelated themes (the number of topics is predetermined and must be known in advance).	
[106]	Bi-LSTM (Bidirectional Long-shortTerm Memory)	Representation and numerical features	Multisemantic document text matching	Have the sequence information in both directions. Usage of gates to regulate the flow of information.	Prone to overfittings. Expensive	-
[94]	Knowledge Graph(KG)	Representation and numerical features	Graph structure	Create a consistent low-dimensional vector space from the knowledge graph's elements and linkages (semantic portrayals that may effectively transmit semantic facts).	Coverage, correctness, and freshness of knowledge graphs	-

Table 7. *Cont.*

Paper	Name	Proximity by	Assessment by	Description	Limitation	Range
[68,107]	IC-based measure	Information Content	Knowledge based	Use the Information Content values to compute semantic similarity between them. Lowest Common Subsume (LCS) which is extracted from the "is a" hierarchy.	Two pairs with the same summation of IC(c1) and IC(c2) will have the same similarity	
[108]	Recall Oriented Understudy for Gisting Evaluation (ROUGE-N)	co-occurrence	determines the proportion of "n-grams" that match the model and reference texts.	ROUGE Recall ROUGE Precision ROUGE F1-Score	Cannot capture synonymous concepts and coverage of topics	0 to 1
[108]	Recall-Oriented Understudy for Gisting Evaluation (ROUGE-SU) Skip Unigram	Co-occurrence	With maximum skip distances of 1, 4, or 9	A candidate phrase is given credit even if it does not contain any word pairs that are co-occurring with its references.	Does not cater to different words that have the same meaning	0 to 1
[108]	ROUGE-L	Representation	String based	Measures longest matching sequence of words using LCS	Does not require consecutive matches	0 to 1

3.4. Comparative Study of Significant Biomedical Text Summarization

ATS is a method of reducing input text documents into a meaningful summary that shows the main gist of the document without altering information [8]. Its approaches are broadly categorized into extractive and abstractive types. The authors [8–10] conducted a detailed review of various techniques of ATS. Due to the always-expanding measure of logical and clinical writing, ATS is an interesting issue in the area of data recovery research, especially in the clinical as well as medical spaces, since it gives a compelling method for consolidating source archives while holding their most educational content [9]. BTS is especially helpful in clinical QA frameworks, where it is necessary to precisely recognize experimentally sound distributed investigations and summarizes specific studies for a particular query type (e.g., intervention and prognosis).

Generally, every ATS process goes through a subsequent step including data collection, text data pre-processing, feature extraction, summarization approach, and ranking techniques as shown in the above Figure 8.

1. Data Collection: Collection of text data from various relevant sources.
2. Text Data Preprocessing: linguistic techniques utilized to pre-process input text documents, including sentence segmentation, punctuation marks removal, filtering stopwords, stemming [29], etc.
3. Feature Extraction: The extraction and representation of sentences is vital for the entire summarization process by discovering topic sentences, essential data traits or attributes within the source document [29].
4. Sentence Preparation: Encode and representation of sentences into real-valued vectors for further summarization process.
5. Summarization Approach: It is the first and important step in text summarization for choosing the approach [8] to be used. A few strategies include picking the main words and lines from the messages, while others include paraphrasing sentences by condensing original contents.

6. Summary: To acquire a superior synopsis of the source record, different calculations, and methods [9] are utilized under different methodologies. It is a stage where sentences are positioned and the high level is picked for incorporation in the synopsis.

This section shows that various ATS systems have been developed in recent years. Table 8 shows a comparative study of these systems with consideration of important steps involved in text summarization mentioned above.

Figure 8. Text summarization process.

Table 8. Comparative analysis of biomedical text summarization systems.

Paper	Supervised/Unsupervised Approach	Model	Semantically Aware Feature Extraction	Classification/Clustering/ Ranking	Performance	Corpus
[109]	Supervised	Graph-based summarizer with named entity recognition(NER)	Maps the words in the linguistic index to the entities in the NER(Named Entity Recognition) index	Extended the LexRank graph-based algorithm with NER [99,110] Entity Rank with graph-based approach	ROUGE scores increased for unweighted, as well as the weighted, Entity Rank	Used Entrez Programming Data from PubMed scientific biomedical abstracts
[111]	Supervised	Itemset based summarizer	Extracted concepts [95] by excluding concepts that are very generic.	Ranking of sentence by adding the support value of the item sets that cover the sentence. Itemset mining using the Apriori algorithm [112,113]	ROUGE metrics	400 biomedical articles from BioMed Central's corpus
[114]	Supervised	Bayesian summarizer	all extracted concepts. use of concepts [95] by excluding generic semantic type frequency-based ranking of features. Helmholtz principle to compute meaningfulness [115] CF-IPF approach classification of features [112]	Naïve Bayes for classification	Bayesian summarizer approach	BioMed Central's corpus
[116]	Unsupervised	Graph-based biomedical text summarizer	Extracted concepts [95] by excluding concepts with aforementioned semantic types. Correlations among multiple concepts using frequent itemset.	Graph-based minimum spanning Tree clustering algorithm [117]	ROUGE scores	400 biomedical articles from BioMed Central's open-access corpus
[118]	Unsupervised	graph-based summarizer	Extracted concepts [95] by excluding concepts that are very generic. Itemset mining.	Graph-based approach-small world network [119]	ROUGE-2	Corpus contains 300 biomedical full-text articles from BioMed Central's corpus.
[13]	Supervised	Extractive query-based summarizer	Sentences and queries are vectorized using the tf-idf approach.	Regression and classification. Support Vector Machine	Classification performs better than regression	BioASQ data set
[12]	Supervised	LSTM Model	classifier to label topics in the history of present illness (HPI) notes	LSTM Model	Precision (P), Recall (R), F1 Score = 0.88	MIMIC-III

Table 8. *Cont.*

Paper	Supervised/Unsupervised Approach	Model	Semantically Aware Feature Extraction	Classification/Clustering/ Ranking	Performance	Corpus
[120]	Hybrid approach	Clustering and Item set mining-based summarizer (CIBS)	Itemset mining Apriori Algorithm [112,113]	Agglomerative hierarchical Clustering algorithm [121]	ROUGE scores For multi-document	Multi document corpus consists of 25 collections, each one containing 300 documents (Pubmed abstract) and a model summary. A single document corpus consists of 400 scientific biomedical articles from BioMed Central's corpus.
[6]	Supervised	Small world network based summarizer	Helmholtz principle [115] to calculate the meaningfulness of the concept	Graph based approach to a small-world network	Rouge	300 biomedical articles from BioMed Central's corpus.
[122]	Unsupervised	Clustering and itemset based summarizer	Concept frequency sentence frequency (CF-SF) Vector space model [123] extracted concepts [95] by excluding generic semantic type	itemset mining using the apriori algorithm [96]. K means clustering [124]	Rouge	100 biomedical full-text papers from the BioMed Central.
[87]	Unsupervised	pre-trained deep language model BERT	Pre-trained BERT on Wikipedia and BookCorpus	Agglomerative hierarchical clustering algorithm [121]	R1 = 0.7639 R2 = 0.3481	Articles from BioMed Central.
[125]	Supervised	MINTS (Multi Indicator Text Summarization Algorithm).	Feature matrix using 5 pointers of significance such as: length of the sentence position, term relevance rate standardized degree centrality, cross-over with global term frequency distribution determined using the Srensen Dicecoefficient/list (DS) as a comparability metric [126]	Apache Lucene [127] Random forests classifier [126] Aggregated ranking of indicators of relevance	ROUGE-1: 0.414 ROUGE-2:0.136 ROUGE-SU4:0.171	Articles from the Colorado Richly Annotated Full Text (CRAFT) corpus [128]. Indexed database of Medline abstracts
[58]	Supervised	Syntax based Negation and Semantic Concept Identification based summarizer	Concept recognition using cTAKES [129]	cTAKES clinical NER [129] using regular expression	Negation Detection Accuracy Concept identification	clinical narrative texts from MIMIC-III critical care database [20] contains 58,976 ICU patients.

Table 8. *Cont.*

Paper	Supervised/Unsupervised Approach	Model	Semantically Aware Feature Extraction	Classification/Clustering/ Ranking	Performance	Corpus
[7]	Unsupervised	Domain specific word embeddings and graph based summarizer	3 different versions of BioBERT pretrained on PubMed abstracts, PubMed Central (PMC) full-text articles, and a combination of both respectively	Page rank algorithm	ROUGE-1 ROUGE-2	Created corpus by retrieving 2000 articles from PubMed Central.
[130]	Unsupervised	Word Embedding Based BiomedicaText Summarizer	Word2vec Pretrained from PubMed, PMC, and recent English Wikipedia dump texts	Graph based Page Rank algorithm	ROUGE-1 ROUGE-2 ROUGE-3 ROUGE-SU4	Corpus of 200 biomedical papers from BioMed central full-text database
[65]	Supervised	Biomed Summarizer	Kera tokenizer [131] The prognosis quality recognition model (AdaBoost MLP) was trained on 5 features title, abstract, article type, publishing journal, and authors. Semantic enrichment using ontologies	Bi-LSTM PICO classifier with two more classes, Aim, and Results. Aggregate score of relevance study type venue credibility freshness	Accuracy of identification of quality articles:95.41% Accuracy of classification:93%	PubMed abstracts
[132]	Supervised	Word Embedding based Maximal Marginal Relevance [34]	Pre-trained word2vec and skip-gram tools from PubMed and PubMed Central (PMC). 5 features from the QSpec system [36]	Maximum Marginal Relevance [34]	F1-score	Clinical Inquiries section of The Journal of Family Practice
[133]	Unsupervised	MultiGBS	MetaMap [134], OGER [135], and SemRep [136] to extract 3 types of relationships semantic, word and co-reference	Multi-layer graph approach with MultiGBS sentence selection algorithm [137]	F-measure ROUGE-L	450 biomedical scientific articles from BioMed Central
[18]	Supervised	Attention based clinical note summarizer	Fine-tuned BERT model used for word embeddings	High Attention Score of sentences calculated by correlating tokens, segments, and positional embeddings	KLD = 0.795 JSD = 0.405	ICD-9 labeled MIMIC-III discharge notes
[84]	Supervised	SciBERT based Summarizer	Pretrained SciBERT [88]	Graph Attention Networks-based graph encoder to encode sentences and word co-occurrence graphs,	kappa/alpha informative = 0.669/0.671 coherence = 0.602/0.605 redundancy = 0.653/0.656 fluency = 0.689/0.692	COVID-19 open research corpus build from PMC, PubMed, WHO Database

3.5. Comparative Study of Popular Works on Biomedical Question Answering Systems

This examination pushes the limits of BTS with a very interesting application that utilizes summarization techniques to develop patient-centric systems such as QA Systems [39], Telemedicine [38], and others that are presently accessible to the overall population.

The enormous expansion in biomedical examination papers implies the quick development of biomedical research. Applying this huge data corpus to extract relevant information for a presented query/question is the QA model. This was popular in the new COVID-19 flare-up, where biomedical specialists and clinicians were attempting to present remote monitoring and medication and beat the odds to accumulate pertinent information to foster viable therapies or antibodies. BQA is the most popular application of text summarization nowadays after the global pandemic of COVID-19 [38,39,84]. As an arising Question Answering task, BQA empowers imaginative applications to actually see, access, and comprehend complex biomedical knowledge [21,25] and generate meaningful results to queries. The basic steps involved in BQA are as follows in Figure 9.

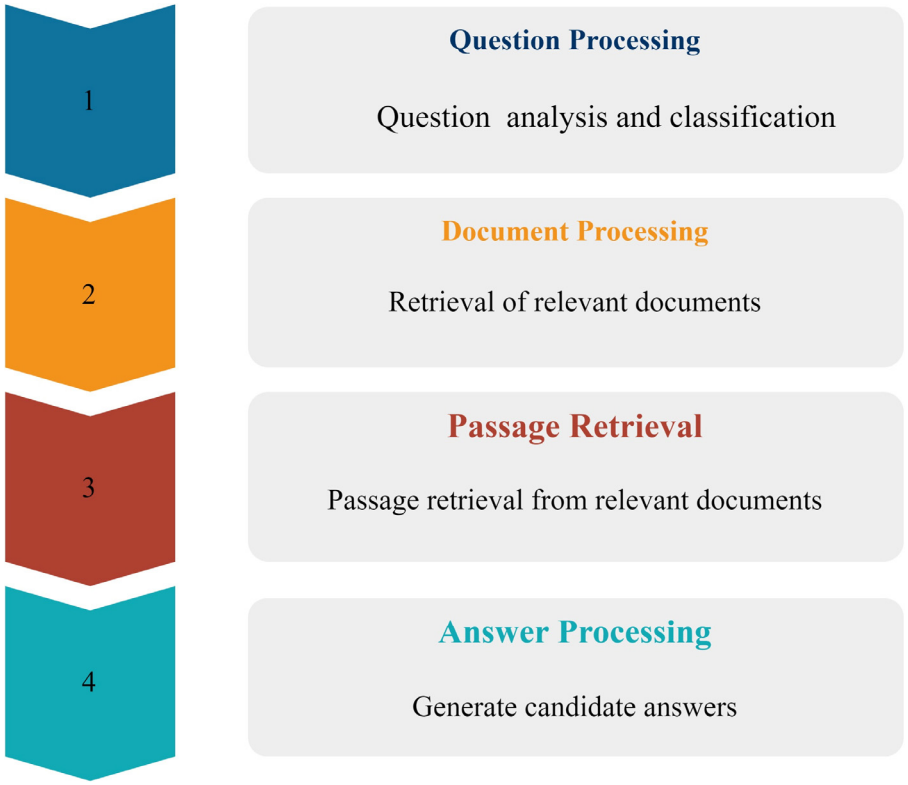

Figure 9. Biomedical question answering process.

1. Question Processing: Performs question analysis and classification to convert it into a search query.
2. Document Processing: Query terms are applied to retrieve a group of related documents.
3. Passage Retrieval: NLP techniques can be utilized to extract groups of passages.
4. Answer Processing: It uses different extraction techniques on the result of the document or passage processing module to present an answer.

In this section, we provide a critical review of recent efforts in BQA represented in Table 9.

Table 9. Comparative analysis of various biomedical question answering systems.

Paper	Dataset Content Type	Question Proc.	Document Proc.	Passage Retrieval	Answer Proc.	Answer Type	Databases
[138]	Clinical	Question analysis with MetaMap Transfer [134] (MMtx) and UMLS [71]. Question classification on the basis of weighted phrase annotation	Use of machine learning classifiers for document classification. Use of cosine similarity for searching the relevant documents	Passage retrieval using similarity vectors	Topic clustering, ranking, and hierarchical answer representation	Passage	1700 abstracts related to pancreatic cancer from PUBMED
[139]	Scientific	Question classification and query modification using NER and SRL [140].	WordNet [141] and Longman's [72] dictionary used with Google interfacing program	NER(Named Entity Recognition) and SRL, Ranking	Linear Answer Ranking Model	Passage	Google
[142]	Scientific	Deep syntactic representation of the questions using Government and Binding parser, FIPS [85]	Document retrieval through PubMed	Rank descriptor belonging to the target set	Rank descriptor belonging to the target set	Candidate answers	5000 MEDLINE abstracts
[143]	Clinical	-	probabilistic relevance model BM25 [144]	Longest Common Subsequence [101]	Topical clustering and ranking	Multiple sentences passages	MEDLINE abstracts, eMedicine documents, clinical guidelines full-text articles, and Wikipedia documents
[145]	Clinical	Question processing with cTAKES clinical text analysis system [129]	Document retrieval and ranking of full text using Lucene indexing [127]	The paragraph level baseline using document level score and paragraph level scores	Rule-based reranking and ML-based reranking	Paragraphs	Medpedia [146] and Cliniques corpus [147]
[148]	Clinical	PICO-based question templates	Lucene indexer [127] for relevant document retrival	top N-matched clinical evidence will be considered as the candidate answers	Probability based score	Paragraphs	Trip Answers website [149]
[150]	Clinical	Customization of question templates in [151]	Web search engines, Google and PubMed	Description Logic(DLF) and UMLS Semantic Network	DLF pattern matching in question and answer	Answer patterns are semantic triples in the form of subject-predicate-object	Google and PubMed

Table 9. *Cont.*

Paper	Dataset Content Type	Question Proc.	Document Proc.	Passage Retrieval	Answer Proc.	Answer Type	Databases
[152]	medical	MESA ontology-based extraction of medical entities, semantic relations, and additional information about the patient	RDF annotations of the source documents and SPARQL queries	RDF annotations with SPARQL queries	Three steps query relaxation, semantic search, and ranking	Factoid, definition	MEDLINE articles
[153]	clinical	Semantically annotating the questions with Mesh [154]	Retrieval of relevant documents databases and knowledge bases using PubMed curators [21]	Annotate passages with Ontological concepts	a ranked list of candidate answers	Factoid or collection of text snippets	Pubmed articles
[35]	clinical	Metamap [134] tool for query construction	PubMed search engine and UMLS similarity for question concept [21,71]. Document reranking using MetaMap [134].	Stanford CoreNLP [155] to retrieve relevant passage	BM25 [144] to rank passages	passage	Pubmed documents
[156]	Clinical Diabetics	Regular expression matching for question answer pair extraction	-	-	Latent semantic Indexing based on similarity calculation and answer ranking [104]	Candidate passage	Historical health data, LMD-FAQ Repository, web of the knowledge base
[157]	clinical/ Examination	Text sequences as an input to the SeaReader model	Apache Lucene [127] followed by BM25 ranking [144]	-	Attention score used to rank answers	passage	National Medical Licensing Examination
[158]	clinical	Annotating questions using Wordnet, SNOMED ontology	-	-	Question-answer template matching using Semantic Acquisition and text implication algorithms [159]	-	500 user questions collected from the medical field
[160]	scientific	Term-based interaction model	Document retrival using BM25 [144] and reranking using one of the model PACER [161], ABEL-DRMM [162]	BCNN [163] is used to score snippets	Relevance score of the document used to select the top K snippets as the answer passage	passage	Articles from MEDLINE /PubMed Baseline 2018 collection

Table 9. *Cont.*

Paper	Dataset Content Type	Question Proc.	Document Proc.	Passage Retrieval	Answer Proc.	Answer Type	Databases
[86]	scientific	question and answer sentence encoding using BiLSTM	SDM sequential dependence model based on the Markov random field model [164]	-	Semantic Matching Model	-	PubMed abstracts
[64]	Clinical	handcrafted lexico syntactic patterns and a machine learning algorithm for question classification of questions [165]	Pubmed search engine and UMLS [71] similarity	Stanford Core NLP and BM25 [144]	Different approaches for different types such as UMLS, and BM25 [71,144]	Yes or no, factoid, list, and summary	Medline Database
[166]	clinical	LSTM and DNN-based query selection to obtain keyword query	Iterative Elastic search	-	weighted Relaxed Word Mover's Distance [98] and Supervised answer candidate reranking using BERT	Passage as well as a factoid	Corpus of abstracts extracted from the PMC
[167]	scientific	Query formulation using NLTK	search engine to retrieve relevant documents	Generates semantic vectors of Question Snippets pairs.	probabilities of Q-A relations and ranking using RNN	snippets	Biomedical literature from PubMed/MedLine
[168]	Consumer health	Question processing with SVM, rule-based method, question frame extraction	More weightage to question focus and type in a query to get a relevant document	IR-based and entailment based answer retrieval using BM25 [144] and Feature based classifier respectively.	conventional team–draft interleaving to score answer sentences	paragraph	LiveQA-Med 2017 and Alexa MedlinePlus collections
[169]	Scientific	-	Elastic search (ES) used with the BM25 [144] to get relevant documents	Neural Ranking Model Deep Rank [170]	aggregation network for ranking	passage	PubMed Articles
[171]	Scientific	-	TF-IDF vectorizer and cosine similarity	Pre trained Distil BERT [89] model	The top 3 answers retrieved based on a weighted score between the retriever score and reader score	passage	CORD-19: Open Research Data set [172]

4. Discussion

4.1. RQ1. What Are the Various Biomedical Databases Available Online for Automated Biomedical Text Summarization?

Biomedical databases play an important role in the creation of knowledge sources and corpus that is used for BTS. Section 3.1 discusses the different universities and highly cited databases [21,22,47] that they host. The review of the literature shows that biomedical databases are classified into four types [41]: scientific, clinical, examination, and consumer. We have studied popular scientific databases such as PubMed [21], PMC [22], clinical databases SEER [50], Physio Net [27], examination databases [19,157], and consumer databases such as MEDLINE Plus [23]. Most of the studies in BTS made use of scientific databases such as the PubMed, Bioportal, Medline, etc. Some of the studies [4,11,12,18,56,173] focus on EHR-based databases such as Physionet [27]. Some of the studies [12,18,58] focus on clinical trials such as The Cochrane Library [25]. Few of the studies [8,65,171] focus on consumer-related content in databases such as those presented in [21–23]; here the content is related to statistics in public health. Moreover, there have been few works [157] in text summarization that made use of examination databases where a corpus was built using medical certification exams, where clinicians are evaluated based on their professional knowledge and ability to make a diagnosis.

Most of the works in BTS that are taken in our study use scientific databases as their knowledge corpus. Studies of the recent literary work show that there has been a good rise in the number of text summarizations on EHR databases. At the outset of the COVID-19 pandemic, we have also seen a rise in the summarization using consumer databases such as public health statistics. However, our study shows that significantly lesser work has been done in BTS using only examination knowledge sets. We found that most of the work has been carried out with the use of the standard datasets, however, most of them used only a single dataset which narrows down the scope of the summarization. There is a scope to broaden the scope of text summarization by using databases from multiple sources to build a knowledge base that can be used to train biomedical text summarizers and classifiers. This will help in diversifying the focus of the knowledge base in the biomedical domain. The creation of a knowledge corpus will serve as a test bed for several QA systems, Medical chatbots, and foster research.

4.2. RQ2. What Are the Different Semantic Enrichment Approaches Used in Biomedical Text Summarization and Their Comparative Evaluation?

In Section 3.2 we have discussed the role of semantic enrichment approaches used in BTS. Word Sense Disambiguation (WSD) [54] poses a huge challenge to BTS hence employing a distributional/corpus-based approach for text representation as textual vectors have been seen in the literature. Vector models such as word2vec [66], fast text [70], and pre-trained vectors such as a glove [32] have been popularly used in the literature work. We have also studied the literary works that established the significance of word vector representation for biomedical disambiguation.

Semantic enrichment for biomedical text is becoming significant because of the huge corpus of biomedical vocabulary. The use of dictionaries [71,72], has been found in highly cited literature work such as [35,64,139]. Application of ontologies [51,76,77] has been cited in the literature work such as [72,75,78,158] employing multiple ontologies for creating a better knowledge corpus. UMLS is discussed in our literary analysis because of its huge application in the BTS literature.

With the advent of DL techniques and emerging progress in the accuracy of deep neural networks, these approaches are used for computing semantic similarity in biomedical text. The ever-increasing size of biomedical text, heterogeneity of textual sources, and lack of syntactic interoperable methods makes DL methods the best fit for this area. The most popular DL approaches [63,89,103] have been found in the recent literature works [31,37,63,65,78,84,86,171]. Use of transformers for finding sequential semantics is rising.

A study of the literature works shows that distributional-based approaches are predominantly used for context vector representation of biomedical text.

The knowledge-based approach is used for semantically enriching biomedical knowledge corpus using sources such as biomedical ontologies [76–83], dictionaries [71], and lexicons [73]. DL-based approaches are used for extracting and classifying the context of the biomedical text in the BTS process. Each of these approaches plays a significant role in BTS. Hence the recent literature work on BTS [65,84,87,116,118–120] employs multiple approaches together for better performance. Although application of deep learning semantic enriching techniques results in average accuracy above 80% with most of the works, application of knowledge-based semantic enrichment approach, promises higher accuracies and manual validation. The challenges in employing single approach suggest the use of multiple semantic enrichment approaches in biomedical text summarization for more accurate summarization.

4.3. RQ3. What Are the Different Text Similarity Metrics Used in Biomedical Text Summarization?

Text summarization techniques focus on finding the similarity between two text units, sentences, paragraphs, or sections. Hence the study of text similarity metrics becomes essential. Study of two popular types of similarity metrics such as lexical and semantic similarity focus on syntax and meaning-based similarity respectively. Although text distance-based metrics are used for comparing the length, sequences, and distribution, in the biomedical domain, we see the emphasis on semantic-based similarity measures. The distance metrics include length-based techniques such as cosine, Euclidean, and word movers. The distribution-based metrics such as KL and JS divergence are also seen to be prominently used in biomedical similarity measurements.

With the increase in the vocabulary size in the biomedical domain, various representation and numerical feature-based similarity metrics such as word2vec are prominently employed in literary works [7,65,67,130].

The pre-trained vectors such as Glove [32] improve the accuracy of the similarity measurement using sub-linear relationships.

BERT employed in [7,18,31,84,85] is chosen in the summarization in biomedical clinical notes, EHR, and texts with long sequences. Summarization techniques also employ corpus-based topic modeling techniques such as LDA, and LSA where an unsupervised approach to creating topics with related text and exhibiting better coherence. BTS works [7,109,111,116,118,125,130] show the use of the ROUGE metric to find the co-occurrence similarity between the text using n-grams grouping for its simplicity in capturing synonymous concepts. The study of the evolution of the vectorization techniques from feature-based to co-occurrence based approaches and their comparative study will support future works in the BTS.

4.4. RQ4. What Are the Different Approaches Used for the Automatic Summarization of Biomedical Text and Their Comparative Analysis?

BTS works have been taken from 2020 to 2022. The techniques used for feature extraction, classification, and ranking are evolved over the period of literature. Most of the works for BTS are using a supervised approach which implied that they have used a labeled dataset such as corpus names [20,128] papers [12,18,58,125].

Most of the BTS techniques [57,93,94] have adopted graph-based approaches [6,7,116,118,133,174–176] to relate semantically similar topics. Most of the approaches [7,174] also used ranking algorithms such as page rank [176] to select the most relevant sentence for a summary generation.

With the evolution of DL methods BTS works [7,31,84,177] are employing transformers such as BioBERT [31] for establishing correlated topics. Word embedding model word2vec [66], and pretrained vector models such as a glove [32] are used to semantically enrich the biomedical text.

Most of the works [65,109,120,125] focus on generating a corpus using only abstracts in place of the entire document. With evolution, semantic enrichment approaches [52,53,56,58] have also been increasingly used for better summary generation.

We have seen the application of semantic metrics such as ROUGE and its variant being increasingly used in BTS as ROUGE is recall oriented and it can measure human-generated summaries with machine-generated summaries in a better way.

Literary work [20,45,157,158] shows various types of corpora used such as scientific, examination, consumer, and clinical.

Some of the works used standard datasets that are available in the public domain whereas many of the works also curate their own datasets using standard databases such as PubMed [7,21,118,120] using APIs.

4.5. RQ5. What Are the Different Approaches Used for Automatic Biomedical Question Answering Systems and Their Comparative Analysis?

Section 3.4 on BTS converges to an interesting recent application BQA system. The presented discussions outline the overview of the BQA process and some very interesting works that have been published and cited during the global pandemic period of COVID-19. The presented literature is an exhaustive list of recent as well as significant literary works. The papers are cited on the basis of the content types of the data sets that are used for making the corpus. The QA systems rely on the techniques used for question processing where we find the question classification carried out with evolving techniques such as handcrafted feature extraction, annotations using word vector models, and forming question templates. The use of DL sequential models such as BiLSTM, RNN, and CNN is seen increasingly in the systems studied. The document processing step focuses on the retrieval of the relevant documents from the knowledge corpus. Various text similarity-finding approaches are employed here. Statistical approaches such as probabilistic relevance [35], ranking methods [7,34,139,167,170,174], SPARQL Protocol, and RDF Query Language [152], and even some papers use the ready-made PUBMED, and UMLS filters are used. Passage retrieval aims to find the exact content of relevance to the question that is under consideration. Here similarity vectors and ranking techniques are predominantly used. LCS sequences [101,143] or the use of tools such as CoreNLP [35,155] is also done. The step in Answer Processing shows predominantly employs ranking techniques such as BM25 [144]. Various similarity metrics such as cosine similarity and word movers' distance [98,166] are seen employed to find the similarity between the extracted summary. The study of the section on BQA suggests the following research gaps that need to be addressed in our system.

1. Lack of access to a biomedical text corpus for summarizing data and its application to evidence-based medicine.
2. Lack of application of semantic enrichment approach for better context-based BTS.
3. Lack of proper heuristics for relevant document screening of biomedical text.

5. Proposed System

The study of the literature unveils the potential of text summarization on biomedical corpus for the design and development of BQA.

The proposed architecture mainly focuses on extractive summarization for domain-specific QA models. We will follow the extractive summarization approach to obtain significant words, phrases, or sentences from available biomedical literature and patients' EHR, clinical trials data, and patients' clinical summary notes which can be specific to syndrome or disease to generate an extractive summary. The study of research gaps in the discussion section shows the scope for the design of automated BQA with unique features as follows:

- Heuristics for sentence extraction
- Document Screening
- Context-Aware Semantic Enrichment

As shown in Figure 10 the outline of the architecture of the proposed system is elaborated in the following steps.

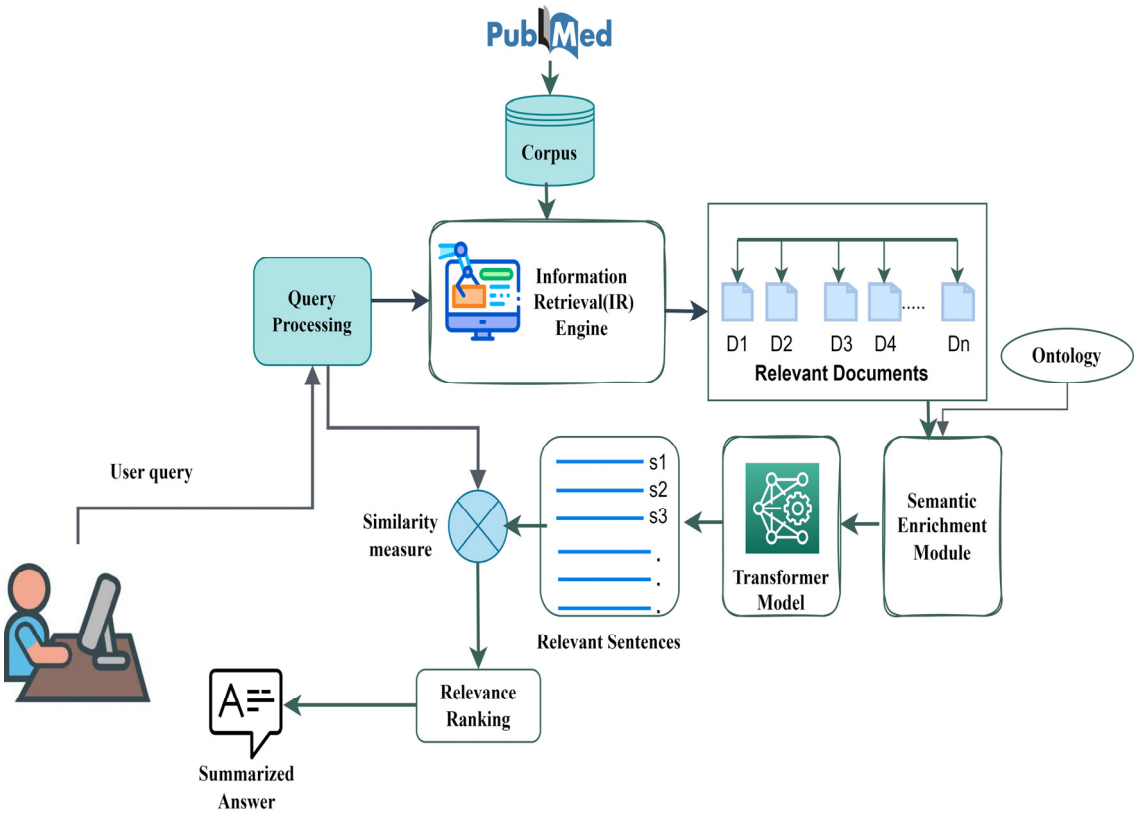

Figure 10. Proposed system architecture.

Step 1: Data Collection

The Data corpus is the knowledge base that serves as a test bed for the question answering system. This module aims to collect data in two ways: primary data collection is through patients' EHR, clinical trials data, and patient's clinical summary notes which can be specific to a syndrome. The secondary data collection can be done from a biomedical knowledge corpus that is publicly available such as PUBMED [21], Medline [23], UMLS [71], etc.

Step 2: Document Screening and Heuristics

To obtain the most relevant documents from various databases such as Pubmed, and Medline we will apply document screening using advanced search builders with various keywords such as Title/Abstract, or subject headings such as Mesh, Boolean operators to enhance search, etc. While collecting data from clinical summary notes we will be focusing on primary data such as 1. Diagnosis 2. Tests and 3. Patients' discharge condition. We suggest following heuristics for content extraction.

Heuristics for content extraction: Too long an input sequence and irrelevant content not only consume extremely high computational power of the computer but also prolong the inference phase. So, to limit the length of input text and extract relevant content we have here proposed two heuristics for content extraction as follows.

Heuristic 1. *As literature works [65,109,120,125] show that salient features and major findings of the complete article are found in the abstract so we are focusing on the abstracts of scientific articles.*

Heuristic 2. *As literature works [7,130,133] show that salient features and major findings of a complete article are found in the conclusion section, we are focusing on the conclusion section of scientific articles.*

Step 3: Data Preprocessing

Once data collection is carried out, then basic text pre-processing will be performed with the following steps to clean, prepare, and transform text data into the required format for further analysis.

- Initially text will be converted to lowercase and split into separate words.
- Stemming to convert word to its original form.
- Lemmatization to convert a word to its meaningful base form.
- Removal of stop words using NLTK library.
- Normalization to convert text into standard form.

Step 4: Semantic Enrichment

Once a user query is given, we find a semantically enriched summary relevant to it [31,63,65,78,84,88]. Literature works show increasing employment of deep learning methods, transformers such as BERT, and its variants such as DistillBert and SciBERT to semantically enrich input text. In our work, semantic enrichment is proposed to be done with the hybridization of traditional deep learning-based transformers and augmenting popularly used biomedical dictionaries, vocabularies, and ontologies such as SNOMED-CT [78], ICD [83], MESH [82], etc., to the database, which will improve the context of generated summary.

Step 5: Topic Modelling

The presented corpus is clustered based on similarity measures so that cohesive sentences are clustered in a single topic. The use of deep learning techniques such as LDA, RNN, and transformers is predominant for the modeling.

Step 6: Summary Generation

This step aims to generate a summary from top-n ranked text sequences using semantic similarity measures specific to a given user query and this is an aggregation of relevant sentences.

The complete flow of work is shown in Figure 10. Our primary work is to build a corpus for which relevant articles are extracted from popular databases such as Pubmed [21] and BioMed Central [47]. The user query will be processed by an Information retrieval engine to get multiple relevant documents from the corpus. The use of text similarity metrics can be an exercise to find relevant documents. Then documents text will be semantically enriched using biomedical ontologies and deep learning transformer-based techniques. After that, the semantically enriched text will be given as input to the transformer model to get relevant sentences. Finally, top-n-ranked text sequences will be included in the final text summary using semantic similarity measures.

6. Limitations

Our review is limited by the shortcomings of the selected literature. The biomedical domain is highly diverse and the use of information retrieval and natural language processing techniques for BTS and BQA systems is an emerging area. Hence, many of the selected articles and works were limited to prototype-based validations. The research survey was limited to 81 works from 2007 to 2021. This may lead to the risk of bias in the overall work. We have employed popularity ranking, citations, and manual screening for the chosen articles on the evolution of biomedical text summarization, semantic enrichment techniques, automated BQA systems, and state-of-the-art literary works done in biomedical text summarization. Review on biomedical text summarization has been limited to the stated three areas, i.e., semantic enrichment techniques, biomedical databases, and similarity measures. The automated biomedical question and answering system is the chosen application area of the BTS. We have chosen this area because of its increased popularity post-pandemic times. The work on biomedical databases and knowledge corpus is limited to chosen standard datasets, ontologies, and knowledge standards. The works on semantic enrichment approaches are limited to the popular ones used for biomedical text only. Although the similarity metrics are outlined for text similarity of ant nature, we have elaborated on the ones that are more relevant to biomedical text only. The works on the

BTS are studied and their comparative analysis is presented based on the stated criteria. These limited criteria may introduce a bias in the study. The works on the BQA are studied and their comparative analysis is presented based on the stated and limited criteria. This work also presents an architecture of MEDIQA: a framework of a BTS-based heuristic QA system for effective and precise answering to the patient's queries of a selected domain.

7. Conclusions and Future Work

The presented review provides a summary of the landmark works in the area of biomedical text summarization with emphasizes on the important and relevant focus areas in the field such as biomedical databases, feature extraction techniques, semantic enrichment approaches, and semantic metrics. To the best of our knowledge, the work is one of its kind to present the advancements in techniques in all the major artifacts of text summarization.

The study of biomedical datasets revealed that most of the work has been carried out with the use of standard datasets which narrows down the scope of the summarization. There is a scope to diversify the focus of the knowledge base in the biomedical domain with the use of databases from multiple domains. The creation of a knowledge corpus will serve as a test bed for several QA systems, medical chatbots, and foster research.

The study of different semantic enrichment approaches was carried out. Although there are significant works in dictionary-based and knowledge-based semantic enrichment approaches, deep learning approaches and their applications in the domain are found increasing.

A study of two popular types of similarity metrics such as lexical and semantic similarity which focus on syntax and meaning-based similarity was carried out. The majority of the studied works on biomedical text summarization show the use of the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) metric to find the co-occurrence similarity between the text using n-grams grouping for its simplicity in capturing synonymous concepts. The work also helps uncover the major research challenges in the area of biomedical text summarization. The review also puts forth the role of deep learning approaches in biomedical text summarization and the increasing relevance of the semantic enrichment approaches in the heterogeneous biomedical data in the literature.

The study of the question–answering systems unveiled research gaps such as lack of access to a biomedical text corpus, lack of application of semantic enrichment approach, and lack of proper heuristics for relevant document screening of biomedical text. This has resulted in proposing a summarization-based question-answering system in the medical domain. Although empirical results could validate the results and accuracy of the framework, its implementation and experimentation are left to future work.

Author Contributions: Conceptualization, D.P., C.K.A. and W.H.L.; methodology, W.H.L., A.S. and S.P.; validation, G.K.S., A.S., S.P. and A.S.; formal analysis, W.H.L., S.P. and C.K.A.; investigation, D.P. and S.P.; resources, A.S. and W.H.L.; data curation, S.P. and D.P.; writing—original draft preparation, D.P., S.P. and C.K.A.; writing—review and editing, A.S. and W.H.L.; visualization, D.P.; supervision, W.H.L., A.S. and S.P.; project administration, S.P. and C.K.A.; funding acquisition, A.S. and W.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Mishra, R.; Bian, J.; Fiszman, M.; Weir, C.R.; Jonnalagadda, S.; Mostafa, J.; Del Fiol, G. Text summarization in the biomedical domain: A systematic review of recent research. *J. Biomed. Inform.* **2014**, *52*, 457–467. [CrossRef] [PubMed]
2. Afantenos, S.; Karkaletsis, V.; Stamatopoulos, P. Summarization from medical documents: A survey. *Artif. Intell. Med.* **2005**, *33*, 157–177. [CrossRef]
3. Moradi, M.; Ghadiri, N. Text Summarization in the Biomedical Domain. *arXiv* **2019**, arXiv:1908.02285. [CrossRef]
4. Wang, M.; Wang, M.; Yu, F.; Yang, Y.; Walker, J.; Mostafa, J. A systematic review of automatic text summarization for biomedical literature and EHRs. *J. Am. Med. Inform. Assoc.* **2021**, *28*, 2287–2297. [CrossRef]

5. Chaves, A.; Kesiku, C.; Garcia-Zapirain, B. Automatic Text Summarization of Biomedical Text Data: A Systematic Review. *Information* **2022**, *13*, 393. [CrossRef]
6. Moradi, M. Small-world networks for summarization of biomedical articles. *arXiv* **2019**, arXiv:1903.02861.
7. Moradi, M.; Dashti, M.; Samwald, M. Summarization of biomedical articles using domain-specific word embeddings and graph ranking. *J. Biomed. Inform.* **2020**, *107*, 103452. [CrossRef]
8. Mridha, M.F.; Lima, A.A.; Nur, K.; Das, S.C.; Hasan, M.; Kabir, M.M. A Survey of Automatic Text Summarization: Progress, Process and Challenges. *IEEE Access* **2021**, *9*, 156043–156070. [CrossRef]
9. Awasthi, I.; Gupta, K. Natural Language Processing (NLP) based Text Summarization—A Survey. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; ISBN 978-1-7281-8501-9.
10. Manish, S.; Disha, M. Techniques and Research in Text Summarization—A Survey. In Proceedings of the 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 4–5 March 2021.
11. Gulden, C.; Kirchner, M.; Schüttler, C.; Hinderer, M.; Kampf, M.; Prokosch, H.-U.; Toddenroth, D. Extractive summarization of clinical trial descriptions. *Int. J. Med. Inform.* **2019**, *129*, 114–121. [CrossRef] [PubMed]
12. Alsentzer, E. Extractive Summarization of EHR Discharge Notes. *arXiv* **2018**, arXiv:1810.12085v1.
13. Kaur, M.; Mollá, D. Supervised Machine Learning for Extractive Query Based Summarisation of Biomedical Data. In Proceedings of the 9th International Workshop on Health Text Mining and Information Analysis (LOUHI 2018), Brussels, Belgium, 31 October 2018; pp. 29–37.
14. Fiszman, M.; Rindflesch, T.C.; Kilicoglu, H. Summarizing drug information in Medline citations. *AMIA Annu. Symp. Proc.* **2006**, *2006*, 254–258.
15. Sackett, D.L. Evidence-based medicine. In *Seminars in Perinatology*; Elsevier: Amsterdam, The Netherlands, 1997; Volume 21, pp. 3–5.
16. Mollá, D.; Santiago-Martínez, M.E.; Sarker, A.; Paris, C. A corpus for research in text processing for evidence-based medicine. In *Language Resources and Evaluation*; Springer Science & Business Media: Dordrecht, The Netherlands, 2015. [CrossRef]
17. Hassanzadeh, H.; Groza, T.; Hunter, J. Identifying scientific artefacts in biomedical literature: The evidence-based medicine use case. *J. Biomed. Inform.* **2014**, *49*, 159–170. [CrossRef]
18. Kanwal, N.; Rizzo, G. Attention-based Clinical Note Summarization. *arXiv* **2021**, arXiv:2104.08942v2.
19. Masic, I. Review of Most Important Biomedical Databases for Searching of Biomedical Scientific Literature. *Donald Sch. J. Ultrasound Obstet. Gynecol.* **2012**, *6*, 343–361. [CrossRef]
20. Johnson, A.E.W.; Pollard, T.J.; Shen, L.; Lehman, L.-W.H.; Feng, M.; Ghassemi, M.; Moody, B.; Szolovits, P.; Celi, L.A.; Mark, R.G. MIMIC-III, a freely accessible critical care database. *Sci. Data* **2016**, *3*, 160035. [CrossRef]
21. Available online: https://pubmed.ncbi.nlm.nih.gov/ (accessed on 26 December 2022).
22. Available online: https://www.ncbi.nlm.nih.gov/pmc/about/intro/ (accessed on 26 December 2022).
23. Available online: https://www.nlm.nih.gov/medline/medline_overview.html (accessed on 26 December 2022).
24. Available online: https://www.elsevier.com/en-in/about (accessed on 26 December 2022).
25. Available online: https://www.cochranelibrary.com/about/about-cochrane-library (accessed on 26 December 2022).
26. Available online: https://www.ebsco.com/products/research-databases/cinahl-database (accessed on 26 December 2022).
27. Available online: https://physionet.org/about/ (accessed on 26 December 2022).
28. Available online: https://pcornet.org/about/ (accessed on 26 December 2022).
29. Feldman, R.; Sanger, J. *The Text Mining Handbook. Advanced Approaches in Analysing Unstructured Data*; Cambridge University Press: New York, NY, USA, 2007; pp. 13–19.
30. Singh, A.; Sharma, A.; Rajput, S.; Bose, A.; Hu, X. An investigation on hybrid particle swarm optimization algorithms for parameter optimization of PV cells. *Electronics* **2022**, *11*, 909. [CrossRef]
31. Lee, J.; Yoon, W.; Kim, S.; Kim, D.; Kim, S.; So, C.H.; Kang, J. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics* **2019**, *36*, 1234–1240. [CrossRef]
32. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global vectors for word representation. In Proceedings of the EMNLP 2014—2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
33. Hliaoutakis, A.; Varelas, G.; Voutsakis, E.; Petrakis, E.; Milios, E. Information retrieval by semantic similarity. *Int. J. Seman. Web Inf. Syst.* **2006**, *2*, 55–73. [CrossRef]
34. Carbonell, J.; Goldstein, J. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In Proceedings of the 21st Annual International ACMSIGIR Conference on Research and Development in Information Retrieval—SIGIR '98, Melbourne, Australia, 24–28 August 1998; pp. 335–336.
35. Sarrouti, M.; El Alaoui, S.O. A passage retrieval method based on probabilistic information retrieval model and UMLS concepts in biomedical question answering. *J. Biomed. Inform.* **2017**, *68*, 96–103. [CrossRef] [PubMed]
36. Sarker, A.; Mollá, D.; Paris, C. Query-oriented evidence extraction to support evidence-based medicine practice. *J. Biomed. Inform.* **2016**, *59*, 169–184. [CrossRef] [PubMed]
37. Jin, D.; Szolovits, P. PICO Element Detection in Medical Text via Deep Neural Networks. In Proceedings of the BioNLP 2018 Workshop, Melbourne, Australia, 24 July 2018.

38. Mutabazi, E.; Ni, J.; Tang, G.; Cao, W. A Review on Medical Textual Question Answering Systems Based on Deep Learning Approaches. *Appl. Sci.* **2021**, *11*, 5456. [CrossRef]
39. Jin, Q.; Yuan, Z.; Xiong, G.; Yu, Q.; Ying, H.; Tan, C.; Chen, M.; Huang, S.; Liu, X.; Yu, S. Biomedical Question Answering: A Survey of Approaches and Challenges. *ACM Comput. Surv.* **2022**, *55*, 1–36. [CrossRef]
40. Kaddari, Z.; Mellah, Y. Biomedical Question Answering: A Survey of Methods and Datasets. In Proceedings of the 2020 Fourth International Conference On Intelligent Computing in Data Sciences (ICDS), Fez, Morocco, 21–23 October 2020. [CrossRef]
41. Jin, Q.; Yuan, Z.; Xiong, G.; Yu, Q.; Tan, C.; Chen, M.; Huang, S.; Liu, X.; Yu, S. Biomedical Question Answering: A Comprehensive Review. *arXiv* **2021**, arXiv:2102.05281.
42. Soares, M.A.C.; Parreiras, F.S. A literature review on question answering techniques, paradigms and systems. *J. King Saud Univ. Comput. Inf. Sci.* **2020**, *32*, 635–646.
43. Kitchenham, B. Guidelines for performing Systematic Literature Reviews in software engineering. *Engineering* **2007**, *45*, 1051.
44. Masic, I. How to Search, Write, Prepare and Publish the Scientific Papers in the Biomedical Journals. *Acta Inform. Med.* **2011**, *19*, 68–79. [CrossRef] [PubMed]
45. Jin, D.; Pan, E.; Oufattole, N.; Weng, W.-H.; Fang, H.; Szolovits, P. What disease does this patient have? A large-scale open domain question answering dataset from medical exams. *Appl. Sci.* **2020**, *11*, 6421. [CrossRef]
46. Available online: https://www.tripdatabase.com/ (accessed on 26 December 2022).
47. Available online: https://www.biomedcentral.com/about (accessed on 26 December 2022).
48. Available online: https://www.embase.com/landing?status=grey (accessed on 26 December 2022).
49. Available online: https://www.ebsco.com/products/research-databases/allied-and-complementary-medicine-database-amed (accessed on 26 December 2022).
50. Available online: https://seer.cancer.gov/ (accessed on 26 December 2022).
51. Available online: https://bioportal.bioontology.org/ (accessed on 26 December 2022).
52. Alam, F.; Afzal, M.; Malik, K.M. Comparative Analysis of Semantic Similarity Techniques for Medical Text. In Proceedings of the 2020 International Conference on Information Networking (ICOIN), Barcelona, Spain, 7–10 January 2020.
53. McInnes, B.T.; Pedersen, T. Evaluating measures of semantic similarity and relatedness to disambiguate terms in biomedical text. *J. Biomed. Inform.* **2013**, *46*, 1116–1124. [CrossRef]
54. Patwardhan, S.; Banerjee, S.; Pedersen, T. Using measures of semantic relatedness for word sense disambiguation. In *The 4th International Conference on Computational Linguistics and Intelligent Text Processing*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 241–257.
55. Sanchez, D. *Domain Ontology Learning from the Web: An Unsupervised, Automatic and Domain Independent Approach*; Akademiker: Catalonia, Spain, 2012.
56. Gøeg, K.R.; Cornet, R.; Andersen, S.K. Clustering clinical models from local electronic health records based on semantic similarity. *J. Biomed. Inform.* **2015**, *54*, 294–304. [CrossRef]
57. Shanavas, N.; Wang, H.; Lin, Z.; Hawe, G. Knowledge-driven graph similarity for text classification. *Int. J. Mach. Learn. Cybern.* **2021**, *12*, 1067–1081. [CrossRef]
58. Weng, W.-H.; Chung, Y.-A.; Tong, S. Clinical Text Summarization with Syntax-Based Negation and Semantic Concept Identification. *arXiv* **2020**, arXiv:2003.00353.
59. Sugumaran, V.; Storey, V.C. Ontologies for conceptual modeling: Their creation, use, and management. *Data Knowl. Eng.* **2002**, *42*, 251–271. [CrossRef]
60. McInnes, B.T.; Pedersen, T. Evaluating semantic similarity and relatedness over the semantic grouping of clinical term pairs. *J. Biomed. Inform.* **2015**, *54*, 329–336. [CrossRef] [PubMed]
61. Sammut, C.; Webb, G.I. (Eds.) *Encyclopedia of Machine Learning*; Springer: Boston, MA, USA, 2011.
62. Jaccard, P. Nouvelles recherches sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.* **1908**, *44*, 223–270.
63. Cai, R.; Zhu, B.; Ji, L.; Hao, T.; Yan, J.; Liu, W. An CNN-LSTM Attention Approach to Understanding User Query Intent from Online Health Communities. In Proceedings of the 2017 IEEE International Conference on Data Mining Workshops (ICDMW), New Orleans, LA, USA, 18–21 November 2017.
64. Sarrouti, M.; El Alaoui, S.O. SemBioNLQA: A semantic biomedical question answering system for retrieving exact and ideal answers to natural language questions. *Artif. Intell. Med.* **2020**, *102*, 101767. [CrossRef]
65. Afzal, M.; Alam, F.; Malik, K.M.; Malik, G.M. Clinical Context–Aware Biomedical Text Summarization Using Deep Neural Network: Model Development and Validation. *J. Med. Internet Res.* **2020**, *22*, e19810. [CrossRef]
66. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
67. Khattak, F.K.; Jeblee, S.; Pou-Prom, C.; Abdalla, M.; Meaney, C.; Rudzicz, F. A survey of word embeddings for clinical text. *J. Biomed. Inform.* **2019**, *100*, 100057. [CrossRef]
68. Resnik, P. Using Information Content to Evaluate Semantic Similarity. Proceedings of the 14th International Joint Conference on Artificial Intelligence. Available online: https://arxiv.org/abs/cmp-lg/9511007 (accessed on 26 December 2022).
69. Le, Q.; Mikolov, T. Distributed representations of sentences and documents. *Int. Conf. Mach. Learn.* **2014**, *32*, 1188–1196.
70. Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching word vectors with subword information. *Trans. Assoc. Comput. Linguist.* **2017**, *5*, 135–146. [CrossRef]

71. National Library of Medicine. UMLS Meta Thesaurus Fact Sheet. Available online: http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html (accessed on 18 May 2016).
72. Boguraev, B.; Briscoe, T.; Carroll, J.; Carter, D.; Grover, C. The derivation of a grammatically indexed lexicon from the Longman Dictionary of Contemporary English. In Proceedings of the 25th conference on Association for Computational Linguistics, Stanford, CA, USA, 6–9 July 1987; pp. 193–200.
73. National Library of Medicine. UMLS Specialist Lexicon Fact Sheet. Available online: http://www.nlm.nih.gov/pubs/factsheets/umlslex.html (accessed on 18 May 2016).
74. Bada, M. Mapping of biomedical text to concepts of lexicons, terminologies, and ontologies. *Methods Mol. Biol.* **2014**, *1159*, 33–45. [CrossRef] [PubMed]
75. Sánchez, D.; Batet, M. Semantic similarity estimation in the biomedical domain: An ontology-based information-theoretic perspective. *J. Biomed. Inform.* **2011**, *44*, 749–759. [CrossRef] [PubMed]
76. Batet, M.; Sánchez, D.; Valls, A.; Gibert, K. Semantic similarity estimation from multiple ontologies. *Appl. Intell.* **2013**, *38*, 29–44. [CrossRef]
77. Jiang, R.; Gan, M.; Dou, X. From ontology to semantic similarity: Calculation of ontology-based semantic similarity. *Sci. World J.* **2013**, *2013*, 793091.
78. SNOMED International. SNOMED—Home—SNOMED International. 2019. Available online: http://www.snomed.org/ (accessed on 6 November 2019).
79. Available online: https://bioportal.bioontology.org/ontologies/RCD (accessed on 26 December 2022).
80. Available online: https://bioportal.bioontology.org/ontologies/NDFRT (accessed on 26 December 2022).
81. Available online: https://bioportal.bioontology.org/ontologies/ICD10 (accessed on 26 December 2022).
82. Available online: https://www.ncbi.nlm.nih.gov/mesh (accessed on 26 December 2022).
83. MedDRA MSSO—MedDRA. Available online: https://www.meddra.org/about-meddra/organisation/msso (accessed on 6 November 2019).
84. Cai, X.; Liu, S.; Yang, L.; Lu, Y.; Zhao, J.; Shen, D.; Liu, T. COVIDSum: A linguistically enriched SciBERT-based summarization model for COVID-19 scientific papers. *J. Biomed. Inform.* **2022**, *127*, 103999. [CrossRef]
85. Wehrli, E. Fips, a deep linguistic multilingual parser. In Proceedings of the ACL Workshop on Deep Linguistic Processing, Prague, Czech Republic, 28 June 2007; pp. 120–127.
86. Noh, J.; Kavuluru, R. Document retrieval for biomedical question answering with neural sentence matching. In Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, 17–20 December 2018; pp. 194–201.
87. Moradi, M.; Samwald, M. Clustering of Deep Contextualized Representations for Summarization of Biomedical Texts. *arXiv* **2019**, arXiv:1908.02286.
88. Beltagy, I.; Lo, K.; Cohan, A. Scibert: A pretrained language model for scientific text. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 9 November 2019; pp. 3615–3620.
89. Sanh, V.; Debut, L.; Chaumond, J.; Wolf, T. DistilBERT, a distilled version ofBERT: Smaller, faster, cheaper and lighter. *arXiv* **2019**, arXiv:1910.01108.
90. Wang, J.; Dong, Y. Measurement of Text Similarity: A Survey. *Information* **2020**, *11*, 421. [CrossRef]
91. Ben Aouicha, M.; Taieb, M.A.H. Computing semantic similarity between biomedical concepts using new information content approach. *J. Biomed. Inform.* **2016**, *59*, 258–275. [CrossRef]
92. Han, M.; Zhang, X.; Yuan, X.; Jiang, J.; Yun, W.; Gao, C. *A Survey on the Techniques, Applications, and Performance of Short Text Semantic Similarity*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2020.
93. Cajiao, A.Z.; Mateus, A.R. Graph-based Similarity for Document Retrieval in the Biomedical Domain. In Proceedings of the 2022 7th International Conference on Machine Learning Technologies (ICMLT), Rome Italy, 11–13 March 2022.
94. Chen, X.; Jia, S.; Xiang, Y. A review: Knowledge reasoning over knowledge graph. *Expert Syst. Appl.* **2020**, *141*, 112948. [CrossRef]
95. Plaza, L.; Díaz, A.; Gervás, P. A semantic graph-based approach to biomedical summarisation. *Artif. Intell. Med.* **2011**, *53*, 1–14. [CrossRef] [PubMed]
96. Deza, M.M.; Deza, E. Encyclopedia of distances. In *Encyclopedia of Distances*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–583.
97. Jaccard, P. The distribution of the flora in the alpine zone. *New Phytol.* **1912**, *11*, 37–50. [CrossRef]
98. Andoni, A.; Indyk, P.; Krauthgamer, R. Earth mover distance over high-dimensional spaces. In Proceedings of the Symposium on Discrete Algorithms, San Francisco, CA, USA, 20–22 January 2008; pp. 343–352.
99. Manning, C.D.; Schütze, H. *Foundations of Statistical Natural Language Processing*; MIT Press: Cambridge, MA, USA, 1999.
100. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86. [CrossRef]
101. Iliopoulos, C.S.; Rahman, M.S. New efficient algorithms for the LCS and constrained LCS problems. *Inf. Process. Lett.* **2008**, *106*, 13–18. [CrossRef]
102. Dice, L.R. Measures of the amount of ecologic association between species. *Ecology* **1945**, *26*, 297–302. [CrossRef]
103. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.

104. Landauer, T.K.; Dumais, S.T. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol. Rev.* **1997**, *104*, 211–240. [CrossRef]
105. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
106. Sak, H.; Senior, A.; Beaufays, F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. *arXiv* **2014**, arXiv:1402.1128.
107. Li, Y.; Bandar, Z.A.; Mclean, D. An approach for measuring semantic similarity between words using multiple information sources. *IEEE Trans Knowl. Data Eng.* **2003**, *15*, 871–882.
108. Lin, C.-Y. ROUGE: A Package for Automatic Evaluation of Summaries. In *Text Summarization Branches Out*; Association for Computational Linguistics: Barcelona, Spain, 2004; pp. 74–81.
109. Schulze, F.; Neves, M. Entity-Supported Summarization of Biomedical Abstracts. In Proceedings of the Fifth Workshop on Building and Evaluating Resources for Biomedical Text Mining (BioTxtM2016), Osaka, Japan, 11–16 December 2016.
110. Aramaki, E.; Miura, Y.; Tonoike, M.; Ohkuma, T.; Mashuichi, H.; Ohe, K. Text2table: Medical text summarization system based on named entity recognition and modality identification. In Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing, Boulder, CO, USA, 4–5 June 2009; pp. 185–192.
111. Moradi, M.; Ghadiri, N. Quantifying the informativeness for biomedical literature summarization: An itemset mining method. *Comput. Methods Programs Biomed.* **2017**, *146*, 77–89. [CrossRef]
112. Agrawal, R.; Imieliński, T.; Swami, A. Mining association rules between sets of items in large databases. *ACM SIGMOD Rec.* **1993**, *22*, 207–216. [CrossRef]
113. Agrawal, R.; Mannila, H.; Srikant, R.; Toivonen, H.; Verkamo, A.I. Fast Discovery of Association Rules. *Adv. Knowl. Discov. Data Min.* **1996**, *12*, 307–328.
114. Moradi, M.; Ghadiri, N. Different approaches for identifying important concepts in probabilistic biomedical text summarization. *Artif. Intell. Med.* **2018**, *84*, 101–116. [CrossRef]
115. Balinsky, A.; Balinsky, H.; Simske, S. *On the Helmholtz Principle for Data Mining*; Hewlett-Packard Development Company, LP.: Palo Alto, CA, USA, 2011.
116. Azadani, M.N.; Ghadiri, N.; Davoodijam, E. Graph-based biomedical text summarization: An itemset mining and sentence clustering approach. *J. Biomed. Inform.* **2018**, *84*, 42–58. [CrossRef]
117. Zhang, W.; Yoshilda, T.; Tang, X.; Wang, Q. Text clustering using frequent itemsets. *Knowl.-Based Syst.* **2010**, *23*, 379–388. [CrossRef]
118. Moradi, M. Frequent Itemsets as Meaningful Events in Graphs for Summarizing Biomedical Texts. In Proceedings of the 2018 8th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 25–26 October 2018; pp. 135–140.
119. Balinsky, H.; Balinsky, A.; Simske, S.J. Automatic text summarization and small-world networks. In Proceedings of the 11th ACM Symposium on Document Engineering, Mountain View, CA, USA, 19–22 September 2011; pp. 175–184.
120. Moradi, M. CIBS: A biomedical text summarizer using topic-based sentence clustering. *J. Biomed. Inform.* **2018**, *88*, 53–61. [CrossRef]
121. Larose, D.T. *Discovering Knowledge in Data: An Introduction to Data Mining*; John Wiley & Sons: Hoboken, NJ, USA, 2014.
122. Rouane, O.; Belhadef, H.; Bouakkaz, M. Combine clustering and frequent itemsets mining to enhance biomedical text summarization. *Expert Syst. Appl.* **2019**, *135*, 362–373. [CrossRef]
123. Salton, G.; Wong, A.; Yang, C.S. A vector space model for automatic indexing. *Commun. ACM* **1975**, *18*, 613–620. [CrossRef]
124. Macqueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California: Los Angeles, CA, USA, 1967; pp. 281–297.
125. Lee, E.K.; Uppal, K. CERC: An interactive content extraction, recognition, and construction tool for clinical and biomedical text. In Proceedings of the 10th International Workshop on Biomedical and Health Informatics, San Diego, CA, USA, 18–20 November 2019.
126. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
127. Apache Lucene. Available online: http://lucene.apache.org (accessed on 26 December 2022).
128. Bada, M.; Eckert, M.; Evans, D.; Garcia, K.; Shipley, K.; Sitnikov, D.; Baumgartner, W.A.; Cohen, K.B.; Verspoor, K.; Blake, J.A.; et al. Concept annotation in the CRAFT corpus. *BMC Bioinform.* **2012**, *9*, 161. [CrossRef]
129. Savova, G.K.; Masanz, J.J.; Ogren, P.V.; Zheng, J.; Sohn, S.; Kipper-Schuler, K.C.; Chute, C.G. Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications. *J. Am. Med. Inform. Assoc.* **2010**, *17*, 507–513. [CrossRef]
130. Rouane, O. Word Embedding-Based Biomedical Text Summarization. In *Emerging Trends in Intelligent Computing and Informatics, Proceedings of the 4th International Conference of Reliable Information and Communication Technology (IRICT2019), Johor, Malaysia, 22–23 September 2019*; Springer: Cham, Switzerland, 2019. [CrossRef]
131. Text Data Preprocessing. Keras. Available online: https://keras.io/preprocessing/text/ (accessed on 7 October 2020).
132. Sarker, A.; Yang, Y.-C.; Al-Garadi, M.A.; Abbas, A. A Light-Weight Text Summarization System for Fast Access to Medical Evidence. *Front. Digit. Health* **2020**, *2*. [CrossRef]
133. Davoodijam, E.; Ghadiri, N.; Shahreza, M.L.; Rinaldi, F. MultiGBS: A multi-layer graph approach to biomedical summarization. *J. Biomed. Inform.* **2021**, *116*, 103706. [CrossRef]
134. MetaMap—A Tool for Recognizing UMLS Concepts in Text. Available online: https://metamap.nlm.nih.gov/ (accessed on 25 April 2019).

135. Basaldella, M.; Furrer, L.; Tasso, C.; Rinaldi, F. Entity recognition in the biomedical domain using a hybrid approach. *J. Biomed. Semant.* **2017**, *8*, 51. [CrossRef] [PubMed]
136. Rindflesch, T.; Fiszman, M. The interaction of domain knowledge and linguistic structure in natural language processing: Interpreting hypernymic propositions in biomedical text. *J. Biomed. Inform.* **2003**, *36*, 462–477. [CrossRef]
137. Rahmede, C.; Iacovacci, J.; Arenas, A.; Bianconi, G. Centralities of nodes and influences of layers in large multiplex networks. *J. Complex Netw.* **2018**, *6*, 733–752. [CrossRef]
138. Zahid, M.A.H.; Mittal, A.; Joshi, R.; Atluri, G. CLINIQA: A Machine Intelligence Based Clinical Question Answering System. *arXiv* **2006**, arXiv:1805.05927.
139. Lin, R.T.; Chiu, J.L.-T.; Dai, H.-J.; Day, M.-Y.; Tsai, R.T.-H.; Hsu, W.-L. Biological question answering with syntactic and semantic feature matching and an improved mean reciprocal ranking measurement. In Proceedings of the 2008 IEEE International Conference on Information Reuse and Integration, Las Vegas, NV, USA, 13–15 July 2008; pp. 184–189.
140. Kogan, Y.; Collier, N.; Pakhomov, S.; Krauthammer, M. Towards Semantic Role Labeling & IE in the Medical Literature. *AMIA Annu. Symp. Proc.* **2005**, *2005*, 410–414.
141. Miller, G.A.; Beckwith, R.; Fellbaum, C.; Gross, D.; Miller, K.J. Introduction to WordNet: An On-line Lexical Database*. *Int. J. Lexicogr.* **2004**, *3*, 235–244. [CrossRef]
142. Gobeill, J.; Patsche, E.; Theodoro, D.; Veuthey, A.-L.; Lovis, C.; Ruch, P. Question answering for biology and medicine. In Proceedings of the 2009 9th International Conference on Information Technology and Applications in Biomedicine, Larnaka, Cyprus, 4–7 November 2009; pp. 1–5.
143. Cao, Y.; Liu, F.; Simpson, P.; Antieau, L.D.; Bennett, A.S.; Cimino, J.; Ely, J.; Yu, H. AskHERMES: An online question answering system for complex clinical questions. *J. Biomed. Inform.* **2011**, *44*, 277–288. [CrossRef]
144. Robertson, S.; Zaragoza, H.; Taylor, M. Simple BM25 extension to multiple weighted fields. In Proceedings of the thirteenth ACM International Conference on Information and Knowledge Management, Washington, DC, USA, 8–13 November 2004.
145. Cairns, B.L.; Nielsen, R.D.; Masanz, J.J.; Martin, J.H.; Palmer, M.S.; Ward, W.H.; Savova, G.K. The mipacq clinical question answering system. In *AMIA Annual Symposium Proceedings*; American Medical Informatics Association: Bethesda, MD, USA, 2011; Volume 2011, p. 171.
146. Ely, J.; Osheroff, J.; Chambliss, M.; Ebell, M.; Rosenbaum, M. Answering Physicians' Clinical Questions: Obstacles and Potential Solutions. *J. Am. Med. Inform. Assoc.* **2005**, *12*, 217–224. [CrossRef]
147. Medpedia. Available online: http://www.medpedia.com/ (accessed on 26 December 2022).
148. Ni, Y.; Zhu, H.; Cai, P.; Zhang, L.; Qui, Z.; Cao, F. CliniQA: Highly Reliable Clinical Question Answering System. *Stud. Health Technol. Inform.* **2012**, *180*, 215–219.
149. Available online: www.tripanswers.org (accessed on 26 December 2022).
150. Athenikos, S.J.; Han, H.; Brooks, A.D. A Framework of a Logic-based Question-Answering System for the Medical Domain (LOQAS-Med). In Proceedings of the 2009 ACM symposium on Applied Computing, Honolulu, HI, USA, 8 March 2009. [CrossRef]
151. NLM Clinical Questions Collection. Available online: http://clinques.nlm.nih.gov/ (accessed on 26 December 2022).
152. Abacha, A.B.; Zweigenbaum, P. MEANS: A medical question-answering system combining NLP techniques and semantic Web technologies. *Inf. Process. Manag.* **2015**, *51*, 570–594. [CrossRef]
153. Balikas, G.; Krithara, A.; Partalas, I.; Paliouras, G. BioASQ: A challenge on large-scale biomedical semantic indexing and question answering. In *International Workshop on Multimodal Retrieval in the Medical Domain*; Springer: Cham, Switzerland, 2015.
154. Peng, S.; You, R.; Wang, H.; Zhai, C.; Mamitsuka, H.; Zhu, S. Deepmesh: Deep semantic representation for improving large-scale mesh indexing. *Bioinformatics* **2016**, *32*, i70–i79. [CrossRef]
155. Manning, C.D.; Surdeanu, M.; Bauer, J.; Finkel, J.; Bethard, S.J.; McClosky, D. The stanford CoreNLP natural language processing toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 23–24 June 2014. [CrossRef]
156. Xie, W.; Ding, R.; Yan, J.; Qu, Y. A Mobile-Based Question-Answering and Early Warning System for Assisting Diabetes Management. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 9163160. [CrossRef]
157. Zhang, X.; Wu, J.; He, Z.; Liu, X.; Su, Y. Medical Exam Question Answering with Large-Scale Reading Comprehension. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018.
158. Zhu, X.; Yang, X.; Chen, H. A Biomedical Question Answering System Based on SNOMED-CT. In Proceedings of the International Conference on Knowledge Science, Engineering and Management, Changchun, China, 17–19 August 2018.
159. Ferrández, Ó.; Micol, D.; Muñoz, R.; Palomar, M. *DLSITE-1: Lexical Analysis for Solving Textual Entailment Recognition*; Kedad, Z., Lammari, N., Métais, E., Meziane, F., Rezgui, Y., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4592, pp. 284–294. [CrossRef]
160. Brokos, G.I.; Liosis, P.; McDonald, R.; Pappas, D.; Androutsopoulos, I. AUEB at BioASQ 6: Document and Snippet Retrieval. *arXiv* **2018**, arXiv:1809.0636.
161. Hui, K.; Yates, A.; Berberich, K.; de Melo, G. PACRR: A position-aware neural IR model for relevance matching. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 9–11 September 2017; pp. 1049–1058.

162. Guo, J.; Fan, Y.; Ai, Q.; Croft, W.B. A deep relevance matching model for ad-hoc retrieval. In Proceedings of the 25th ACM International Conference on Information and Knowledge Management, Indianapolis, IN, USA, 24–28 October 2016; pp. 55–64.
163. Yin, W.; Schutze, H.; Xiang, B.; Zhou, B. ABCNN: Attention-based convolutional neural network for modeling sentence pairs. *Trans. Assoc. Comput. Linguist.* **2016**, *4*, 259–272. [CrossRef]
164. Metzler, D.; Croft, W.B. A Markov random field model for term dependencies. In Proceedings of the 28th Annual International ACM SIGIR Conference. ACM, Salvador, Brazil, 15–19 August 2005; pp. 472–479.
165. Sarrouti, M.; Alaoui, S.O.E. A machine learning-based method for question type classification in biomedical question answering. *Methods Inf. Med.* **2017**, *56*, 209–216. [CrossRef]
166. Ozyurt, I.B.; Bandrowski, A.; Grethe, J.S. Bio-AnswerFinder: A system to find answers to questions from biomedical texts. *Database* **2020**, *2020*, baz137. [CrossRef]
167. Yan, Y.; Zhang, B.; Li, X.; Liu, Z. List-wise learning to rank biomedical question-answer pairs with deep ranking recursive autoencoders. *PLoS ONE* **2020**, *15*, e0242061. [CrossRef] [PubMed]
168. Dina, D.F.; Yassine, M.; Asma, B.A. Consumer health information and question answering: Helping consumers find answers to their health-related information needs. *J. Am. Med. Inform. Assoc.* **2020**, *27*, 194–201.
169. Almeida, T.; Matos, S. *Calling Attention to Passages for Biomedical Question Answering*; Springer Nature: Cham, Switzerland, 2020; pp. 69–77.
170. McDonald, R.; Brokos, G.I.; Androutsopoulos, I. Deep Relevance Ranking Using Enhanced Document-Query Interactions. *arXiv* **2018**, arXiv:1809.01682.
171. Alzubi, J.A.; Jain, R.; Singh, A.; Parwekar, P.; Gupta, M. COBERT: COVID-19 Question Answering System Using BERT. *Arab. J. Sci. Eng.* **2021**, 1–11. [CrossRef] [PubMed]
172. Available online: https://www.kaggle.com/allen-institutefor-ai/CORD-19-research-challenge (accessed on 26 December 2022).
173. Liang, J.; Tsou, C.-H. A Novel System for Extractive Clinical Note Summarization using EHR Data. In Proceedings of the 2nd Clinical Natural Language Processing Workshop, Minneapolis, MN, USA, 7 June 2019; pp. 46–54.
174. Gupta, S.; Sharaff, A.; Nagwani, N.K. *Biomedical Text Summarization: A Graph-Based Ranking Approach*; Advances in Intelligent Systems and Computing; Springer: Singapore, 2021; Volume 1354.
175. Gupta, S.; Sharaff, A. Frequent item-set mining and clustering based ranked biomedical text summarization. *J. Supercomput.* **2022**, *79*, 139–159. [CrossRef]
176. Erkan, G.; Radev, D.R. LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. *J. Artif. Intell. Res.* **2004**, *22*, 457–479. [CrossRef]
177. Debnath, P.; Castillo, O.; Kumam, P. (Eds.) *Soft Computing: Recent Advances and Applications in Engineering and Mathematical Sciences*; CRC Press: Boca Raton, FL, USA, 2023.