This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Unleashing Competitive Intelligence: News Mining Analysis on Technology Trends and Digital Health Driving Healthcare Innovation

Enrique Cano-Marin, Salvador Sánchez-Alonso, and Marçal Mora-Cantallops

***Abstract*—In the rapidly evolving digital health landscape, technology plays a pivotal role in transforming the healthcare industry. With the exponential growth of data, uncovering valuable insights has become a daunting task. In today's data-driven world, healthcare businesses must leverage emerging technologies to stay informed about trends in their field. This research article presents a novel approach to deriving business insights in digital health enabled by technology, including artificial intelligence, and other cutting-edge advancements. We propose a methodology that utilizes news mining techniques and the global data on events, location, and tone database as the primary data source. By employing natural language processing, we developed a practical way of extracting relevant insights from vast amounts of public data. We implemented named-entity recognition (NER) enriched with the DBpedia knowledge base and relationship extraction. In addition, we leveraged graph analytics to identify and analyze the most significant concept relationships within the text corpus and their evolution in time. By integrating these advanced techniques, healthcare businesses can extract actionable insights from public datasets, empowering them to stay abreast of emerging trends and advancements in digital health, such as telehealth, precision medicine, or medical imaging.**

***Index Terms*—Digital health, emerging technologies, entity extraction, graph analytics, healthcare, innovation, natural language processing (NLP), news mining.**

I. Introduction

TECHNOLOGY and innovation have played a critical role in shaping the landscape of healthcare over the past few decades. From advancements in medical imaging technologies to the emergence of personalized medicine and telemedicine services for patient monitoring, rapid technological progress has the potential to revolutionize healthcare delivery and significantly improve patient outcomes [1].

Simultaneously, the ongoing digitization trend within the healthcare industry has created numerous opportunities to enhance the effectiveness of diagnostic and therapeutic processes. This is in line with the research of Gërguri-Rashiti et al. [2], who concluded that technology adoption and the development of innovation activities have positive impacts on firm performance. This trend is evident in the rising numbers of technology entrepreneurs entering the healthcare sector, aiming to capitalize on these opportunities [3]. Significant contributions to this transformative shift include the use of artificial intelligence (AI) for early detection and diagnosis [4], of remote surgical procedures, and of telemedicine tools for patient advice and monitoring [5]. Advances in AI, including machine learning (ML), natural language processing (NLP), blockchain, cloud computing, and the Internet of Things (IoT) have catalyzed the efficient collection and analysis of vast amounts of data (Big Data) [6]. The technologies encompassing AI, blockchain, cloud computing, and Big Data analytics, collectively known as the ABCD technologies [7], are enablers of Industry 4.0 [8]. This has facilitated the delivery of healthcare services with greater efficiency, increased optimization, and unlocked new opportunities [9].

The exponential growth of digital data in recent years has made Big Data a strategic advantage for firms and a facilitator of knowledge creation [10]. However, access to data and the insights derived from its analysis remain challenging. As a reference, in the United States (US), only between 20% and 29% of unstructured patient data are estimated to be accessible and available for analysis [11] and, according to the work in [12], approximately 80% of enterprise information assets (growing at a rate of 65% per year, partially motivated by the development of IoT) are estimated to remain unstructured. Meanwhile, challenges such as data access and severe ethical and social concerns regarding privacy, accuracy, and safety must also be addressed [13]. Keeping up with the volume of public information available can also be challenging.

Access to data is crucial not only for generating valuable insights and driving evidence-based decision-making in healthcare, but also for business and competitive intelligence [14]. News mining enables access to a vast amount of public data from various news sources, offering public information that can be leveraged for strategic purposes, for example, by identifying business collaborations, partnerships, and competitors [15], predicting the profitability of businesses [15], [16] or stocks prices [17] and by identifying technology trends [18] or their environmental sustainability [19]. These data empower businesses to make informed decisions, develop competitive strategies, and stay ahead of the curve in a rapidly evolving industry. By harnessing the power of news mining for business and

Manuscript received 16 June 2023; revised 4 September 2023 and 13 October 2023; accepted 15 October 2023. Review of this manuscript was arranged by Department Editor D. Ribeiro-Soriano. *(Corresponding author: Enrique Cano-Marin.)*

The authors are with the Computer Science Department, University of Alcalá, 28801 Alcalá de Henares, Spain (e-mail: enrique.canom@uah.es; salvador.sanchez@uah.es; marcal.mora@uah.es).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TEM.2023.3326233.

Digital Object Identifier 10.1109/TEM.2023.3326233

competitive intelligence, organizations can gain a competitive edge, capitalize on emerging opportunities, and confidently navigate the complexities of the healthcare market.

Despite the existing research on leveraging social media for healthcare insights, as detailed in our previous research on the use of social media analytics in healthcare [20], there are still significant gaps in the literature that require further research and the exploration of different data sources. Recent studies on news mining analysis in healthcare have focused mainly on the impact of COVID-19 and public health [21], [22], [23], and the impact of the digital transformation during the pandemic [24]. However, no studies have leveraged news mining to gain insights into healthcare businesses. Therefore, we propose a new approach, utilizing news mining analysis to gain insights into the role of technology in enabling innovation in healthcare and, thereby, provide competitive intelligence for business organizations.

This article delves into the impact of technology on innovation in healthcare by utilizing news mining techniques and analyzing data gathered from the Global Data on Events, Location, and Tone (GDELT) database [25]. The main objective of the study is to develop an NLP-based approach for recognizing underlying knowledge and patterns within a collection of news articles to grasp the influence of technology on driving healthcare advancements. Through this research, we intend to build upon current understandings of technology's role in facilitating digital health and offer new perspectives on both the difficulties and possibilities in this field. Our contribution also extends to the NLP area, as we establish a systematic process involving entity extraction and graph analysis. This pipeline holds potential for broad application in news extraction across various domains, beyond healthcare innovation.

The rest of the article is organized as follows. Section II provides a comprehensive review of the existing literature using news mining to generate insights and the goals of this research. In Section III, we detail the methodology implemented for this study, which relies on the use of NLP techniques. Section IV presents the results obtained from our study and the discussion on the validation of the hypotheses. Finally, Section V concludes this article, including a critical assessment of the limitations of our study and an overview of potential future research opportunities in this field.

II. Theoretical Framework

The healthcare industry is undergoing continuous transformation, driven by technological advancements that accelerate digital health innovation [26]. The rise of digital health has further expanded the possibilities for innovation in healthcare. With the proliferation of wearable devices, mobile apps, and telemedicine platforms, patients can now monitor their health remotely, access medical advice and support from anywhere, and receive personalized treatments tailored to their specific needs [27]. Digital health offers numerous business opportunities for innovation and growth [28]. The objective of digital health is to enhance healthcare efficiency, accessibility, and effectiveness by utilizing digital technology to gather, analyze, store, and exchange health data. Healthcare companies can leverage new technologies to develop innovative products and services, improve operational efficiency, and deliver better outcomes for patients and healthcare professionals [29]. However, the digital health landscape is complex and rapidly evolving, and companies must navigate a range of challenges related to patient safety [30], data privacy and regulatory compliance [31], and interoperability and equitable access [32].

Digital health encompasses a range of technologies that include telemedicine, medical imaging, mobile health (mHealth) apps, wearable technology, medical devices, electronic health records (EHRs), and other innovative digital technologies. These advances aim to enhance patient care by enabling more effective and efficient healthcare delivery, particularly in the context of precision medicine [33]. The progress of wearable technology and medical devices, coupled with the utilization of the IoT, has made personalized healthcare services a realistic ambition. These technologies are catalysts for the real-time collection of health data from patients through mHealth apps [34]. These apps and devices also aid physicians and healthcare professionals in various ways [35], supporting the discovery of new drugs during clinical trials [36], the diagnosis of diseases [37], and the treatment of patients. We also underline the progress made in medical imaging in the fields of radiology, pathology, and oncology with the development of AI and ML, with models that can even outperform imaging specialists [38]. In addition, it is worth mentioning the impact of the coronavirus disease (COVID-19) pandemic in accelerating the adoption of telemedicine [39] to mitigate the risks associated with this disease and facilitate access to health services through the use of technology. Overall, another milestone in digital health is the advancement of EHRs as comprehensive repositories of patient health information [40]. Through interoperability, digital health technologies enable the integration and sharing of patient health information across various healthcare providers, improving the coordination and continuity of care. This interoperability-driven approach in digital health has the potential to overcome the fragmentation of health data and promote a more holistic view of patient health, ultimately enhancing healthcare outcomes and patient experiences [41].

Technology serves as the foundation for the evolution of digital health, disrupting the healthcare industry and ushering in innovative solutions that enhance patient care [42]. As pointed out by Sigov et al. [43], Industry 4.0 plays a pivotal role in shaping the digital health landscape [44]. This paradigm entails the integration of cutting-edge technologies, such as AI/ML, IoT, Big Data analytics, cloud computing, blockchain, metaverse, and robotics to enable data-driven decision-making, automation, and improved connectivity [9]. Within this context, the present research study aims to evaluate the interplay between emerging technology trends and the progressive evolution of digital health. One notable technological advancement is the rapid and widespread adoption of blockchain in recent years. Blockchain offers compelling advantages, including enhanced interoperability, privacy, security, and the immutability of patient data [45]. These attributes make it an appealing solution for addressing the healthcare sector's challenges in managing sensitive patient information. Furthermore, the convergence of the metaverse with virtual reality, augmented reality, and mixed reality technologies—collectively referred to as XR technologies—alongside blockchain, opens up exciting possibilities for healthcare applications [46]. This synergy enables the creation of immersive experiences and innovative services such as telemedicine, including medical rehabilitation and mental health needs [47]. The combination of XR technologies and the trust and security provided by blockchain offers immense potential for transforming healthcare delivery.

By exploring the link between technology trends and the dynamic development of digital health, this research study aims to contribute to a deeper understanding of the evolving healthcare landscape and drive the advancement of patient-centered care.

A. *News Mining for Competitive Intelligence*

In today's data-rich business environment, social media analytics has become a critical tool for businesses seeking to gain insights and stay ahead of the competition, and healthcare is no exception. News mining consists of the extraction of valuable insights, latent knowledge, and trends from news articles or blogs [48], and it is often classified as part of social media analytics, as both involve analyzing online-accessible unstructured text [49]. By leveraging news mining techniques, businesses can track market trends, monitor their brand reputation, and identify potential risks and opportunities [50]. Besides, innovation has been linked to the usage of text-mining analytics for predicting its evolution [51]. Along these lines, Li et al. [18] combined news mining with patent data analysis to identify trends and emerging technologies applied to the perovskite solar cell technology. In addition, de Lima-Santos and Ceron [52] completed a news mining analysis on AI in journalism, highlighting the usage of computer vision for fact-checking and ML for recommendations, optimization, and strategic planning (for example, for predicting reader churn rate).

One major challenge in the news mining and social media analytics is the need to process large amounts of text data in a way that is accurate, efficient, and scalable [53]. This requires advanced techniques for NLP, ML, and information retrieval, as well as specialized tools and software platforms that can handle the volume and complexity of the data [54]. The quality and quantity of news articles available for analysis can vary significantly across different sources and topics [55]. Biases, inaccuracies, and omissions in news reports can introduce noise and affect the validity of the analysis, hence the importance of ensuring both quality and relevance. In this line, Choi et al. [56] leveraged NLP and text-mining analysis to predict the quality of news articles, based on journalistic values such as accuracy, objectivity, impartiality, and fairness. Along these lines, we highlight the research in [57], which explored the usage of AI to ensure quality and trust to mitigate the rapid spread of disinformation and suggested the use of NLP for trends detection and knowledge integration. In addition, the availability of data may be limited to specific sources or languages, restricting the comprehensiveness and representativeness of the analysis [58]. It is crucial to consider these limitations and critically evaluate the sources and scope of news coverage when conducting news mining analysis, to ensure accurate and reliable insights [59].

Finally, integrating news mining with existing business intelligence systems and decision-making processes can be challenging [16], particularly in complex industries such as healthcare. This requires careful consideration of factors such as data privacy, security, and governance, as well as effective communication and collaboration between stakeholders in different parts of the organization. Despite these challenges, news mining has enormous potential to generate valuable business insights into healthcare innovation by identifying emerging trends, technologies, and regulatory developments. As such, it is an essential tool for companies seeking to stay ahead of the curve in this rapidly evolving industry. Therefore, we formulate the following hypothesis.

H1: The availability of the full text of news articles is a critical factor affecting the accuracy and completeness of news mining analysis for business intelligence.

News mining can be precious for businesses seeking to stay informed about the current progress in their industries, including emerging trends [60], competitor analysis [61], and regulatory changes [62]. In this context, we highlight the research in [63], which evaluated practical case studies in which businesses across various industries leveraged social media analytics to gain a competitive advantage. Along these lines, with the help of advanced ML and NLP techniques, businesses can gain a competitive advantage [14] when extracting actionable insights from news articles, empowering them to make informed data-driven decisions. To exemplify this potential, Lin and Hsu [64] evaluated the use of text-mining techniques applied to news reports for facilitating business leaders' decision-making processes about corporate social responsibility. In addition, combining these data with other public and quantitative data sources leads to more robust analyses, which have been proven effective when applied to stock market price forecasting [16]. In the field of healthcare, Zolnoori et al. [59] completed an analysis of public health concerns, from obesity and asthma to smoking, depression, and alcohol, analyzing over 3 000 000 articles for over a decade (2007–2017), concluding the utility of text-mining techniques such as topic modeling and sentiment analysis for generating insights that support policymakers and help to develop public health strategies [59]. Along these lines, Bertl [65] extracted and processed over 300 000 news articles from GDELT to identify cybersecurity threats, vulnerabilities, and attacks related to digital health. The COVID-19 outbreak also prompted increased interest in news mining analysis in public health, as reflected in the work [66], [67], which assessed GDELT news articles to evaluate the impact of the pandemic. Therefore, the hypothesis 2 (H2) is as follows.

H2: News mining analysis can provide valuable business insights by identifying emerging trends and technologies.

We also propose that the dynamic nature of relationships and entities extracted from news articles plays a crucial role in generating insights related to healthcare innovation. These concepts and entities are extracted by leveraging named-entity recognition (NER), in line with the work in [68]. As news articles evolve over time, the connections between different entities, such as organizations, technologies, and trends, may change or develop. These evolving relationships can significantly influence the accuracy, completeness, and relevance of insights derived from news mining analysis in the domain of healthcare innovation. Hence, our third hypothesis is as follows.

H3: The evolution of relationships over time within news articles significantly impacts the generation of insights.

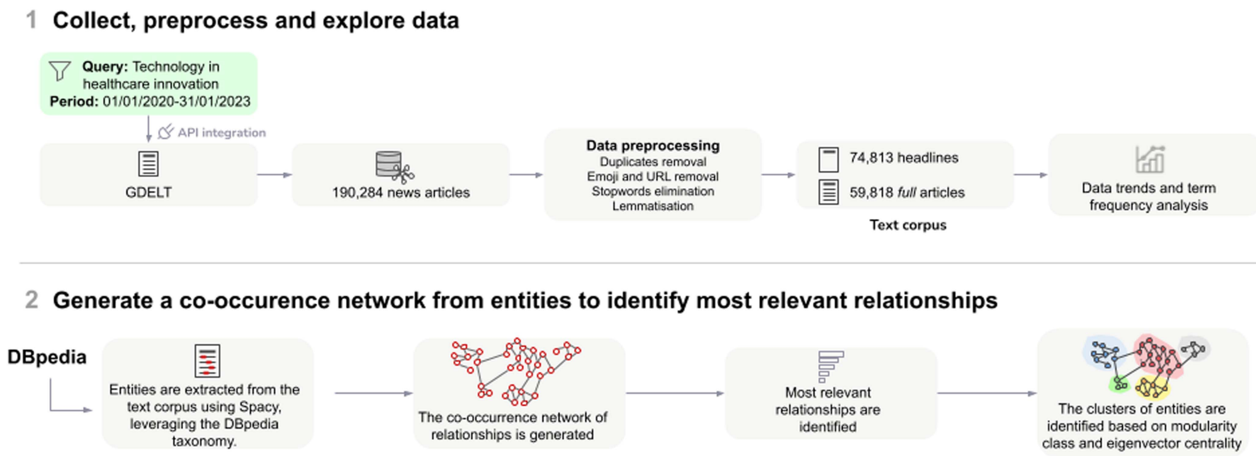

Fig. 1. Methodology flowchart.

III. Methodology

This research article employed NLP techniques to analyze automatically the content of news articles about healthcare innovation. The goal was to generate competitive intelligence by identifying trends and integrating knowledge [57]. This section provides a clear overview of the study's precise procedures, whose steps were selecting the data source, preprocessing textual data, and utilizing NLP techniques such as named-entity extraction (NER), as in the research of the work in [69], relationship extraction [22], and clustering using graph-based models. Fig. 1 presents a schematic flowchart of the methodology implemented in this research, which was applied to a dataset on technology in healthcare innovation as a practical contribution to its empirical validation. It could be potentially implemented in various domains by defining a different search strategy.

A. Data Collection and Data Preprocessing

For this analysis, we used the Global Database of Events, Language, and Tone (GDELT) as our primary data source. GDELT is a public database that collects news articles across regions worldwide, making it a valuable resource for researchers interested in studying media coverage and text-mining analysis [25]. The GDELT project built the global knowledge graph (GKG) of global media news, and it is continuously updated every 15 min [70]. By using GDELT as our data source, we could access a diverse range of news articles worldwide, providing a comprehensive and global perspective on the role of technology in healthcare innovation.

To collect our data, we applied a search query in GDELT using the following terms.

"sourcelang:english (innovation OR technology) (healthcare OR health) (near10:"technology healthcare" OR near 10:"technology health") (theme: WB_1331_HEALTH_TECHNOLOGIES OR theme: WB_376_INNOVATION_TECHNOLOGY_AND_ENTREPRENEURSHIP)."

On the one hand, the language of the retrieved news articles is English. On the other hand, we leveraged the taxonomy defined in the GDELT GKG Themes to filter the news articles depending on their topic [70]. This search strategy enabled us to gather a comprehensive and diverse set of news articles that discuss the intersection of healthcare and digital innovation from various sources and perspectives. We specifically extracted a subset of articles from GDELT that contained the keywords "innovation," "technology," "health," and "healthcare."

The resulting dataset included the headlines of articles, the link to the article, the source, and when these were published, in the period between January 1, 2020, and January 30, 2023. This period was defined to ensure that recent advancements in the field were also captured. In total, we collected 190 284 news reports, which provided a rich and diverse source of data for our analysis. The data were retrieved through the publicly available application programming interface (API) and the resulting news reports were stored in a MongoDB database. MongoDB is a NoSQL document database that has the following advantages over traditional relational databases. As mentioned in [71], one of the primary advantages of MongoDB is its flexibility, enabling the storage and retrieval of unstructured and semistructured data, including news articles dynamically and efficiently. In addition, MongoDB's document-based model enables faster and easier scaling, as it eliminates the need for complex and costly join operations. Another advantage is its ability to handle large amounts of data, making it well-suited for Big Data applications.

At this point, duplicated news articles were removed, after which 74 813 remained in the scope of the research. The full-text version of each remaining article was also collected, if available. To complete the preparation of the corpus, all headlines and full-text articles were preprocessed by removing stopwords, URLs, and emojis, and performing lemmatization on the text [72].

To validate the relevance of the dataset containing news articles for the proposed analysis and the validity of insights in the context of technology in healthcare innovation, a group of six experts in the field evaluated the dataset. Initially, a set of news articles was extracted from the corpus and ranked using a Likert scale ranging from 1 to 5 [56], where 1 indicated low relevance and 5 indicated high relevance. We validated the evaluation by calculating the interrater agreement using Krippendorff's alpha, a suitable metric for assessing agreement among multiple raters and multiple categories [73]. In addition, we calculated the sample size required to ensure a confidence interval of 95% with a margin error of 10%, as defined in [74]

$$S = \frac{\frac{z^2 \cdot p \cdot (1-p)}{e^2}}{1 + \frac{z^2 \cdot p \cdot (1-p)}{e^2 \cdot N}} \quad (1)$$

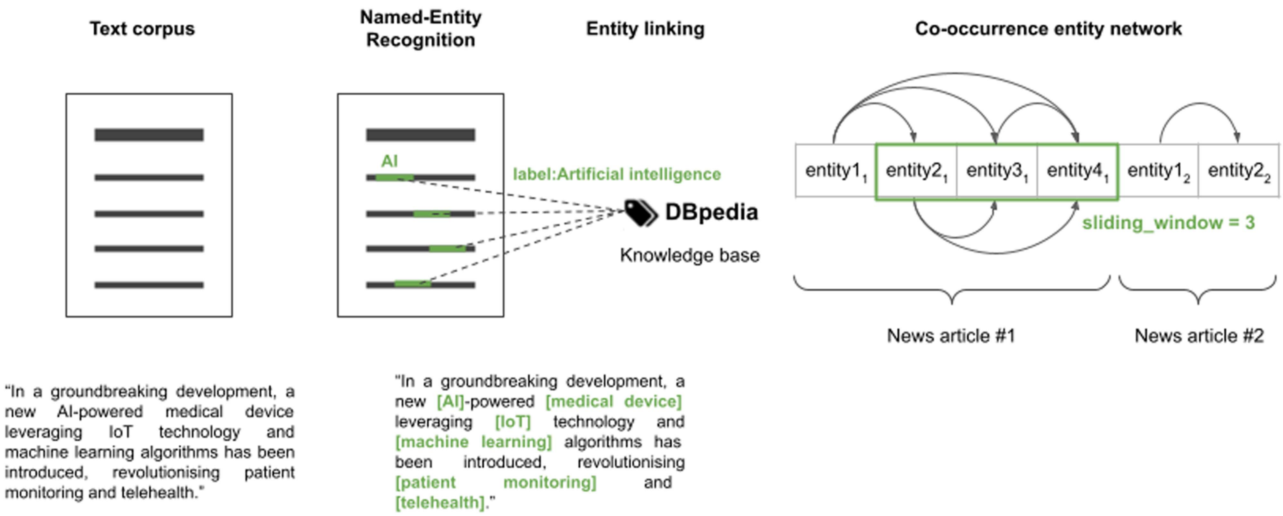

Fig. 2. Relation analysis based on entity extraction and entity linking.

where N is the size of the population, z is the desired level of confidence (in standard deviations), equals to 1.96 if the confidence interval is 95%, p represents the estimated proportion or probability of the event occurring in the population, with a value of 0.5, and e is the desired margin of error.

Equation (1) yields a sample size of 96. As a result of evaluating 100 sample articles, the experts agreed that the dataset has a relevance of 73.90% with a 95% confidence interval and a margin of error lower than 10%. In addition, there was a moderate interrater agreement with Krippendorff's alpha ($\alpha = 0.58$). Therefore, we confirm that the dataset is valid for achieving the goals defined in this study and drawing conclusions from the results and insights obtained.

B. NER, Entity Linking, and Relationship Extraction

Extracting valuable insights from unstructured texts may be a cumbersome task. One of the techniques used to tackle this challenge is NER. NER involves identifying and classifying named entities in text into predefined categories such as person names, organization names, locations, dates, and other specific types [75]. The goal of NER is to extract and label these named entities in order to understand better the information present in the text. It is vital to differentiate NER from entity linking, also known as named-entity disambiguation, which is defined as the task of linking or connecting named entities mentioned in the text to their corresponding entities in a knowledge base or a specific reference source. The goal of entity linking is to resolve the ambiguity that may arise due to multiple entities sharing the same name or similar names [76]. For example, the term "AI" is linked to the concept of "artificial intelligence" from the knowledge base, acting as a unique identifier. As identified by Abhishek et al. [77], we used DBpedia applied to news mining. DBpedia is a large-scale ontology in English extracted from Wikipedia [78]. In practice, we used DBpedia Spotlight for spaCy [79]. spaCy is an open-source Python library for various NLP tasks, such as NER. In order to ensure a high-quality entity recognition process, the entities were only extracted if the similarity between the entity and the potential candidate from DBpedia identified by spaCy was higher than 95%. The entities were extracted and linked to the concepts from the ontology in both headlines and full-text news articles. Fig. 2 schematizes the implemented NLP pipeline for entity recognition, linking, and relationship extraction in the news corpus.

Evolving from NER and entity linking, relation extraction is defined as the NLP task that involves automatically identifying and extracting relationships or associations between entities mentioned in text. This technique is widely used in the creation of knowledge graphs [80], which are defined as structured representations of knowledge that capture relationships between entities and concepts. To identify the relation between the concepts, an entity-based co-occurrence network is generated. Two entities are considered co-occurrent if both appear in the same document (either headline or full-text) in the same context. The context is defined by a sliding-window approach, which involves moving a fixed-size window across a text corpus, capturing the co-occurrence relationships among them. In general, co-occurrence networks are widely used in a myriad of research fields, from bibliometrics [81], legal [82], or social network analysis and graph-based topic modeling techniques [83]. In this research, the sliding window is set to 3. In order to build the co-occurrence network, the weight of the relationship varies with the distance between entities, down from a maximum distance of 3, as per the defined sliding window. In the network, the nodes represent the entities, the edges, and the relations between entities.

In order to identify the main clusters of entities, we leveraged the Louvain algorithm for community detection proposed in [84]. This algorithm effectively identifies communities or groups of nodes within a network that demonstrate a higher degree of interconnectedness than nodes outside of these groups. The algorithm's operation is based on the concept of modularity, where the modularity class is defined as a measure of the network's partitionability. Modularity quantifies the extent to which a network can be divided into communities by evaluating the difference between the number of edges within communities and the expected number of edges in a randomly connected network. Each cluster is assigned a modularity class by the algorithm. Once the nodes are organized into clusters, we

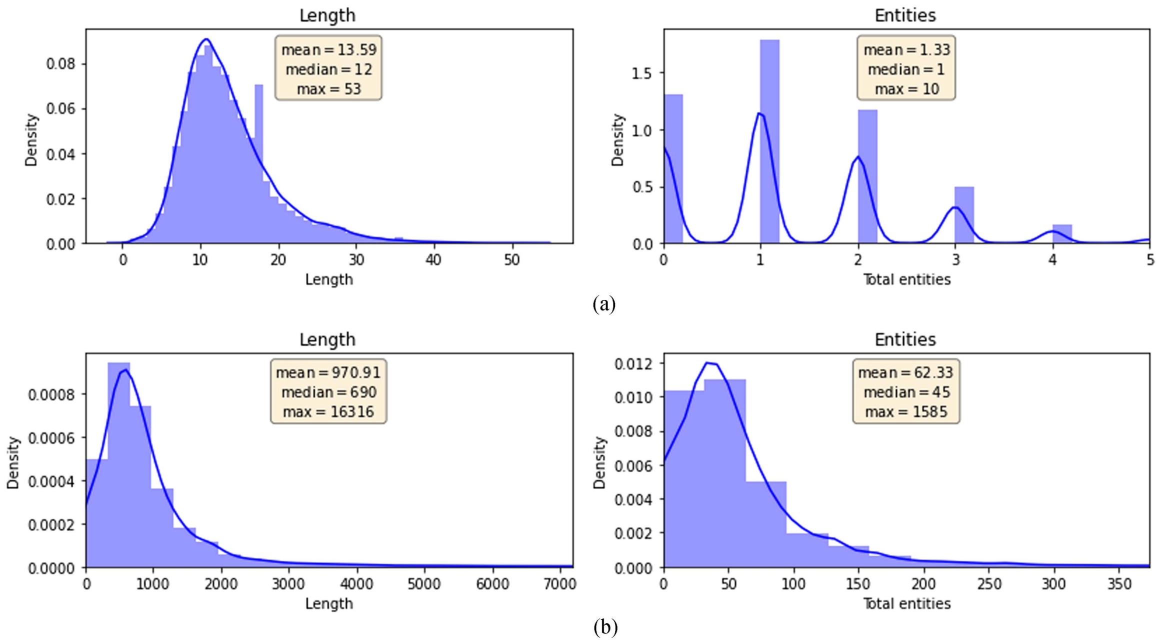

Fig. 3. Length of documents and entities extracted from headlines and full articles (percentile = 99%).

determine the most significant entities by calculating their eigenvector centrality. Eigenvector centrality provides a measure of a node's influence within a network by considering both its direct connections and the centrality of the nodes to which it is connected [85]. We used NetworkX and Gephi for network analysis and clustering purposes. While NetworkX, a Python library, allowed us to generate, manipulate, and analyze complex network structures and their dynamics [86], Gephi, an open-source software, provided a robust platform for visually exploring networks. With its built-in algorithms and interactive visualizations, Gephi offered unique deeper insights into network structure and dynamics [87].

IV. Findings

In this section, we discuss the outcomes of the experiments and analyses conducted to fulfill the goals regarding the potential of news mining for generating insights on the enabling role of technology in healthcare innovation. This is a practical contribution of the proposed methodology for gaining competitive intelligence from news articles, and it can be conceptually extended to various domains.

A. Criticality of Headlines in News Mining Analysis

Firstly, aligned with H1, we conducted a comprehensive comparison of the information available for analysis between headlines and full-text articles. Headlines capture readers' attention and encourage them to read a news article [88]. Therefore, our assumption was to test whether valuable insights could be generated from headlines, bearing in mind their public availability in contrast with the increase of paywalls in newspapers with the presence of digital newspapers [89]. These paywalls limit access to the full-text version of an article and, based on the research in [90], the content available upon payment is more diverse. In addition, the analysis of headlines would be faster and less computationally expensive. In the present research, it is crucial to emphasize that a significant majority of the news articles in our corpus, precisely 80.3%, had an accompanying full-text version, resulting in a substantial collection of 59 818 complete news articles. The disparity in content volume becomes apparent when examining Fig. 3, which illustrates that headlines consist of an average of 14 words while the corresponding articles are substantially longer, comprising 971 words on average (representing a staggering 70-fold increase). Moreover, our analysis of NER and entity linking revealed a striking disparity between headlines and full articles. On average, a headline contains only one entity, whereas full articles exhibit an average of 62 entities, with a median of 45. In order to complete the analysis of headlines, we could potentially leverage NLP techniques focused on the analysis of short and sparse text, such as the collected from social media, as per the research in [91]. However, given these findings, we decided to proceed with the analysis using the corpus composed exclusively of news articles that possess an available full version.

B. Business Insights About Emerging Trends and Technologies

In line with H2, we built the co-occurrence network of entities extracted from the full-text version of news articles, as described in the methodology. As a result, we obtained a network formed by 117 734 nodes and 3 490 307 edges, where nodes represent entities based on the DBpedia knowledge base and the edges the relationships between these concepts.

First, the entities are clusterized as described in Table I, analogously as in graph-based topic modeling techniques [83], [92], [93]. We observe that the nodes are grouped into six main clusters, representing 97.52% of the total number of entities in the network. The most relevant and influential entities are extracted by sorting each community by eigenvector centrality. The COVID-19 pandemic is the most represented concept, with 23.39% of entities belonging to this community. These entities represent various aspects of the pandemic, including the virus

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE I
CLUSTERS OF ENTITY RELATIONSHIP NETWORK

Description	Entities	Eigenvector centrality	%
COVID-19 pandemic	Pandemic, Coronavirus disease 2019, Severe acute respiratory syndrome coronavirus 2, China, Vaccine, Coronavirus disease, India, European Union, Coronavirus, Ebola virus disease, Asia, Africa	185.85	23.39
Healthcare systems	Telehealth, Public Health, Medicine, Mental health, Physician, Communication, Medicaid, Health system, Primary care, Health Insurance	184.99	23.28
Technology in healthcare	Artificial intelligence, Digital media, Digital health, Pharmaceutical industry, Biology, Information technology, Medical device, Biotechnology, Supply chain, Research and development, Machine Learning, Applied Digital Data Systems, Electronic health record, Health care, Cloud computing	142.36	17.91
Medical research	Medication, Cancer, Oncology, Clinical trial, Diabetes, Pipeline, Cardiovascular disease, Biopharmaceutical, HIV, Cardiology, Pain, High-throughput screening, Cancer Screening, Patent, Precision Medicine	136.84	17.22
Healthcare market	CEO, United States, FDA, General officer, NASDAQ, Intellectual property, Stock, Private equity, Insurance	124.97	15.72
Environmental health and sustainability	Agriculture, Global warming, Renewable energy, Earth, Greenhouse gas, NASA, Energy development, Recycling, Wind power, Food safety, Sustainable energy, Solar energy, Fossil fuel, Sustainability, Carbon footprint	12.31	1.54

TABLE II
TOP COVID-19 AND NON-COVID-19-RELATED RELATIONSHIPS IN THE ENTITY-BASED CO-OCCURRENCE NETWORK

Source	Destination	Weight
Coronavirus disease 2019	Pandemic	26 316
Severe acute respiratory syndrome coronavirus 2	Pandemic	20 519
Severe acute respiratory syndrome coronavirus 2	Vaccine	17 215
Coronavirus disease 2019	Vaccine	12 437
Vaccine	Severe acute respiratory syndrome coronavirus 2	12 283
Pandemic	Coronavirus disease 2019	9 604
Severe acute respiratory syndrome coronavirus 2	Virus	9 498
Artificial intelligence	Machine learning	9 254
Pandemic	Severe acute respiratory syndrome coronavirus 2	8 711
China	Virus	8 709
Pandemic	Telehealth	7 923
...	...	...
Digital health	Digital media	2 299
Digital health	Telehealth	2 225
Telehealth	Content management system	2 044
Medical imaging	Health informatics	1 903
Digital media	Artificial intelligence	1 755
Digital media	Telehealth	1 726

itself (severe acute respiratory syndrome Coronavirus 2), related terms such as vaccines and geographical locations that represent its spread worldwide as a public health emergency originated in Wuhan, China [94], along with the challenges that were faced in the mitigation of the disease, the unbalanced distribution of resources, and the inequitable access to the vaccine in different regions (India, Europe, Asia, Africa). In addition, according to Table II, 9 out of 10 of the most relevant relations extracted from the corpus are related to the COVID-19 pandemic. We emphasize the identified connection in [95] between the rise of telehealth and the progression of the pandemic: they conducted a systematic literature review on the role of telehealth as an effective way to mitigate the COVID-19 outbreak. Telehealth enabled patients to access health services while reducing the risk of contagion, which was crucial due to the highly infectious nature of the virus.

The second most-represented concept identified, at 23.38%, was healthcare systems, which play a critical role in providing essential medical services and support to individuals and communities. Public health is a fundamental component of healthcare systems, focusing on disease prevention, promotion of well-being and mental health, and treatment of illnesses. The inclusion of telehealth underlines the growing importance of remote healthcare delivery and its potential to improve access to medical services, as highlighted in the case of the COVID-19 pandemic. Primary care serves as the foundation of healthcare systems, offering essential preventive care, health maintenance, and coordination of care for individuals. Here, physicians are integral to healthcare systems, as they diagnose, treat, and manage various health conditions. In the case of health insurance and Medicaid, a government program in the US that supports healthcare for individuals and families with low income, this reflects the relevance of ensuring access to healthcare services by providing financial coverage and reducing the burden of medical expenses for patients.

Third, we observe the role of technology in accelerating the digital transformation of health services. Technologies such as AI, ML, and cloud computing play a crucial role in advancing digital health, particularly in the maturation of medical devices and EHRs. Technology also acts as an enabler for the digital transformation of the pharmaceutical industry and its entire value chain, encompassing research and development, manufacturing, and the supply chain. This transformation requires adaptation to fast-paced innovation and the restructuring of business processes and organizational structures [96]. The rapid advancement of technology is driving a profound digital transformation within the healthcare sector: Cutting-edge technologies such as AI, ML, and cloud computing are playing pivotal roles in the evolution of digital health services. These technologies are instrumental in the design of innovative medical devices and the establishment of EHRs that enhance patient care and data management. From research and development processes to manufacturing and supply chain management, technology has the power to revolutionize and streamline operations in the pharmaceutical sector. This digital transformation enables more efficient drug discovery, enhanced quality control, optimized manufacturing processes, and a more agile and transparent supply chain.

The fourth most-represented concept identified, at 17.22%, is medical research. In healthcare, medication plays a crucial role, as represented in fields, such as oncology, diabetes, and cardiovascular disease. Clinical trials are conducted to test new medications while precision medicine tailors treatments based on individual characteristics. The biopharmaceutical industry drives innovation and research. High-throughput and cancer screening aid in early detection. Patents protect intellectual property. The pipeline of new drugs and clinical trials contributes to advancing medicine and improving patient outcomes. The healthcare market is influenced by the US, linked to the role of the Food and Drug Administration (FDA), the regulatory agency responsible for ensuring the safety and efficacy of food, drugs, and medical devices, and the NASDAQ stock exchange, where companies list their shares for public trading in the US. Sustainability was the fifth-most represented concept, which underscores the environment's profound impact on human health. In this connection, the news mined contains many mentions of global warming, pollution, and carbon footprint, and the efforts being made to mitigate their effects, such as green energy sources (e.g., solar and wind power).

In Table II, we observe that the link between AI and ML is the most relevant entity relationship for terms unrelated to the COVID-19 pandemic. The top relationships unrelated to the COVID-19 pandemic are presented below the line of dots. In addition to the advancements in telehealth, it is important to highlight the progress made in medical imaging, which is closely intertwined with the development of computer vision and AI [97]. These technological advancements have led to practical diagnostic applications, particularly in fields such as ophthalmology and radiology.

Digital health is revolutionizing the healthcare landscape by integrating disruptive technologies such as AI, blockchain, cloud computing, Big Data, and other innovative solutions, empowering patients and healthcare providers to enhance access, efficiency, and personalized care delivery. Therefore, we focused on evaluating the role of technology in enabling digital health. To do so, we analyzed the relationships between technology concepts and digital health trends, in line with the research in [33]. The Sankey diagram in Fig. 4 illustrates the flow between various technology trends and their corresponding digital health trends based on the weight of their relationships, with a minimum weight of 100. The thickness of the links in the diagram represents the strength of the relationship between the technologies and the areas of digital health where they are applied: wider links indicate stronger relationships.

Fig. 4 shows that telehealth is the digital health trend with the broadest presence. From an infrastructure perspective, telehealth is built on the foundations of cloud computing and 5G, whereas XR technologies, IoT, and metaverse are enablers for its implementation. It is worth mentioning that precision medicine is intrinsically related to Big Data, supported by AI and ML in the analysis of the massive amount of unstructured and structured data collected from heterogeneous data sources such as patient information, scientific research, and real-world data. Medical imaging has also been disrupted in recent years, by adopting AI-based techniques in computer vision, with the implementation of algorithms that outperform even expert physicians on diagnosis [38]. In the case of wearable technology, we observe the link with IoT intended for collecting patient health data in real time that can be further processed by leveraging AI and ML and the role of XR technologies, supporting the user experience. In the development of EHR, AI is a building block technology for making EHR a reality, minimizing the interoperability issues in data management in healthcare and easing access to their data for patients. It is enabled by NLP for processing digitized paper-based information and reports from physicians, and leveraging automation for minimizing this cumbersome and repetitive task for healthcare professionals. It is important to mention the potential of blockchain in enabling EHR, ensuring the privacy, immutability, auditability, and consistency of the data [98] by leveraging the benefits of this type of distributed ledger technology (DLT), in line with the research in [41].

In the case of medical devices, this technology is closely related to the development of IoT and AI/ML, analogously as in the case of wearable technology. Medical devices are also linked to 3-D printing, especially for creating prototypes and low-cost solutions that reduce the access barrier to these kinds of devices, along with the ability to create sophisticated geometric configurations tailored to the specific anatomical features of individual patients. The application of 3-D printing in healthcare has seen notable advances, which has prompted regulatory agencies to oversee and guide its usage [99]. Cloud computing too plays a key role in the development of medical devices for supporting the devices' connectivity, the storage and processing of the collected data, and the infrastructure required for serving the applications that provide service to users, including both patients and healthcare professionals. Along these lines, Klonoff [100]

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

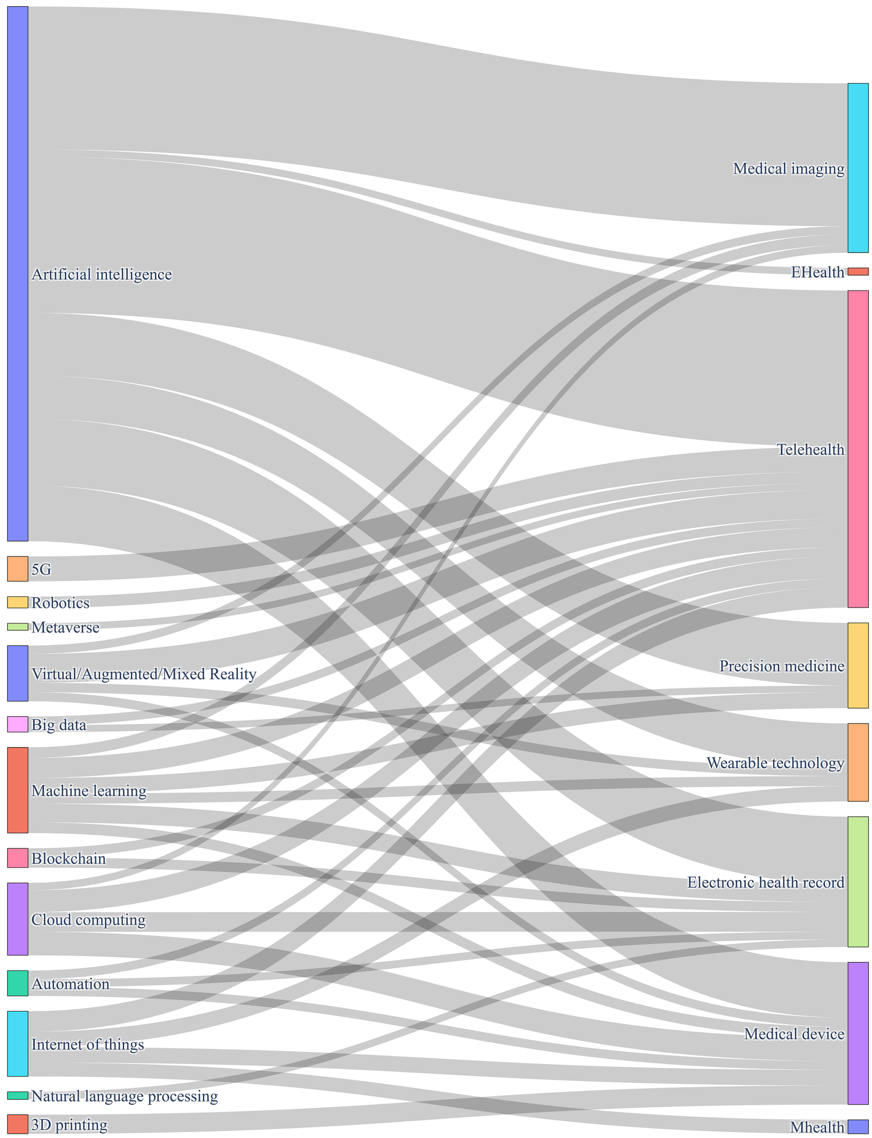

Fig. 4. Relationship between technology trends and digital health trends.

researched the IoT-based cloud computing architectures supporting data processing from diabetes devices, which reflects the relationships identified in the text corpus. It is worth noting the concept of software as a medical device (SaMD), defined by the FDA as the software designed to fulfill one or more medical purposes independently, separate from any hardware medical device [101]. The rise of AI/ML techniques has sparked greater interest in SaMDs, prompting regulatory efforts to govern the use of AI/ML, focusing on potential impacts and consequences, and ensuring safety and effectiveness [102].

Finally, eHealth and mHealth are the digital health fields with the lowest presence in the corpus. As mHealth involves using mobile devices to gather real-time health data from patients, its connection to the advancement of medical devices and IoT is evident, as discussed in the research conducted by Almotiri et al. [103].

From a technology standpoint, we observe that NLP, blockchain, metaverse, robotics, and automation are underrepresented in the news mining corpus. This is a limitation that may partially bias the news mining analysis and which has already been studied in academia by authors such as Bourgeois et al. [104], who ran a coverage analysis to measure this effect by using a corpus of news articles from GDELT, and Buckingham et al. [19], who also leveraged GDELT data to generate nearly real-time insights in the context of environmental studies.

The set of radar charts in Fig. 5 depicts the normalized distribution of interest for each technology trend and its links with the leading digital health areas. In the case of AI, they reveal that medical imaging and telehealth are the most prominent categories, in line with the results presented in Table II. These categories are followed by EHR and precision medicine, linked to the capabilities for processing increasing amounts of available data (Big Data) by leveraging AI techniques such as NLP or ML [105]. Similarly, the second chart shows that telehealth and EHR dominate the ML domain, indicating their significance in leveraging ML for healthcare applications, ML plays a significant role in precision medicine by leveraging computational techniques to uncover intricate patterns in data. This enables the identification of predictive models and classification algorithms that can be utilized to make accurate predictions and

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

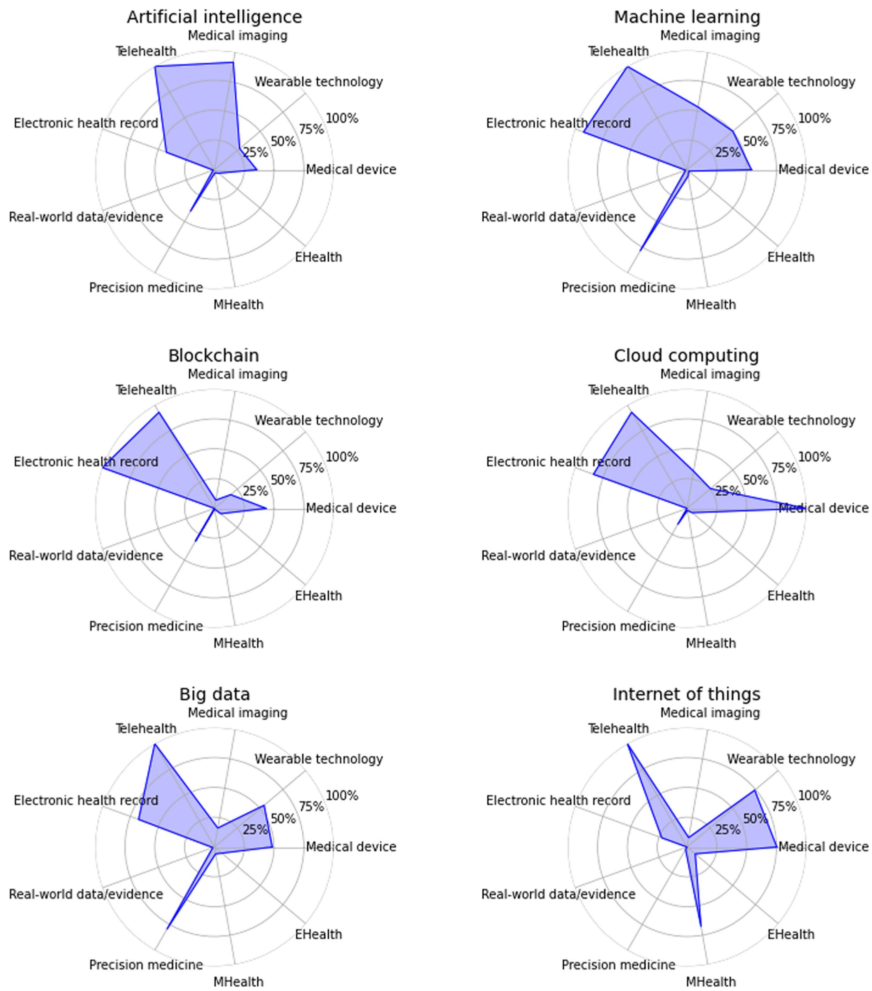

Fig. 5. Digital health trends normalized radar chart by technology trend.

conduct advanced exploratory data analysis [106]. As a branch of AI, in combination with computer vision techniques, ML has the potential to improve clinical practices in diagnostics while leveraging medical images. However, in line with the research in [107], there are still limitations in the use of ML in medical imaging that need to be addressed to ensure the benefits to patients, such as biases in the datasets or the choice of model evaluation. On the other hand, RWD/RWE and e-health exhibit lower percentages, suggesting a lesser emphasis on ML applications and a lower degree of presence in the news coverage. In the realm of blockchain technology, the third chart highlights the prevalence of telehealth as the top category. The integration of blockchain technology for health services is still in its early stages of adoption, as there are still challenges to be faced, such as data security and data interoperability across healthcare systems [108], as reflected in categories such as RWD/RWE and precision medicine appear to have limited representation. This raises questions about the potential benefits and challenges of incorporating blockchain in healthcare contexts that rely heavily on real-world evidence and precision medicine. Following the work in [109], this technology has enormous potential in telehealth as it can enhance the security and privacy of health data by utilizing smart contracts to improve compliance with patients while also enabling the quick settlement of healthcare transactions for continuous remote monitoring. It can also help detect fraud in physician credentials and medical testing kits and ensure data security and patient privacy, making it suitable for the digitization and automation of telehealth services. The fourth chart underscores the dominance of telehealth as the most significant category within the cloud computing domain, along with EHR and medical devices, which also hold notable percentages. As shown in [110], cloud computing, in combination with other Industry 4.0 technologies such as Big Data and IoT, can transform the healthcare ecosystem. We see this link in the strong presence of all these technologies for the development of medical devices that allow personalized quality services to be provided to patients, enabling the launch of *Healthcare as a Service* [111], which heavily relies on the analysis of the data collected from patients. Finally, IoT is connected to wearable and medical devices that facilitate mHealth in the adoption of telehealth. Nazir et al. [112] argued that IoT, in conjunction with mobile computing, can facilitate mHealth, ensuring privacy and security in health IoT devices, and contributing to the advancement of smart hospitals and remote patient monitoring [113].

One of the implications of these findings is that categories that demonstrate limited presence should be explored further and invested in, as they may hold untapped potential for technological advancements in healthcare. Moreover, the charts can help

TABLE III
EVOLUTION IN TIME OF MAIN RELATIONSHIPS FOR TECHNOLOGY TRENDS

		AI		ML		Blockchain		Cloud computing		Big data		Internet of Things	
Year	**Half-Year**	**Top terms**	**Weight**	**Top terms**	**Weight**	**Top terms**	**Weight**	**Top terms**	**Weight**	**Top terms**	**Weight**	**Top terms**	**Weight**
2020	H1	ML	2,019	AI	2,019	AI	310	AI	400	AI	372	AI	201
		Coronavirus disease 2019	1,439	Automation	77	Supply chain	206	Microsoft	261	Cloud computing	116	Computer security	152
		Telehealth	314	Drug discovery	57	Cryptocurrency	174	IoT	145	Public health	69	Cloud computing	145
	H2	ML	1,383	AI	1,383	Bitcoin	178	AI	338	AI	381	AI	252
		Pandemic	705	Cloud computing	131	AI	145	ML	131	Cloud computing	103	5G	129
		Telehealth	305	Amazon Web Services	78	Digital media	124	Big Data	77	ML	68	Pandemic	46
2021	H1	ML	2,320	AI	2,320	AI	220	AI	437	AI	328	AI	359
		FDA	285	FDA	202	Cryptocurrency	131	Information technology	200	IQVIA	48	Computer security	112
		Cloud computing	269	Drug discovery	64	Privacy	51	Pandemic	89	Privacy	31	Cloud computing	43
	H2	ML	2,507	AI	2507	AI	145	AI	431	AI	406	AI	244
		Digital media	589	Natural language processing	67	Supply chain	103	Microsoft	215	ML	76	ML	37
		Big Data	270	Cancer	59	Smart contract	86	Computer security	69	Digital health	47	Wireless	37
2022	H1	ML	1,880	AI	2,698	AI	193	AI	440	AI	354	AI	302
		Digital health	574	Digital media	87	Digital media	71	Digital media	234	IQVIA	99	mHealth	88
		Drug discovery	334	Deep learning	62	Metaverse	66	Digital transformation	107	Predictive analytics	38	Blockchain	41
	H2	ML	2,128	AI	2,128	AI	156	AI	347	AI	279	AI	308
		Radiology	587	Pharmaceutical industry	59	Cryptocurrency	144	Computer security	60	ML	42	Blockchain	107
		Drug discovery	584	Biology	54	IoT	59	IoT	55	Pharmaceutical industry	36	Telehealth	105
2023	H1	ML	488	AI	488	Drug discovery	97	Software as a service	30	AI	33	Olea	34
		Radiology	255	Digital media	61	Privacy	53	Digital health	19	Cloud computing	9	AI	19
		Medical imaging	212	Aida	23	Health care	20	Electronic health record	19	Electronic health record	8	Supply chain	13

decision-makers, researchers, and practitioners identify areas where technology integration and innovation can have the greatest impact, leading to improved healthcare delivery, enhanced patient outcomes, and more efficient healthcare systems.

C. Insights From the Evolution of Entity Relations in News Mining

The results presented in Table III provide quantitative observations that have significant implications for the technological landscape and its impact on various domains in healthcare innovation. First, the dominance of AI and ML as the most prominent terms throughout the period underscores their continued significance and influence. This underlines the widespread adoption and integration of AI and ML technologies that reflect their transformative potential and suggests that they are critical drivers of innovation and advancement in healthcare. Second, it is highlighted the interdependence of AI with the different technologies. AI appears as a top term not only within the AI and ML categories, respectively, but also in blockchain, cloud computing, Big Data, and the IoT in its evolution in time. This interrelation highlights the crucial role of AI in shaping and advancing other technology domains and signifies the increasing reliance on AI as a foundational technology that drives innovation and progress in various sectors, in line with the work in [114]. Third, we observe the effect of the COVID-19 pandemic during 2020, in line with the increase in interest in telehealth. Fourth, we observe the maturity of AI in developing specific healthcare applications with noticeable results, such as drug discovery [115], radiology, and medical imaging [116]. Researchers such as Gupta et al. [117] emphasized the potential AI has to accelerate and optimize drug discovery, leveraging techniques such as ML, deep learning, and the recently developed generative AI models such as large-language models (LLMs) [118]. All of these AI techniques allow for data harmonization from heterogeneous sources and fuel collaborative data networks for better clinical trial data management [119]. AI is also contributing to the development of personalized healthcare (PHC) [120]. Furthermore, in the case of blockchain, cloud computing, Big Data, and the IoT, these present a growing importance in the digital landscape. This indicates a shift toward data-driven approaches, decentralized systems, and interconnected devices, all of which are central to the digital transformation and the development of intelligent ecosystems. In the healthcare industry, specific terms such as telehealth, drug discovery, and digital health underscore the increasing significance of technology. The prominence of these terms aligns with the ongoing digital transformation in healthcare, where AI, ML, and the IoT are being leveraged to improve patient care, enable remote monitoring, and drive advancements in personalized medicine. Initially, blockchain emerged as a concept that could revolutionize various domains, including finance and supply chain management. However, it has evolved from foundational applications, such as cryptocurrency and smart contracts to practical applications in drug discovery, healthcare, and privacy. This signifies the growing recognition and acceptance of blockchain as a transformative technology. It reflects the healthcare industry's growing interest in leveraging blockchain for various purposes, such as secure health data exchange, disease prediction, remote patient monitoring, and tracking [121] or supply chain [122] while ensuring privacy protection and enabling interoperability [123]. The increased prominence of blockchain in healthcare highlights its maturation as a technology and its potential to address critical challenges within the industry.

The inclusion of regulatory terms such as FDA highlights the influence of external factors on technology trends, emphasizing the impact of regulatory bodies on shaping the development and adoption of innovative technologies for healthcare services.

While the benefits of adopting these cutting-edge technologies are evident, there is an obvious need as well for robust regulatory frameworks that ensure patient privacy, transparency in decision-making, and bias minimization [124].

In addition, these technologies are being adopted across the entire value chain of healthcare systems: the supply chain, the pharmaceutical industry, and diagnostics (e.g., radiology). This signals the diverse range of possible applications and the potential for technology to revolutionize business and processes in healthcare beyond general technology domains. Finally, the fluctuations in term weights across different years demonstrate the dynamic nature of technology trends. New terms emerge, and the importance of existing terms may vary over time, reflecting the evolving landscape of technology and its priorities. This dynamic nature emphasizes the need for continuous monitoring and adaptation to stay abreast of the latest technological advancements. Hence, news mining plays a crucial role in our attempts to understand and analyze the dynamic technology trends in healthcare. By extracting relevant information from open-access news articles, news mining provides insights into the evolving landscape of technology, and its impact on healthcare services.

V. Conclusion

In summary, this research study offers a robust NLP-driven methodology for effective entity extraction, linking, and relationship extraction in the news mining analysis. In the context of healthcare innovation, this method lays the foundation for generating competitive intelligence and empowering strategic development through real-time trend and pattern analysis. Furthermore, this methodology exhibits potential for expansion into diverse domains (apart from healthcare) by harnessing the power of news articles and NLP.

A vital consideration arising from this analysis is the availability of data in the context of news mining. The findings of this research highlight the importance of utilizing publicly available news data sources to extract valuable insights. The abundance of news data provides researchers with a unique opportunity to leverage text-mining analytics techniques for generating competitive intelligence and conducting comprehensive analyses. The accessibility of news data enhances the transparency and repeatability of research findings, enabling researchers to validate and build upon previous studies. Therefore, it is essential to remark on the wealth of publicly available high-quality news data, as it is a valuable resource for scientific investigations.

In addition, as a practical contribution to the field of healthcare, we also offer useful insights for business leaders into the role of emerging technologies as catalysts for innovation in digital health. Specifically, we emphasize the value of utilizing NLP and text-mining analytics techniques applied to news mining in effectively generating competitive intelligence from unstructured data collected from public sources. The findings of this study emphasize the potential that emerges from integrating emerging technologies, i.e., AI/ML and blockchain, with healthcare data. This integration not only enhances decision-making processes but also contributes to improved patient outcomes within the digital health landscape. Furthermore, this research sheds light on the evolving trends in digital health through the lens of news mining analytics, specifically in areas such as telehealth, medical imaging, and medical devices, where the integration of various technologies plays a pivotal role in advancing healthcare practices.

A. Limitations

This article is not without limitations. First, the use of a different search strategy—defining another search query, varying the time frame, or using another data source—might well have led to slightly different results.

Second, the breadth and depth of analysis are potentially limited by the inherent sparsity of data in news and by the fact that the availability of relevant news articles or data points may vary across different technological advancements or underrepresented niche areas.

Third, the accuracy and completeness of the information extracted from the news articles could be influenced by factors such as noise, bias, or unbalanced topic coverage per topic or location. These factors could introduce potential limitations in the generalization of the findings.

Finally, this study focused solely on English-language news articles, which may introduce a language bias and limit the inclusion of insights from non-English sources.

B. Future Lines of Research

In light of the findings presented in this study, several opportunities for future research could be explored. First, in the field of healthcare, it would be valuable to develop a mechanism for identifying incipient and latent trends and relationships between digital health and technology concepts by conducting a time-series forecasting analysis. This analysis could predict the evolution of concepts over time and uncover the potential for long-term bonding. This, in turn, would foster a more profound grasp of how technology catalyzes innovative advancements in healthcare. Furthermore, the technique of entity recognition and relation extraction could be expanded upon to construct contextual knowledge graphs, adding another layer of depth to the analysis.

Second, integrating data from reports by policymakers, consultancies, and global organizations at various levels (local, national, and international) could strengthen the news mining analysis. This would reveal more insights, offer a broader perspective, and alleviate current sparsity issues.

Third, considering the disruptive development of LLMs in the field of NLP, an opportunity for research lies in creating a multilingual pipeline that automates the entire process of news mining analysis. This pipeline would enable nontechnical users such as business managers to provide human-understandable instructions and automatically integrate data collection, storage, preprocessing, analysis, insights generation, and results presentation. It would also automatically generate the required code for analysis and ensure the explainability of the results. The pipeline would contain the constraints and security mechanisms required to avoid such limitations of these technologies as hallucinations, guaranteeing the reliability of the results.

Finally, expanding and validating the proposed research with diverse news article datasets beyond healthcare could offer valuable insights. Exploring how technology facilitates innovation across sectors such as industries, businesses, entrepreneurship, sustainability, education, politics, and society will yield a more comprehensive understanding of its influence.

COMPETING INTERESTS

Enrique Cano-Marin is employed by Roche. There are no other potential conflicts of interest. The work presented herein was performed solely by the listed authors.

REFERENCES

[1] D. Aloini, E. Benevento, A. Stefanini, and P. Zerbino, "Transforming healthcare ecosystems through blockchain: Opportunities and capabilities for business process innovation," *Technovation*, vol. 119, 2023, Art. no. 102557.

[2] S. Gërguri-Rashiti, V. Ramadani, H. Abazi-Alili, L.-P. Dana, and V. Ratten, "ICT innovation and firm performance: The transition economies context," *Thunderbird Int. Bus. Rev.*, vol. 59, no. 1, pp. 93–102, 2017.

[3] I. Kulkov et al., "Technology entrepreneurship in healthcare: Challenges and opportunities for value creation," *J. Innov. Knowl.*, vol. 8, no. 2, 2023, Art. no. 100365. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2444569X23000616

[4] O. Ali, W. Abdelbaki, A. Shrestha, E. Elbasi, M. A. A. Alryalat, and Y. K. Dwivedi, "A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities," *J. Innov. Knowl.*, vol. 8, no. 1, 2023, Art. no. 100333. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2444569X2300029X

[5] L. Cannavacciuolo, G. Capaldo, and C. Ponsiglione, "Digital innovation and organizational changes in the healthcare sector: Multiple case studies of telemedicine project implementation," *Technovation*, vol. 120, 2023, Art. no. 102550.

[6] G. Manogaran, D. Lopez, C. Thota, K. M. Abbas, S. Pyne, and R. Sundarasekar, "Big Data analytics in healthcare Internet of Things," in *Innovative Healthcare Systems for the 21st Century*. Cham, Switzerland: Springer, pp. 263–284, 2017.

[7] S. Akter, K. Michael, M. R. Uddin, G. McCarthy, and M. Rahman, "Transforming business using digital innovations: The application of AI, blockchain, cloud and data analytics," *Ann. Operations Res.*, vol. 308, pp. 1–33, 2022.

[8] C. Bai, P. Dallasega, G. Orzes, and J. Sarkis, "Industry 4.0 technologies assessment: A sustainability perspective," *Int. J. Prod. Econ.*, vol. 229, 2020, Art. no. 107776.

[9] M. Karatas, L. Eriskin, M. Deveci, D. Pamucar, and H. Garg, "Big Data for healthcare Industry 4.0: Applications, challenges and future perspectives," *Expert Syst. Appl.*, vol. 200, 2022, Art. no. 116912.

[10] A. Bertello, A. Ferraris, S. Bresciani, and P. De Bernardi, "Big Data analytics (BDA) and degree of internationalization: The interplay between governance of BDA infrastructure and BDA capabilities," *J. Manage. Governance*, vol. 25, pp. 1035–1055, 2021.

[11] Statista, "Level of unstructured health data in the US," 2022. Accessed: Jun. 10, 2023. [Online]. Available: https://www.statista.com/statistics/1377521/level-of-unstructured-health-data-in-the-us/

[12] B. Gärtner and M. R. Hiebl, "Issues with Big Data," in *The Routledge Companion to Accounting Information Systems*. Evanston, IL, USA: Routledge, 2017, pp. 161–172.

[13] S. Kraus, F. Schiavone, A. Pluzhnikova, and A. C. Invernizzi, "Digital transformation in healthcare: Analyzing the current state-of-research," *J. Bus. Res.*, vol. 123, pp. 557–567, 2021.

[14] M.-F. Hsu, T.-M. Chang, and S.-J. Lin, "News-based soft information as a corporate competitive advantage," *Technol. Econ. Develop. Economy*, vol. 26, no. 1, pp. 48–70, 2020.

[15] T.-M. Chang, M.-F. Hsu, and S.-J. Lin, "Integrated news mining technique and AI-based mechanism for corporate performance forecasting," *Inf. Sci.*, vol. 424, pp. 273–286, 2018.

[16] M. N. Ashtiani and B. Raahmei, "News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review," *Expert Syst. Appl.*, vol. 217, 2023, Art. no. 119509.

[17] P.-Y. Hao, C.-F. Kung, C.-Y. Chang, and J.-B. Ou, "Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane," *Appl. Soft Comput.*, vol. 98, 2021, Art. no. 106806.

[18] X. Li, Q. Xie, and L. Huang, "Identifying the development trends of emerging technologies using patent analysis and web news data mining: The case of perovskite solar cell technology," *IEEE Trans. Eng. Manage.*, vol. 69, no. 6, pp. 2603–2618, Dec. 2022.

[19] K. Buckingham, J. Brandt, W. Anderson, L. F. do Amaral, and R. Singh, "The untapped potential of mining news media events for understanding environmental change," *Curr. Opin. Environ. Sustainability*, vol. 45, pp. 92–99, 2020.

[20] E. Cano-Marin, M. Mora-Cantallops, and S. Sánchez-Alonso, "Twitter as a predictive system: A systematic literature review," *J. Bus. Res.*, vol. 157, 2023, Art. no. 113561.

[21] K. Krawczyk et al., "Quantifying online news media coverage of the COVID-19 pandemic: Text mining study and resource," *J. Med. Internet Res.*, vol. 23, no. 6, 2021, Art. no. e28253.

[22] N. Sadman, N. Anjum, K. D. Gupta, and M. P. Mahmud, "Understanding the pandemic through mining COVID news using natural language processing," in *Proc. IEEE 11th Annu. Comput. Commun. Workshop Conf.*, 2021, pp. 0362–0367.

[23] K. Jahanbin et al., "Using Twitter and web news mining to predict the Monkeypox outbreak," *Asian Pacific J. Trop. Med.*, vol. 15, no. 5, 2022, Art. no. 236.

[24] V. Ramadani, A. Istrefi-Jahja, J. Zeqiri, and D. Ribeiro-Soriano, "COVID-19 and SMEs digital transformation," *IEEE Trans. Eng. Manage.*, vol. 70, no. 8, pp. 2864–2873, Aug. 2023.

[25] K. Leetaru and P. A. Schrodt, "GDELT: Global data on events, location, and tone, 1979–2012," *in Proc. ISA Annu. Conv.*, vol. 2, no. 4, 2013, pp. 1–49.

[26] G. E. Iyawa, M. Herselman, and A. Botha, "Digital health innovation ecosystems: From systematic literature review to conceptual framework," *Procedia Comput. Sci.*, vol. 100, pp. 244–252, 2016.

[27] S. C. Mathews, M. J. McShea, C. L. Hanley, A. Ravitz, A. B. Labrique, and A. B. Cohen, "Digital health: A path to validation," *NPJ Digit. Med.*, vol. 2, no. 1, 2019, Art. no. 38.

[28] J. Cancela, I. Charlafti, S. Colloud, and C. Wu, *Digital Health in the Era of Personalized Healthcare: Opportunities and Challenges for Bringing Research and Patient Care to a New Level*. San Diego, CA, USA: Digit. Health, 2021, pp. 7–31.

[29] V. Ramadani, H. Abazi-Alili, L.-P. Dana, G. Rexhepi, and S. Ibraimi, "The impact of knowledge spillovers and innovation on firm-performance: Findings from the Balkans countries," *Int. Entrepreneurship Manage. J.*, vol. 13, pp. 299–325, 2017.

[30] T. L. Lewis and J. C. Wyatt, "mHealth and mobile medical apps: A framework to assess risk and promote safer use," *J. Med. Internet Res.*, vol. 16, no. 9, 2014, Art. no. e3133.

[31] M. Paul, L. Maglaras, M. A. Ferrag, and I. AlMomani, *Digitization of Healthcare Sector: A Study on Privacy and Security Concerns*, Tamworth, U.K.: ICT Exp., 2023.

[32] S. Canali, V. Schiaffonati, and A. Aliverti, "Challenges and recommendations for wearable devices in digital health: Data quality, interoperability, health equity, fairness," *PLoS Digit. Health*, vol. 1, no. 10, 2022, Art. no. e0000104.

[33] N. Kasoju et al., "Digital health: Trends, opportunities and challenges in medical devices, pharma and bio-technology," *CSI Trans. ICT*, vol. 11, pp. 1–20, 2023.

[34] S. Chakraborty, V. Bhatt, and T. Chakravorty, "Big-Data, IoT wearable and mHealth cloud platform integration triads-a logical way to patient-health monitoring," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 3, pp. 388–394, 2020.

[35] P. Galetsi, K. Katsaliaki, and S. Kumar, "Assessing technology innovation of mobile health apps for medical care providers," *IEEE Trans. Eng. Manage.*, vol. 70, no. 8, pp. 2809–2826, Aug. 2023.

[36] B. Munos et al., "Mobile health: The power of wearables, sensors, and apps to transform clinical trials," *Ann. New York Acad. Sci.*, vol. 1375, no. 1, pp. 3–18, 2016.

[37] P. M. Kumar, S. Lokesh, R. Varatharajan, G. C. Babu, and P. Parthasarathy, "Cloud and IoT based disease prediction and diagnosis system for healthcare using fuzzy neural classifier," *Future Gener. Comput. Syst.*, vol. 86, pp. 527–534, 2018.

[38] A. Barragán-Montero et al., "Artificial intelligence and machine learning for medical imaging: A technology review," *Physica Medica*, vol. 83, pp. 242–256, 2021.

[39] S. Ye et al., "Telemedicine expansion during the COVID-19 pandemic and the potential for technology-driven disparities," *J. Gen. Intern. Med.*, vol. 36, pp. 256–258, 2021.

[40] J. S. Alpert et al., "The electronic medical record in 2016: Advantages and disadvantages," *Digit. Med.*, vol. 2, no. 2, 2016, Art. no. 48.

[41] R. Jabbar, N. Fetais, M. Krichen, and K. Barkaoui, "Blockchain technology for healthcare: Enhancing shared electronic health record interoperability and integrity," in *Proc. IEEE Int. Conf. Inform. IoT Enabling Technol.*, 2020, pp. 310–317.

[42] M. Muhos, M. Saarela, D. Foit, and L. Rasochova, "Management priorities of digital health service start-ups in California," *Int. Entrepreneurship Manage. J.*, vol. 15, pp. 43–62, 2019.
[43] A. Sigov, L. Ratkin, L. A. Ivanov, and L. D. Xu, "Emerging enabling technologies for Industry 4.0 and beyond," *Inf. Syst. Front.*, vol. 52, pp. 1–11, 2022.
[44] P. P. Jayaraman, A. R. M. Forkan, A. Morshed, P. D. Haghighi, and Y.-B. Kang, "Healthcare 4.0: A review of frontiers in digital health," *Wiley Interdiscipl. Rev.: Data Mining Knowl. Discov.*, vol. 10, no. 2, 2020, Art. no. e1350.
[45] S. Ramzan, A. Aqdus, V. Ravi, D. Koundal, R. Amin, and M. A. Al Ghamdi, "Healthcare applications using blockchain technology: Motivations and challenges," *IEEE Trans. Eng. Manage.*, vol. 70, no. 8, pp. 2874–2890, Aug. 2023.
[46] B. B. Gupta, A. Gaurav, A. A. Albeshri, and D. Alsalman, "New paradigms of sustainable entrepreneurship in metaverse: A micro-level perspective," *Int. Entrepreneurship Manage. J.*, vol. 19, pp. 1449–1465, 2023.
[47] A. Garavand and N. Aslani, "Metaverse phenomenon and its impact on health: A scoping review," *Informat. Med. Unlocked*, vol. 32, 2022, Art. no. 101029.
[48] J. A. Iglesias, A. Tiemblo, A. Ledezma, and A. Sanchis, "Web news mining in an evolving framework," *Inf. Fusion*, vol. 28, pp. 90–98, 2016.
[49] B. Batrinca and P. C. Treleaven, "Social media analytics: A survey of techniques, tools and platforms," *AI Soc.*, vol. 30, pp. 89–116, 2015.
[50] S. Bhadani, I. Verma, and L. Dey, "Mining financial risk events from news and assessing their impact on stocks," in *Proc. Workshop Mining Data Financial Appl.*, 2020, pp. 85–100.
[51] X. Zhou et al., "Identifying and assessing innovation pathways for emerging technologies: A hybrid approach based on text mining and altmetrics," *IEEE Trans. Eng. Manage.*, vol. 68, no. 5, pp. 1360–1371, Oct. 2021.
[52] M.-F. de Lima-Santos and W. Ceron, "Artificial intelligence in news media: Current perceptions and future outlook," *Journalism Media*, vol. 3, no. 1, pp. 13–26, 2021.
[53] I. Vayansky and S. A. Kumar, "A review of topic modeling methods," *Inf. Syst.*, vol. 94, 2020, Art. no. 101582.
[54] T. R. Rao, P. Mitra, R. Bhatt, and A. Goswami, "The Big Data system, components, tools, and technologies: A survey," *Knowl. Inf. Syst.*, vol. 60, pp. 1165–1245, 2019.
[55] P. Bachmann, M. Eisenegger, and D. Ingenhoff, "Defining and measuring news media quality: Comparing the content perspective and the audience perspective," *Int. J. Press/Politics*, vol. 27, no. 1, pp. 9–37, 2022.
[56] S. Choi, H. Shin, and S.-S. Kang, "Predicting audience-rated news quality: Using survey, text mining, and neural network methods," *Digit. Journalism*, vol. 9, no. 1, pp. 84–105, 2021.
[57] A. L. Opdahl et al., "Trustworthy journalism through AI," *Data Knowl. Eng.*, vol. 146, 2023, Art. no. 102182.
[58] H. D. Wu, "Systemic determinants of international news coverage: A comparison of 38 countries," *J. Commun.*, vol. 50, no. 2, pp. 110–130, 2000.
[59] M. Zolnoori et al., "Mining news media for understanding public health concerns," *J. Clin. Transl. Sci.*, vol. 5, no. 1, 2021, Art. no. e1.
[60] N. T. Viet and V. Gneushev, "Analyzing and forecasting emerging technology trends by mining web news," in *Proc. Conf. Creativity Intell. Technol. Data Sci.*, 2021, pp. 55–69.
[61] Z. Ma, G. Pant, and O. R. Sheng, "Mining competitor relationships from online news: A network-based approach," *Electron. Commerce Res. Appl.*, vol. 10, no. 4, pp. 418–427, 2011.
[62] R. Murali, S. Patnaik, and S. Cranefield, "Mining international political norms from the GDELT database," in *Coordination, Organizations, Institutions, Norms, and Ethics for Governance of Multi-Agent Systems XIII: International Workshops COIN 2017 and COINE 2020*. Berlin, Germany: Springer, 2021, pp. 35–56.
[63] P. Nanda and V. Kumar, "Decision analytics for competitive advantage: Cases on using social media analytics," *Int. J. Technol., Policy, Manage.*, vol. 23, no. 4, pp. 372–386, 2023.
[64] S.-J. Lin and M.-F. Hsu, "Decision making by extracting soft information from CSR news report," *Technological Econ. Develop. Economy*, vol. 24, no. 4, pp. 1344–1361, 2018.
[65] M. Bertl, "News analysis for the detection of cyber security issues in digital healthcare: A text mining approach to uncover actors, attack methods and technologies for cyber defense," *Young Inf. Scientist*, vol. 4, pp. 1–15, 2019.
[66] O. David Williams, K. C. Yung, and K. A. Grépin, "The failure of private health services: COVID-19 induced crises in low-and middle-income country (LMIC) health systems," *Glob. Public Health*, vol. 16, no. 8/9, pp. 1320–1333, 2021.
[67] S. Cossin et al., "Public health and epidemiology informatics: Recent research trends moving toward public health data science," *Yearbook Med. Inform.*, vol. 29, no. 1, pp. 231–234, 2020.
[68] Z. Nasar, S. W. Jaffry, and M. K. Malik, "Named entity recognition and relation extraction: State-of-the-art," *ACM Comput. Surv.*, vol. 54, no. 1, pp. 1–39, 2021.
[69] V. Ng, E. E. Rees, J. Niu, A. Zaghool, H. Ghiasbeglou, and A. Verster, "Application of natural language processing algorithms for extracting information from news articles in event-based surveillance," *Canada Communicable Dis. Rep.= Releve des Mal. Transmissibles au Canada*, vol. 46, no. 6, pp. 186–191, 2020.
[70] C. Zheng, "Comparisons of the city brand influence of global cities: Word-embedding based semantic mining and clustering analysis on the Big Data of GDELT global news knowledge graph," *Sustainability*, vol. 12, no. 16, 2020, Art. no. 6294.
[71] D. Chauhan and K. Bansal, "Using the advantages of NOSQL: A case study on MongoDB," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 5, no. 2, pp. 90–93, 2017.
[72] S. Vijayarani et al., "Preprocessing techniques for text mining-an overview," *Int. J. Comput. Sci. Commun. Netw.*, vol. 5, no. 1, pp. 7–16, 2015.
[73] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA, USA: Sage, 2018.
[74] S. Smith, "Determining sample size: How to ensure you get the correct sample size," *Qualtrics*, 2013. Accessed: Apr. 15, 2015. [Online]. Available: http://www.qualtrics.Com/blog/determining-sample-size/
[75] M. Marrero, J. Urbano, S. Sánchez-Cuadrado, J. Morato, and J. M. Gómez-Berbís, "Named entity recognition: Fallacies, challenges and opportunities," *Comput. Standards Interfaces*, vol. 35, no. 5, pp. 482–489, 2013.
[76] B. Hachey, W. Radford, J. Nothman, M. Honnibal, and J. R. Curran, "Evaluating entity linking with Wikipedia," *Artif. Intell.*, vol. 194, pp. 130–150, 2013.
[77] K. Abhishek, V. Pratihar, S. K. Shandilya, S. Tiwari, V. K. Ranjan, and S. Tripathi, "An intelligent approach for mining knowledge graphs of online news," *Int. J. Comput. Appl.*, vol. 44, no. 9, pp. 838–846, 2022.
[78] J. Lehmann et al., "DBpedia—A large-scale, multilingual knowledge base extracted from Wikipedia," *Semantic Web*, vol. 6, no. 2, pp. 167–195, 2015.
[79] P. N. Mendes, M. Jakob, A. García-Silva, and C. Bizer, "DBpedia spotlight: Shedding light on the web of documents," in *Proc. 7th Int. Conf. Semantic Syst.*, 2011, pp. 1–8.
[80] Z. Geng, Y. Zhang, and Y. Han, "Joint entity and relation extraction model based on rich semantics," *Neurocomputing*, vol. 429, pp. 132–140, 2021.
[81] S. Grayson, K. Wade, G. Meaney, and D. Greene, "The sense and sensibility of different sliding windows in constructing co-occurrence networks from literature," in *Proc. Comput. Hist. Data-Driven Humanities: 2nd IFIP WG 12.7 Int. Workshop*, 2016, pp. 65–77.
[82] E. Sulis, L. Humphreys, F. Vernero, I. A. Amantea, D. Audrito, and L. Di Caro, "Exploiting co-occurrence networks for classification of implicit inter-relationships in legal texts," *Inf. Syst.*, vol. 106, 2022, Art. no. 101821.
[83] D. Paranyushkin, "InfraNodus: Generating insight using text network analysis," in *Proc. World Wide Web Conf.*, 2019, pp. 3584–3589.
[84] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech.: Theory Exp.*, vol. 2008, no. 10, 2008, Art. no. P10008.
[85] P. Bonacich, "Some unique properties of eigenvector centrality," *Social Netw.*, vol. 29, no. 4, pp. 555–564, 2007.
[86] M. Z. Al-Taie and S. Kadry, *Python for Graph and Network Analysis*. Berlin, Germany: Springer, 2017.
[87] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proc. Int. AAAI Conf. Web Social Media*, 2009, pp. 361–362.
[88] M. Haider and A. Gandomi, "When Big Data made the headlines: Mining the text of Big Data coverage in the news media," *Int. J. Serv. Technol. Manage.*, vol. 27, no. 1/2, pp. 23–50, 2021.
[89] R. Benson, "Paywalls and public knowledge: How can journalism provide quality news for everyone?," *Journalism*, vol. 20, no. 1, pp. 146–149, 2019.

[90] G. Joris, J. Hendrickx, S. Vercoutere, O. De Clercq, and L. De Marez, "Paying for news diversity? A topic diversity analysis of free and paywalled online news," *Journalism Pract.*, pp. 1–20, 2023, doi: 10.1080/17512786.2023.2191863.

[91] X. Wu, C. Li, Y. Zhu, and Y. Miao, "Short text topic modeling with topic distribution quantization and negative sampling decoder," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2020, pp. 1772–1782.

[92] W. Lin, X. Wu, Z. Wang, X. Wan, and H. Li, "Topic network analysis based on co-occurrence time series clustering," *Mathematics*, vol. 10, no. 16, 2022, Art. no. 2846.

[93] A. K. Jayaraman and S. Abirami, "Opinion-based co-occurrence network for identifying the most influential product features," *J. Eng. Res.*, vol. 8, no. 4, 2020, doi: 10.36909/jer.v8i4.8369.

[94] M. Lipsitch, D. L. Swerdlow, and L. Finelli, "Defining the epidemiology of COVID-19—Studies needed," *New England J. Med.*, vol. 382, no. 13, pp. 1194–1196, 2020.

[95] E. Monaghesh and A. Hajizadeh, "The role of telehealth during COVID-19 outbreak: A systematic review based on current evidence," *BMC Public Health*, vol. 20, pp. 1–9, 2020.

[96] F. Tagscherer and C.-C. Carbon, "Leadership for successful digitalization: A literature review on companies' internal and external aspects of digitalization," *Sustain. Technol. Entrepreneurship*, vol. 2, no. 2, 2023, Art. no. 100039. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2773032823000032

[97] O. Oren, B. J. Gersh, and D. L. Bhatt, "Artificial intelligence in medical imaging: Switching from radiographic pathological data to clinically meaningful endpoints," *Lancet Digit. Health*, vol. 2, no. 9, pp. e486–e488, 2020.

[98] J. M. G. Martínez, P. Carracedo, D. Gorgues Comas, and C. H. Siemens, "An analysis of the blockchain and COVID-19 research landscape using a bibliometric study," *Sustain. Technol. Entrepreneurship*, vol. 1, no. 1, 2022, Art. no. 100006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2773032822000062

[99] H. Zhou and S. B. Bhaduri, "3D printing in the research and development of medical devices," in *Biomaterials in Translational Medicine*. New York, NY, USA: Elsevier, 2019, pp. 269–289.

[100] D. C. Klonoff, "Fog computing and edge computing architectures for processing data from diabetes devices connected to the medical Internet of Things," *J. Diabetes Sci. Technol.*, vol. 11, no. 4, pp. 647–652, 2017.

[101] U. S. Food and Drug Administration, "Software as a medical device (SAMD)," 2018. Accessed: Jun. 07, 2023. [Online]. Available: https://www.fda.gov/medical-devices/digital-health/software-medical-device-samd

[102] S. Gerke, B. Babic, T. Evgeniou, and I. G. Cohen, "The need for a system view to regulate artificial intelligence/machine learning-based software as medical device," *npj Digit. Med.*, vol. 3, no. 1, 2020, Art. no. 53.

[103] S. H. Almotiri, M. A. Khan, and M. A. Alghamdi, "Mobile health (m-Health) system in the context of IoT," in *Proc. IEEE 4th Int. Conf. Future Internet Things Cloud Workshops*, 2016, pp. 39–42.

[104] D. Bourgeois, J. Rappaz, and K. Aberer, "Selection bias in news coverage: Learning it, fighting it," in *Proc. Companion Web Conf. Companion*, 2018, pp. 535–543.

[105] Y. Juhn and H. Liu, "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research," *J. Allergy Clin. Immunol.*, vol. 145, no. 2, pp. 463–469, 2020.

[106] S. J. MacEachern and N. D. Forkert, "Machine learning for precision medicine," *Genome*, vol. 64, no. 4, pp. 416–425, 2021.

[107] G. Varoquaux and V. Cheplygina, "Machine learning for medical imaging: Methodological failures and recommendations for the future," *npj Digit. Med.*, vol. 5, no. 1, 2022, Art. no. 48.

[108] P. Aithal, A. Aithal, and E. Dias, "Blockchain technology-current status and future research opportunities in various areas of healthcare industry," *Int. J. Health Sci. Pharm.*, vol. 5, no. 1, pp. 130–150, 2021.

[109] R. W. Ahmad, K. Salah, R. Jayaraman, I. Yaqoob, S. Ellahham, and M. Omar, "The role of blockchain technology in telehealth and telemedicine," *Int. J. Med. Inform.*, vol. 148, 2021, Art. no. 104399.

[110] G. Aceto, V. Persico, and A. Pescapé, "Industry 4.0 and health: Internet of Things, Big Data, and cloud computing for Healthcare 4.0," *J. Ind. Inf. Integration*, vol. 18, 2020, Art. no. 100129.

[111] P. D. Kaur and I. Chana, "Cloud based intelligent system for delivering health care as a service," *Comput. Methods Programs Biomed.*, vol. 113, no. 1, pp. 346–359, 2014.

[112] S. Nazir, Y. Ali, N. Ullah, and I. García-Magariño, "Internet of Things for healthcare using effects of mobile computing: A systematic literature review," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–20, 2019.

[113] K. Ullah, M. A. Shah, and S. Zhang, "Effective ways to use Internet of Things in the field of medical and smart health care," in *Proc. IEEE Int. Conf. Intell. Syst. Eng.*, 2016, pp. 372–379.

[114] Y. K. Dwivedi et al., "Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy," *Int. J. Inf. Manage.*, vol. 57, 2021, Art. no. 101994.

[115] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, and T. Blaschke, "The rise of deep learning in drug discovery," *Drug Discov. Today*, vol. 23, no. 6, pp. 1241–1250, 2018.

[116] E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," *Nature Med.*, vol. 25, no. 1, pp. 44–56, 2019.

[117] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. K. Ambasta, and P. Kumar, "Artificial intelligence to deep learning: Machine intelligence approach for drug discovery," *Mol. Diversity*, vol. 25, pp. 1315–1360, 2021.

[118] J.-P. Vert, "How will generative AI disrupt data science in drug discovery ?," *Nature Biotechnol.*, vol. 41, pp. 750–751, 2023.

[119] S. Kolluri, J. Lin, R. Liu, Y. Zhang, and W. Zhang, "Machine learning and artificial intelligence in pharmaceutical research and development: A review," *AAPS J.*, vol. 24, pp. 1–10, 2022.

[120] A. Shaban-Nejad, M. Michalowski, and D. L. Buckeridge, "Health intelligence: How artificial intelligence transforms population and personalized health," *npj Digit. Med.*, vol. 1, no. 1, 2018, Art. no. 53.

[121] K. Azbeg, O. Ouchetto, S. Andaloussi, and L. Fetjah, "A taxonomic review of the use of IoT and blockchain in healthcare applications," *IRBM*, vol. 43, no. 5, pp. 511–519, 2022.

[122] M. Fiore et al., "Blockchain for the healthcare supply chain: A systematic literature review," *Appl. Sci.*, vol. 13, no. 2, 2023, Art. no. 686.

[123] R. Saranya and A. Murugan, "A systematic review of enabling blockchain in healthcare system: Analysis, current status, challenges and future direction," *Mater. Today: Proc.*, vol. 80, pp. 3010–3015, 2023.

[124] B. Murdoch, "Privacy and artificial intelligence: Challenges for protecting health information in a new era," *BMC Med. Ethics*, vol. 22, no. 1, pp. 1–5, 2021.

Enrique Cano-Marin received the M.Sc. degree in robotics and automation from the University Carlos III of Madrid, Getafe, Spain, in 2016. He is currently working toward the Ph.D. degree in information and knowledge engineering with the Department of Computer Science, University of Alcalá, Alcalá de Henares, Spain.

He is currently an Industrial Engineer. He teaches as an Assistant Lecturer with the University of Alcalá. In addition to his academic pursuits, he is currently a Global IT Product Manager with a multinational pharmaceutical company. He has authored or coauthored several papers in high-impact international journals, in the fields of artificial intelligence, natural language processing, graph analytics, social media analytics, social network analysis, and information retrieval.

Salvador Sánchez-Alonso received the Ph.D. degree in computer science from the Universidad Politcnica de Madrid, Madrid, Spain, in 2005, and the graduate degree in documentation management from the University of Alcala, Madrid, in 2011.

He is currently an Associate Professor and a Senior Member of the Information Engineering Group, a research unit of the Computer Science Department, University of Alcalá, Alcalá de Henares, Spain. He was a Software Engineer with a software solutions company. With many high-impact factor publications in the last ten years, his current research interests include social network analysis applications, web science, and blockchain applications.

Marçal Mora-Cantallops received the Ph.D. degree in computer science from the University of Alcala, Alcalá de Henares, Spain, in 2018.

He is currently teaching and researching with the University of Alcalá, Alcalá de Henares, Spain. He is also an Industrial and Computer Engineer. Since 2018, he has been an Assistant Professor with the Computer Science Department, University of Alcalá, and also a member of the I4 (intelligence, innovation, internet, and information) Research Group. His research interests include machine learning, social network analysis, data science, ethical AI, and game studies, having authored and published multiple related articles in the past few years.