BLOCKCHAIN: A VIEW FROM THE PROTECTION OF SENSITIVE DATA IN THE HEALTH SECTOR.

BLOCKCHAIN: UNA MIRADA DESDE LA PROTECCIÓN DE DATOS SENSIBLES EN EL SECTOR SALUD


*Universidad Nacional Abierta y a Distancia UNAD de Colombia, Escuela de Ciencias Básicas, Tecnología e Ingeniería. Calle 14 # 28 - 45, Pasto, Nariño, Colombia. (1)(3443700)
E-mail: {edgar.dulce, james.betancourt, francisco.solarte, carlos.rosero}@unad.edu.co

Resumen: Se seleccionaron las bases de datos bibliográficas ACM Digital Library, Scopus, IEEE Explore y Web of Science, que sean compatibles con la disciplina de estudio. Entre los principales hallazgos se encuentran diferentes tipos de Blockchain que se aplican de muchas maneras para respaldar positivamente aspectos de las soluciones de atención médica, además, se están estudiando múltiples áreas de la atención médica, por ejemplo, soluciones para la gestión del entrenamiento físico y la nutrición, seguimiento de pacientes, predicción de futuras enfermedades, aseguramiento de información sensible en salud, manejo de Historia Clínica Electrónica EHR y Registro Personal de Salud PHR, identidad digital farmacéutica, control de grandes poblaciones, Covid-19, entre otros. Por otro lado, se identifican múltiples limitaciones o retos a abordar sobre la tecnología Blockchain en el entorno sanitario, tales como: falta de marcos legales y regulatorios, adopción de estándares a nivel nacional e internacional, tamaño de los datos sanitarios, problemas de interoperabilidad, sus mecanismos de consenso, falta de casos prácticos y experimentos con casos reales, escalabilidad y rendimiento, entre otros.

Palabras clave: Blockchain, Estudio de Caso, Salud, EHR, PHR, Seguridad

Abstract: The bibliographic databases ACM Digital Library, Scopus, IEEE Explore and Web of Science, which are compatible with the discipline of study, were selected. Among the main findings are different types of Blockchain being applied in many ways to positively support aspects of healthcare solutions, in addition, multiple areas of healthcare are being studied, for example, solutions for the management of physical training and nutrition, patient tracking, prediction of future diseases, assurance of sensitive health information, management of Electronic Medical Records EHR and Personal Health Records PHR, pharmaceutical digital identity, controlling large populations, Covid-19, among others. On the other hand, multiple limitations or challenges to be addressed about Blockchain technology in the healthcare environment are identified, such as: lack of legal and regulatory frameworks, adoption of standards at national and international level, size of healthcare data, interoperability problems, its consensus mechanisms, lack of practical cases and experiments with real cases, scalability and performance, among others.

Keywords: Blockchain, Case Study, Health, EHR, PHR, Safety
1. INTRODUCTION

In today's globalized world, it is a necessity for everyone to have access to quality health services (diagnosis, treatment and prevention) in an efficient, safe and transparent manner. At present, the percentage of universal health coverage is 50% (O. M. of Health, 2021). For this purpose, technologies are being developed every day to increase the coverage and quality of hospital services, and without them, medical centers would be inefficient and lose credibility (Aggelidis & Chatzoglou, 2009). In (Moen et al., 2013) explores some of the challenges involved in the adoption of healthcare technologies, such as storage, consultation and transmission of associated data. Also, (Lin et al., 2018) the implementation of health-oriented technology is associated with a decrease in morbidity and mortality.

In recent years, it has been shown that the exchange of health information between medical institutions benefits the medical industry, leading to better patient care and service. In this process, today there are multiple and varied problems, derived from the uncertainty of having pertinent, complete, available and reliable information.

The healthcare field is a topic that is researched daily from different approaches, one of them, is the relationship that exists between Blockchain technology and the management of Electronic Health Records (EHR) and Personal Health Records (PHR), improving the availability, traceability, confidentiality, and integrity of information. Blockchain was first introduced by Satoshi Nakamoto in an article on Bitcoin. (Nakamoto, n.d.).

Blockchain applications have been studied in financial environments (where it started) as well as in other growing areas of TIC. The World Economic Forum in 2015 (WEF, 2015), published that by 2025, 10% of the world's gross domestic product will be stored using Blockchain. Now, considered a mainstream technology, used in different industries and use cases, such as identity management, contracts, supply chain, insurance, health care, voting, among others (C. Burniske, E. Vaughn, J. Shelton, 2016).

In brief, blockchain consists of a technology capable of building an open and distributed database (Yang et al., 2019). At the heart of Blockchain technology is a distributed ledger that is immutably maintained collaboratively (Yang et al., 2019), this ledger is composed of interlocking data structures called blocks (hence its name blockchain), ordered chronologically and correlatively (Huang et al., 2019). Each block contains a timestamp, a hash pointing to the previous block and the data of each transaction (which, if put in context, would be the data concerning health) (Esposito et al., 2018). The generation or obtaining of new blocks to build the chain is achieved through consensus mechanisms (Agbo et al., 2019). The information cannot be reproduced without the participation of all its nodes and can be transmitted confidentially without a trusted third party (Janssens et al., 2020). The technological component in Blockchain is quite disruptive (Mettler, 2016), it can be shaped in different ways by the actors and the needs of each sector where it is implemented (Yang et al., 2019).

Benefits of Blockchain in the Health Applications Ecosystem. Conceptually Blockchain is secure, providing decentralized confidentiality, integrity, consistency and resistance to intentional or unintentional attacks. The benefits of Blockchain in a healthcare ecosystem are (Esposito et al., 2018): i) It generates secure transmission without a trusted third party, generating privacy. ii) Patients have control over their data. iii) Data such as medical records over Blockchain are complete, consistent, timely, accurate and easy to distribute. iv) Visible changes and immutable insertions for all members of the patient's network. In addition, any unauthorized modification is easy to detect. v) Ensures privacy and anonymity in their transactions (Peterson et al., 2017).

In this article, we study how Blockchain technology is being applied within case studies oriented to people's health, through the Systematic Mapping method proposed by Petersen et al (K. Petersen, R. Feldt, S. Mujtaba, 2008). The results show that there is an increasing amount of research on solutions, technologies and areas of health that are supported by Blockchain, oriented to improve people's health.

2. METHODOLOGY

Description of the Research Protocol. This study uses the methodology of Petersen et al (K. Petersen, R. Feldt, S. Mujtaba, 2008) for the
realization of the SMS. The objective of the study is to explore how Blockchain technology is supporting solutions provided in the healthcare sector. The SMS allowed to identify and classify research topics in Blockchain technology solutions related to people's health. It is composed of methodical steps to search, interpret, synthesize and analyze the information present in published articles related to the topic of study (O. Revelo, C. Collazos, 2018). The use of this technique aims to provide an overview of the area of interest and minimize the chances of error during the review process, it also allows to have a better control of the review activity and eliminate possible errors that may lead to erroneous or inaccurate conclusions. In addition, different facets of the scheme can be combined to answer more specific research questions. An SMS is divided into several sequential steps viz: Defining research questions, Execution of the search, Selection of relevant articles, Keyword search, and Mapping process and Data extraction. As can be seen in the following illustration:

*Figure 1. The systematic mapping process (K. Petersen, R. Feldt, S. Mujtaba, 2008)*

3. APPLICATION OF LA METHODOLOGY OF PETERSEN

**Research questions.** The first step in an SMS is to define the research questions RQs to be answered by the study. The following RQs were defined:

RQ1: What type of Blockchain (private, public, other) do the authors use to support sensitive information management in healthcare?
RQ2: What are the areas of healthcare where Blockchain has been applied?
RQ3: What are the emerging limitations and challenges that the literature raises for this area of research?
RQ4: What countries are they from and in what year was the research related to the object of study conducted?

**Conducting the search.** The second step is to search and collect all related research papers based on a specific search term. The English language terms "(Blockchain OR "block chain") AND (health OR healthcare) AND ("case study" OR "case studies" OR "field study" OR experiments)) were chosen. This phrase relates Blockchain technology to the words Health and "Healthcare" and various synonyms for the term "case study".

Then, taking the analysis of the results of (T. Dyba, T. Dingsoyr, 2007) who provide a list of important DBs in the field of Computer Science and Engineering, four DBs were selected to perform the search: ACM Digital Library, Scopus, IEEE Explore and Web of Science. A bubble diagram was used to report the frequencies, shown in Figure 2, in which the total number of articles categorized by each of the databases can be seen.

*Figure 2. Visualization of the number of articles obtained, by type in each database.*

**Search for relevant articles.** Table 1 Shows the inclusion (I) and exclusion (E) criteria defined to add or remove relevant/irrelevant articles.

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>I1: Papers published in the last five (5) years (2017 - 2021).</td>
<td>1: Technical reports, abstracts, surveys (gray literature), secondary studies (SMS).</td>
</tr>
<tr>
<td>I2: If there are papers related to the same study, only the most recent one is selected.</td>
<td>E2: Articles in languages other than English or Spanish.</td>
</tr>
<tr>
<td>I3: If an article describes more than one study, each study is evaluated individually.</td>
<td>E3: Articles not related to Blockchain, health, case studies or synonyms.</td>
</tr>
<tr>
<td>I4: If there are short and complete versions of the same study, we select complete.</td>
<td>E4: Only Article (Journal), Conference Paper and Early Access Article type papers.</td>
</tr>
</tbody>
</table>
**Keyword Search.** The next step is to find relevant articles through abstracts and keywords. For this stage, the process defined by Petersen et al (K. Petersen, R. Feldt, S. Mujtaba, 2008) was used. Titles and abstracts (and sometimes the introduction and conclusions) were reviewed to identify those considered irrelevant to the topics studied. The entire contents of the articles were then carefully read, and the data contained in them were extracted, analyzed and categorized.

**Mapping and Data Mining Process.** The articles were analyzed and classified according to categories created to separate the research contributions of each article (Results section). The data extracted from the articles were stored and subjected to qualitative and quantitative analysis. This analysis aimed to find evidence to answer the research questions defined in the Conceiving the Search section. To organize the findings and document the data extraction process, a spreadsheet was used, which also allowed for further statistical analysis.

6. **RESULTS AND DISCUSSION**

The aim of this paper is to explore how Blockchain technology is supporting solutions provided in the healthcare sector, describing different initiatives and discussing their social implications and future perspectives. Figure 3 shows the process indicated in the section Description of the research protocol. Initially, 284 articles were obtained. Of these, 42 duplicates were identified, which were eliminated, leaving 242. Next, the inclusion and exclusion criteria (I1 + E1 + E2) were applied, including those published in the last five (5) years and excluding those that did not correspond to papers, congresses or book chapters, and those written in languages other than English or Spanish, leaving 205 articles. The inclusion and exclusion criteria (I2, I3, I4, E3, E4) were applied to the 205 articles, reading their title and abstract, identifying 87 papers. A complete reading of these 87 articles was carried out and all the inclusion and exclusion criteria were applied again, obtaining 21 articles that were used as evidence to answer the RQs. For the review and analysis of the 21 papers and to avoid bias on the part of the researcher, we had the support of an external expert on the topics investigated.

![Figure 3. Summary of the item filtering process.](image)

Figure 4 shows the number of articles by type of publication and DB. The number of Journal Articles (13 articles) are the most frequent, followed by Conference Papers (7 articles) and, finally, book chapters (1 article).

![Figure 4. Number of final articles organized by database and type.](image)

**RQ1: What type of Blockchain (private, public, other) do the authors use to support sensitive health information management?** Figure 5 shows a summary of the types of Blockchain used in each of the 21 articles analyzed. There is no bias for the use of one type or another, some authors mention that the selection depends on the situation and type of solution offered, for example, in some cases where public data is handled, a Blockchain of this type would be the best option, or in the case of projects that have different types of data protection, you can select between a private or authorized Blockchain and, if several entities are involved in the project, they can opt for consortium Blockchain.

![Diagram showing the process of filtering items.](image)

---

1 https://sites.google.com/unad.edu.co/blockchain-salud/inicio

**Universidad de Pamplona**

I. I. D. T. A.
RQ2: What are the areas of healthcare in which Blockchain has been applied? A classification scheme, studied by Köttertzsch and Weyers (Köttertzsch & Weyers, 2016), was used, which groups categories and subcharacteristics that help classify systems, challenges and future trends in research (Table 2). These categories were extracted from the final articles and gradually aggregated. The classification itself was used as a means of validation. For each category, at least one contribution could be assigned to each characteristic. Seven classification categories are proposed, as can be seen in the following table:

<table>
<thead>
<tr>
<th>Type of Service</th>
<th>Context</th>
<th>User</th>
<th>Information Technology</th>
<th>Adaptation</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring Intervention Support</td>
<td>Senior Users</td>
<td>Location Movement Activities Health User Information Communication Devices</td>
<td>Sensors Wireless sensors WNS networks Camera Login Radio frequency Laser</td>
<td>User based Automatic complete</td>
<td>Prototype Concept Complete</td>
</tr>
</tbody>
</table>

The most basic category, Contribution Type, includes "systems", "services" and "studies". Application Context covers a variety of topics grouped into 23 characteristics. Within this category, "intervention" includes care, prevention and rehabilitation measures; "Presence" includes virtual and real presence; "tools" includes frameworks, technologies or outcomes to support people. Applications differ in the User group, including "older adults", "care staff", "caregivers" or "aides". The systems require or collect Information about the User's "location", "activities", "health" or "information". These aspects can be transferred to the system by different Technologies, including motion or pressure "Sensors"; information recorded by "cameras" or "radio frequency technologies" such as RFID or NCF. Systems also use different means of Adaptation, "user based", "automatic" or "manual". Finally, the literature includes systems, studies and services in certain stages (Status), which can be grouped into "prototypes", "proof of concept" and "ongoing" or "completed" research.

To highlight some of the contributions (Yaqoob et al., 2021) (Bedin et al., 2021) (Shyru et al., 2021) (Yang et al., 2019) (Khurram & Sardar, 2020), focus on the management of sensitive health (EL Azaouiti et al., 2021) information. Focuses on the storage of the entire medical history of patients (Attaallah et al., 2021) (Zhang & Kuo, 2021) (Santos et al., 2021) works on privacy and protection of sensitive health data. Researches on smart and healthy cities. Patient identity and access control to patient data (Jamal et al., 2021) focuses on patient fitness and data-driven body mass index. (Rupasinghe et al., 2019) Investigates the topic of fall prevention. Studies the topic of diabetes and cardiovascular diseases (Azogu et al., 2019) handles interoperability and data integration issues. The pandemic unleashed worldwide related to Covid-19 has prompted studies concerning patient identification (Attaallah et al., 2021) and patient confidentiality (Mohsin et al., 2021).

Considering the Status category in Table 2, only 6 out of the 21 final papers develop prototypes and experiments (Shyru et al., 2021) (Khurram & Sardar, 2020) (Rajput et al., 2021) (Zhang & Kuo, 2021) (Jamil et al., 2021) (Meena et al., 2019), which shows the need for much more research and delving into functional proposals in the Blockchain field.

RQ3: What are the emerging limitations and challenges that the literature poses for this area of research?

Blockchain technology is at an early stage and requires addressing many considerations to implement it successfully. This knowledge can help researchers and practitioners better understand
the obstacles in the way of Blockchain implementation. In this section, several challenges/constraints have been summarized based on the different requirements of the healthcare sector.

**Lack of legal frameworks and standardization.** The lack of legal and regulatory frameworks is one of the main concerns, in addition to the lack of universal standardization for EHR or PHR sharing (Kötteritzsch & Weyers, 2016) (Sharma & Joshi, 2021) (Rajput et al., 2021) (Kouhizadeh et al., 2021) (Hussien et al., 2021) (Hosseini Bamakan et al., 2021).

**Lack of real use cases.** Within the literature there is a very large gap in the lack of real cases or close-to-reality experiment (Yaqoob et al., 2021) (Rajput et al., 2021) (Santos et al., 2021) (Azogu et al., 2019), only 7 of the 21 final papers mention contributions of these types.

**Security issues.** There are several problems in this regard: for example (Bedin et al., 2021) companies are often reluctant to cooperate because the disclosure of operational information may be seen as a threat rather than an opportunity for growth. (Kötteritzsch & Weyers, 2016) (Gupta et al., 2021) Cryptographic key management. (Kouhizadeh et al., 2021) (Hosseini Bamakan et al., 2021) The transparency of a Blockchain allows participants to see some data, leading to issues of confidentiality, synchronization and scalability. (Khurram & Sardar, 2020) Undesirable events such as DDoS attacks, 51% attacks, fraud, collusion, manipulation, and sabotage can occur. Patient identity theft (Khurram & Sardar, 2020).

**Interoperability.** (Yaqoob et al., 2021) (Kouhizadeh et al., 2021) (Santos et al., 2021) (Hussien et al., 2021) (Gupta et al., 2021) The healthcare ecosystem is very broad, the lack of interoperability between different healthcare providers and other entities is increasing, this is made more serious by the lack of adoption of universal standards for data processing and exchange.

**Size of healthcare data.** Blockchain was created to share financial data, which are relatively small, in the healthcare sector different types of data can be found, for example, a clinical history or a diagnostic image can occupy several GB on a storage medium (Hussien et al., 2021) (Hosseini Bamakan et al., 2021) (Gupta et al., 2021).

**Energy inefficiency.** In public networks using the proof-of-work consensus mechanism (Zhang & Kuo, 2021) (Gupta et al., 2021), there are problems of high energy consumption due to the processing required to add new blocks.

**Different consensus mechanisms.** The consensus mechanism makes it possible to add new blocks to the chain, every day new algorithms (Zubaydi et al., 2019) are created that enter the Blockchain world and increase the uncertainty of selection (Jamil et al., 2021) (Gupta et al., 2021).

**Controlling access to sensitive information.** One of the main concerns is that the patient should always have control of his data and, in this sense, that he should grant access to his data only to authorized personnel (Sun et al., 2020) (Khurram & Sardar, 2020) (Santos et al., 2021) (Hosseini Bamakan et al., 2021).

It is possible to evidence other challenges, such as: cost of Blockchain implementation (Hosseini Bamakan et al., 2021), lack of personnel specialized in Blockchain (Sharma & Joshi, 2021), performance (Jamil et al., 2021) (Santos et al., 2021), decentralized storage (Yaqoob et al., 2021) (Gupta et al., 2021), data manipulation (Khurram & Sardar, 2020), among others.

**RQ4: From which countries did the research related to the object of study come and in which year was it conducted?** 100% of the studies on the subject were published in the last three years (Figure 6 (a)). This gradually increasing trend indicates the growing importance and potential of the research topic. One third of the final papers (7 out of 21) were part of academic events (Figure 4), showing a positive result, since novel studies of general interest are published in these events. On the other hand, it can be seen in Figure 6(b) that the 21 final papers come from institutions represented in 11 countries, with participation from 5 continents, this indicates that Blockchain in relation to the health of people around the world.
DISCUSSION

RQ1. There is no inclination for a particular type of Blockchain, although the authorized Blockchain type is increasingly used, which has a mixture of two independent types, precisely to obtain the relevant advantages of each of them. In addition, it is observed in Figure 6, that practically the 4 most common types of Blockchain are being used. The authors conclude that the type of Blockchain depends directly on the context of use and the level of privacy required by the proposed solution. Each of the Blockchain types has its advantages and disadvantages, their main difference is the scope, access to transactions and the consensus mechanism they use. In (Shynu et al., 2021) (Khurram & Sardar, 2020) (Rupasinghe et al., 2019), a combination of a private blockchain is used to store healthcare data and a consortium blockchain is used to store the pointers to that information.

RQ2. The results show that there are many characteristics and categories in which Blockchain is being taken into account in relation to health. Table 2 shows 7 categories in which the different types of solutions oriented to provide services that support people’s health can be framed, these are: type, context, user, information, technology, adaptation and state. The articles were analyzed, depending on the contributions in each category, ranging from support of physical or cognitive activities and help in coping with impairments and assistive support (Jamil et al., 2021) (Rupasinghe et al., 2019). Prediction of future disease (EL Azzaoui et al., 2021). Most approaches focus on accident management through monitoring (Rajput et al., 2021). Pharmaceutical cold chain assurance (Hosseini Bamakan et al., 2021). Large number of monitoring systems leads to increased use of sensors (Gupta et al., 2021), safer use of electronic healthcare devices (Zhang & Kuo, 2021). Ambient Assisted Living (AAL) (Sharma & Joshi, 2021), chronic disease monitoring through telemedicine (Shynu et al., 2021) (Khurram & Sardar, 2020). Tele-surgery (Gupta et al., 2021). Another important topic is digitally supported physical or cognitive therapy (Jamil et al., 2021) to combat malnutrition and diabetes. Although some of the contribution’s present prototypes (5), some individual support products were already on the market (1) and some had not yet passed a conceptual phase (15).

RQ3. It is not foreign to the object of study, that Blockchain is used as a technology that provides many security alternatives oriented to people’s health, but that there are many challenges and limitations that need to be investigated and expanded for future work. The above is evident in the development of RQ3, where many limitations and challenges specific to the technology are mentioned, and which directly impact the confidentiality, integrity and availability, of EHRs and PHRs, essential when providing a health service to people (Yaqoob et al., 2021) (Bedin et al., 2021) (Khurram & Sardar, 2020) (Kouhizadeh et al., 2021) (Hosseini Bamakan et al., 2021) (Gupta et al., 2021), where patients will have control of their own health data and can decide how it is used. The exchange of data between healthcare stakeholders will be easier, better controlled, transparent and reliable, leading to better delivery of healthcare services and quality of life for patients. In this sense, it is necessary to take into account multiple considerations that can affect people’s health and psychology.

7. CONCLUSIONS

Most of the current research on Blockchain technology in the healthcare field focuses on issues...
such as storage, management, security, and privacy of EHRs and PHRs handled in healthcare ecosystems. There is also a concern by several authors in addressing solutions for interoperability and standardization of the entire healthcare ecosystem, likewise, performance and scalability of the proposed systems remain an important gap that needs to be addressed. Blockchain has many security features, including confidentiality, anonymity, decentralization of information, integrity, availability, non-repudiation and traceability, which make it an ideal technology to support services, systems and products, which are required in the healthcare field to provide better care towards people. There is no common denominator for the type of Blockchain used in the analyzed works, although taking into account the type of data handled in the proposed solutions, most of them prefer authorized Blockchain; in addition, different consensus mechanisms are mentioned, which depend on the scope, types of data handled, and permissions granted for the data. On the other hand, there are solutions that, due to their complexity or design, integrate two different types of Blockchain, taking advantage of the benefits of each of them. There are many challenges surrounding Blockchain technology, one of them is the lack of legal and regulatory frameworks, which allow providing clear elements for future developments that implement Blockchain in support of healthcare. Likewise, it is necessary to deepen the proof of concept and generation of practical products that implement Blockchain within their solutions, although several solutions to the challenges and limitations have been presented, many of them are only brief suggestions of ideas and lack a concrete evaluation on their effectiveness. This paper discussed the main aspects of Blockchain technology, leading to the next step of developing a practical use case using Blockchain technology in relation to EHRs.

REFERENCES


C. Burniske, E. Vaughn, J. Shelton, and A. C. (2016), HOW BLOCKCHAIN TECHNOLOGY CAN ENHANCE EHR OPERABILITY.


