applied sciences

MDPI

Article

ShinyAnonymizer Enhanced Version and Beyond: A Further Exploration of Privacy-Preserving Solutions in Health Data Management †

Marios Vardalachakis, Nikos Papadakis and Manolis Tampouratzis *

Department of Electrical and Computer Engineering (ECE), Hellenic Mediterranean University (HMU), GR 71004 Heraklion, Crete, Greece; mtp237@edu.hmu.gr (M.V.); npapadak@hmu.gr (N.P.)

* Correspondence: tampouratzis@hmu.gr

† This article is an extended version of our paper published in the 5rd International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE), Heraklion, Crete, Greece, 2–4 May 2019.

Abstract: Healthcare institutions generate massive amounts of valuable patient data in the digital age. Finding the right balance between patient privacy and the demand for data-driven medical enhancements is essential. Since data privacy has become increasingly important, robust technologies must be developed to safeguard private data and allow meaningful exploration. This issue was addressed by ShinyAnonymizer, which was first created to anonymize health data. It achieves this by rendering anonymization methods easily available to users. The enhanced version of ShinyAnonymizer, with an essential improvement in performance, is presented in this study. We explain the merging of data analysis, visualization, and privacy-focused statistics paradigms with data anonymization, hashing, and encryption, offering researchers and data analysts an extensive collection of tools for trustworthy data management.

Keywords: healthcare data; data privacy; data anonymization; hashing; data encryption; statistics; visualization; privacy-preserving data analysis; ShinyAnonymizer

check for updates

Citation: Vardalachakis, M.; Papadakis, N.; Tampouratzis, M. ShinyAnonymizer Enhanced Version and Beyond: A Further Exploration of Privacy-Preserving Solutions in Health Data Management. *Appl. Sci.* **2024**, *14*, 6921. https://doi.org/10.3390/app14166921

Academic Editor: Luigi Portinale

Received: 2 July 2024
Revised: 3 August 2024
Accepted: 5 August 2024
Published: 7 August 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Background and Context

Health data administration is becoming more complex, and robust approaches are being sought to protect privacy. Health professionals can produce comprehensive patient records with easier-to-track patient health, offer real-time data access, and may help save money [1]. With the growing popularity of electronic medical records and the continual development of healthcare technology, protecting critical patient information has emerged as an essential issue. The broad consequences of data breaches in the healthcare sector include compromised trust in the entire system and individuals' confidentiality [2,3]. Developing and enhancing privacy-preserving tools remain crucial as we navigate this challenging environment [4]. This study explores the world of health data management with a particular focus on privacy, demonstrating the latest advances in ShinyAnonymizer, a tool designed to tackle the complex task of integrating data availability with high privacy rules. Robust privacy-preserving methods are even more critical when considering the strict requirements for data protection established by law, including the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) [5,6]. It may be challenging in current technologies to find the right balance between the value of data and privacy. Data can become less useful for research and analysis if anonymization methods that rely on the past are applied [7]. Although advanced methods, including hashing and encryption, provide more and more extraordinary privacy promises, their execution can be physically difficult and time-consuming. The development and enhancement of methods that guarantee privacy remain vital as we traverse this complex scenario

with a focus on privacy; this study dives beyond the topic of health data management and presents the latest current advances in ShinyAnonymizer, a tool developed to facilitate the complex task of integrating rigorous privacy laws with data availability. To offer various methods for anonymization simply accessible to users, ShinyAnonymizer was initially created to anonymize health data. Also, by merging data anonymization, hashing, and encryption with the analysis of data, visualization, and privacy-focused statistically significant paradigms, the new version presented in this research produces significant productivity improvements. This allows data analysts and researchers to use an extensive selection of effective data management solutions.

1.2. The Significance of Privacy in Health Data Privacy

Privacy in health data management is beyond just a law or moral need and is an essential element, supporting people's trust in healthcare systems [8]. Since health records are extremely valuable, processing them needs to be conducted with care and accuracy. Patients must be happy that their private medical records remain secure, and doctors need methods that encourage the sharing of information for research and medication and maintain the highest standards of privacy. Given this background, ShinyAnonymizer emerged as a remarkable solution, initially created to anonymize health data but later expanded to address larger healthcare industry privacy issues [9].

1.3. Introduction to ShinyAnonymizer and Its Role

ShinyAnonymizer, in its initial creation, was an asset for anonymizing health data, enabling medical professionals and researchers to interact with datasets while respecting individual privacy. As we move forward, the reason for an enhanced and better version of this tool becomes obvious. The ShinyAnonymizer Enhanced Version presented in this article goes beyond simple anonymization by incorporating advanced features and fixing issues observed in its counterpart. The strategy is to establish ShinyAnonymizer as an extensive, innovative solution for privacy-preserving health data management.

1.4. Research Objectives

This article has two primary objectives: initially, to determine and resolve the limitations of the present privacy-preserving solutions in the healthcare industry, and secondly, to showcase an enhanced version of ShinyAnonymizer, illustrating its development in reducing privacy problems. This study adds to the ongoing conversation about privacy in health data management by meeting these objectives. It provides an affordable option for healthcare professionals and investigators that seeks an essential equilibrium between data availability and individual privacy.

1.5. Overview of the Study

In the subsequent sections, this study will be an in-depth literature examination, giving insights into past events and developing trends surrounding privacy issues in health data. The evolving landscape of privacy-preserving technologies is reviewed, highlighting their advantages and visible shortcomings. The section on methodology outlines the study's approach, providing goals for enhancing ShinyAnonymizer. After that, the study goes into the technical details of the updated version, covering significant modifications and enhancements, and it ends with a comparison to the initial version. The results section compares the enhanced tool's effectiveness with previous solutions and real-world application scenarios. The discussion section explores the implications of the enhancements and their potential integration with advanced privacy-preserving systems.

2. Literature Review

2.1. The Growing Awareness of Privacy Issues in Health Data

The rise of privacy concerns regarding health data has its roots in the history of healthcare procedures and the increasing reliance on digital technology. The shift from

study-based healthcare records to digital health records (EHRs) has enhanced data accessibility, while boosting the risk of privacy breaches. The digitization of health data has resulted in dramatic changes, enabling the simple exchange of data but also requiring robust security measures for preserving people's confidential data. Knowing the past progression of privacy concerns is crucial for explaining present day difficulties in health data management [10].

2.2. The Present State of Privacy Laws in Healthcare

In response to increased privacy concerns, separate laws have been created to regulate health data management. The Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union are key examples of the laws that control the way healthcare firms manage patient data [11]. Still, the constant evolution of medical technology and activities calls into doubt the usefulness of these limitations.

2.3. Current Privacy-Preserving Tools

An in-depth review of today's privacy-preserving tools must be conducted to identify the landscape's advantages and disadvantages. Several applications for anonymizing and securing health data have been developed, each with unique benefits and features. Popular technologies include k-anonymity algorithms, differential privacy mechanisms, and encryption approaches. Despite their advantages, these innovations often face expanding constraints, complex deployment, and low adaptation to evolving privacy requirements [12].

2.4. Conclusion of Literature Review

Lastly, the literature review highlights the evolving nature of privacy issues related to health data, monitoring their development from earlier procedures to present issues. The regulatory environment plays a significant role in establishing privacy standards, but current approaches have limits that call for additional research. This sets the foundation for the study's future years parts, which highlight the ShinyAnonymizer Enhanced Version as an option that seeks to close present holes and enhance the requirements of privacy-preserving solutions in health data management.

3. ShinyAnonymizer: Key Features and Functionality

3.1. Key Modifications and Updates

Various modifications and improvements have been added to ShinyAnonymizer to fix the limits discovered in its initial usage. The enhanced version includes modern features that provide a more complete solution for managing health data while protecting privacy, moving beyond traditional anonymization. The anonymization technique has been enhanced, hash and data encryption techniques have been boosted, and its general speed and capacity have been upgraded.

3.2. Description of the Features of the Enhanced Version

ShinyAnonymizer has been enhanced with several enhancements that boost the utility's functionality. One significant development is enhancing support for various data types, especially the structured and unstructured information frequently found in electronic health records. The application can now handle an increased number of healthcare data settings, ensuring the protection of sensitive data while maintaining the data's usability for research and analysis. The enhanced version contains dynamic anonymization techniques, which allow real-time adjustments to evolving privacy requirements and new security risks.

3.3. Critique Concerning the Original Version

Contrasting the enhancements made to ShinyAnonymizer's first version is crucial in assessing the modifications made. This comparison emphasizes visible improvements in efficiency, functionality, and privacy protection and explains what has been achieved.

The efficacy of the enhanced version for handling complex healthcare data scenarios is compared with the weaknesses of the initial program, providing an overview of the advancements accomplished in boosting privacy protection.

3.4. Technical Further Details

The technical specifics of the modifications implemented in ShinyAnonymizer are explained in great detail. A detailed examination of the changes applied to the anonymization algorithm is additionally covered, as well as the techniques adopted to give a higher level of privacy while preserving the importance of the data. The latest security processes utilized to avoid disclosing private medical information are shown, in addition to practices concerning authorization and authentication for users. More factors about its freedom and effectiveness are discussed, displaying the upgraded version's potential for handling many significant and unique datasets.

3.5. Overview of the Research Design

The goal of the research's design is to review the enhanced ShinyAnonymizer and deal with the weaknesses in the identified privacy-preserving solutions. Initially, an in-depth examination of the current privacy-preserving methods and tools forms an element of the research. This examination addresses their effectiveness, limitations, and flexibility in adapting to increasing privacy issues in healthcare data management. The goal of giving an overview of the enhancement phases is to determine specific constraints and challenges that the enhanced version of ShinyAnonymizer must address.

Following an analytical methodology involving quantitative and qualitative research techniques, an enhanced ShinyAnonymizer was created. Examining the enhanced tool in relation to past versions and other privacy-preserving options and evaluating its performance are instances of quantitative methodologies. The measurements are organized around the system effectiveness in broad terms, efficiency, and anonymization efficiency. Detailed comments on the tool's practical utility and areas that require further improvement are gathered, utilizing qualitative approaches that involve data analysts and healthcare experts engaging with the application.

Utilizing three methods, the enhancements will continue to be good theoretically and practically. In addition, a complete evaluation step is integrated into the research design, after the enhanced ShinyAnonymizer is placed under examination in real-life scenarios. This involves identifying the software in multiple healthcare environments, while assessing its use and practical impact on data privacy. What is gathered via these deployments is analyzed, to find out how effectively the tool integrates data utility with security issues. The results of this research help to enhance the tool by providing insightful knowledge as to how it can be connected with the healthcare industry's present privacy-preserving technologies.

3.6. Description of the Research Methodology

The research utilizes methodologies that involve an in-depth examination of the latest and most current privacy-preserving technologies, an in-depth overview of issues of privacy regarding health data, and an in-depth evaluation of ShinyAnonymizer's advances in technology and goals. Each one of these methodologies enables a deeper understanding of the problems related to health data management, ensuring that the ShinyAnonymizer Enhanced Version addresses the theoretical as well as practical aspects of privacy protection.

3.7. Selection Criteria for Enhancements in Methodologies of ShinyAnonymizer

ShinyAnonymizer's enhancements are the outcome of a careful evaluation of methodologies that brings a consideration of the shortcomings demonstrated in previous solutions, while additionally applying those methodologies to the growing requirements of healthcare data management. These requirements comprise the ability to handle various data types, flexibility, functionality, and adaptability to potential privacy risks. The research group collaborates alongside privacy experts, healthcare professionals, and technology specialists

to ensure an in-depth investigation and the right choice of methodologies for participation in the enhanced release.

3.8. Data Sources Apis

A diverse method is utilized for gathering the data required for this research, grabbing from various resources to create a big and accurate dataset. Healthcare organizations and facilities that provide permission to view de-identified health records with associated information act as the primary sources of data. Working together with hospitals, clinics, and academic organizations makes it possible to collect real-world health data that demonstrate the intricate details of the healthcare industry.

The research effort additionally explores the publicly available datasets about medical and health studies. Public datasets from trustworthy sources, such as worldwide medical databases, medical agencies, and academic organizations, provide significant information for research and replication. Such databases could include information regarding demographics, health diagnoses, outcomes from therapy, and anonymized medical records. The findings of the research are open and reliable, as datasets that are publicly accessible were utilized.

These independent sources present helpful perspectives on the practical uses of privacy-preserving technologies, as well as a better understanding of the challenges that are faced in health data management. With the combination of qualitative and quantitative data from multiple sources, this study intends to offer an in-depth and insightful examination of the enhanced ShinyAnonymizer's effectiveness in fixing concerns about privacy in the management of health data.

The data provided in our study include a carefully created fake medical database aimed at reducing the challenges encountered when obtaining access to real electronic medical records (EMRs). EMRbots.org is an adaptable system that simulates the characteristics of real medical databases. It consists of economic data, statistics about patients, admission information, and an extensive variety of lab conclusions. Given its flexible characteristics, users can develop an extensive variety of artificial patient populations that include various genders, backgrounds, illnesses, and other features. This allows simpler-to-evaluate machine learning algorithms, the conducting of studies in educational institutions, and it can occasionally indicate to potential buyers and investors how novel EMR management platforms can be established [13].

Finally, the generated patient database has been configured with significant information contained in core-filled tables. Individual patient IDs, gender, date of birth, race, marital status, language, and socioeconomic status are all contained in the "PatientCorePopulatedTable". Additional APIs have been included in this layer to allow access to additional data sources, such as relational databases, CSV files, and Excel files. Also, the appropriate API and REST functions may be utilized to dynamically obtain data from external data sources and implement various hashing, encryption, and anonymization techniques.

3.9. Background of Data Anonymization Techniques

Data anonymization is the first implementation of our methodology to protect sensitive information and maintain individual privacy. Data anonymization's primary goal is to conceal personally identifiable data (PII) in databases, so that criminals have difficulty linking particular pieces of data to specific individuals. Initially, individual identifiers within the dataset, such as names and social security numbers, are deleted. However, deleting identifiable elements can often be unnecessary. The next step is to use these complex techniques to prevent re-identification, so data cannot be quickly linked to specific individuals. Utility concepts and quantitative privacy models are combined with anonymization methods to assess and regulate risks [14]. As privacy models analyze the prospect of prospective privacy violations, the utility theory examines the anonymized data's significance to its intended purpose. These models assist with determining the proper anonymization methods by defining standards for reasonable danger rates. Multiple techniques are utilized

based on the type of data and the level of risk, like generalization, removing information, suppression, and bottom coding.

3.10. Technical Details of Anonymization Techniques

- *Generalization*
  a. Description: Generalization means substituting particular data values with higher but exact categories. For instance, age ranges can be utilized instead of real ages. This method reduces the resolution of the data, rendering it harder to determine individuals while maintaining important aggregate data.
  b. Application: Generalization efficiently handles sensitive variables like age, wealth, or geographic location. It is a usual approach for scenarios when unique individual data is not necessary, but patterns or trends must be analyzed.
- *Removing Information*
  a. Description: This method requires eliminating specific identifiable data from the dataset. For instance, it removes names, social security numbers, and other direct identifiers from a database.
  b. Application: Removal is easy, but it might not always be sufficient, mainly since indirect identifiers can still be utilized to re-identify an individual. It is often employed together with other anonymization techniques to enhance privacy.
- *Suppression*
  a. Description: Suppression involves concealing or hiding particular data values. For instance, sensitive variables can be replaced with a common placeholder (e.g., substituting a patient's true gender with an asterisk '*') or parts of data (e.g., only disclosing the initial three digits of a social security number).
  b. Application: Suppression is often used to protect highly sensitive data that cannot be generalized without drastically decreasing the data utility. It is beneficial when just a tiny amount must be preserved while permitting a specific data analysis.
- *Bottom Coding*
  a. Description: Bottom coding allows data to be divided into more significant categories, particularly with variables with an extensive range of values. Earnings levels, for example, may be categorized as "low," "medium," or "high" instead of offering actual statistics.
  b. Application: Bottom coding minimizes the risk of re-identification by grouping individual data points into larger groups, rendering it less likely that every individual data will be wiped out. It is often employed with different methods to guarantee data privacy while permitting helpful research.

All the techniques were chosen, to find the right balance between privacy and data utility. Once methods of anonymization are used, an extensive evaluation process is required. Evaluation means deciding if personal characteristics have been successfully removed by the methods of anonymization and if the data still have enough quality to be utilized for what was intended. The process makes sure that each person's privacy is well safeguarded, despite the anonymized data's analytical value. After confirmation, the anonymized dataset may be transmitted or maintained without risks. This planned strategy optimizes the privacy and security of data by permitting businesses to effectively utilize their data, while following legal and ethical guidelines [15].

Also, compared to perturbation, which demands adding noise to data, generalization swaps bigger groups for specific items [16]. All the methods were chosen to find the right balance between privacy and data utility. Once methods of anonymization are used, an extensive evaluation process is required. Evaluation means deciding if the methods of anonymization have successfully removed personal characteristics and if the data still has enough quality to be utilized for what was intended. The process ensures that each

person's privacy is well safeguarded despite the anonymized data's analytical value. After confirmation, the anonymized dataset may be transmitted or maintained without risks. This planned strategy optimizes the privacy and security of data by permitting businesses to effectively utilize their data, while following legal and ethical guidelines [17].

A brief description of our anonymization process is provided in Figure 1: Beginning with the database identification, the data anonymization process takes place immediately. Initially, businesses determine the specific database or datasets required to be anonymized. As a base for the process's future steps, this requires a basic understanding of the data format and the different types of data it holds. Risk assessment follows next, soon after database identification. In this step, the data's sensitivity is analyzed, and potential risks from the data's exposure are identified. Risk assessment also assists in realizing the possible data use and guides the choice of the proper methods of anonymization to mitigate these risks. Next in this process is the selection of anonymization methods when the risk assessment is complete. Effective strategies are chosen in the present scenario, based on the dangers discovered and the kind of data.

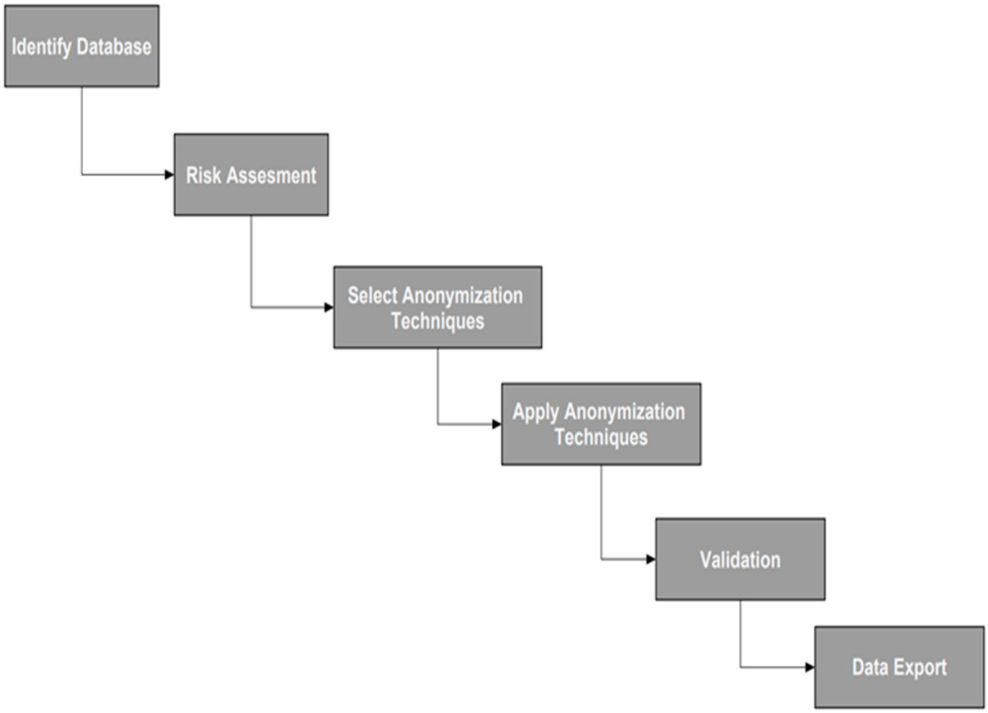

Figure 1. Data anonymization methods are utilized to execute a conceptual process.

There are alternative techniques, such as encryption, masking, and generalization. The above methods are selected and then used to convert the data, to preserve privacy in the anonymization process. Validation takes place on the data after the implementation of anonymization processes. This critical step makes sure that anonymization has been utilized properly and that the data can continue to be used for its intended objectives. With validation, the predictive power of the anonymized data is preserved, while unique identities are secured. Providing the anonymized dataset for allocation or storage is the final step, referred to as data export. Data security and privacy rules confirm that the data can be used for the reasons it was first obtained.

A scenario of integrating several transformation models is provided in Table 1. In this scenario, Table 1 shows the initial dataset of the healthcare organization in command. Specific patient data, including gender, race, marital status, language preferences, and unique identifiers (PatientId), are included in this data collection. For example, a record with the number FB2ABB23-C9D0-4D09-8464-49B70B982F0F0 defines a married male patient of African-American race. While these broad data are needed for internal processes

and analysis, the risk of personal identification poses serious issues regarding privacy. The business utilizes anonymization methods, as shown in Table 2, to mitigate these issues with confidentiality. By applying suppression and removing information methodologies, the PatientGender and PatientRace columns in this particular situation have been deleted and swapped with general placeholders (*), which limits the monitoring of specific features. Bottom coding has been utilized to modify the PatientMarital-Status, combine different situations into greater categories (e.g., 2, 2, 2), and expand the PatientLanguage into a specific range (45,000), indicating averaged language data. While preserving the dataset's value for research and analysis, this anonymized copy maintains the privacy of sensitive individual data.

Table 1. Original Dataset.

PatientId	PatientGender	PatientRace	PatientMaritalStatus	PatientLanguage
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	Male	Unknown	Married	Icelandic
64182B95-EB72-4E2B-BE77-8050B71498CE	Male	African-American	Separated	English
DB22A4D9-7E4D-485C-916A-9CD1386507B	Female	Asian	Married	English

Table 2. Anonymized Dataset.

PatientId	PatientGender (Suppression)	PatientRace (Rem. Information)	PatientMaritalStatus (Bottom Coding)	PatientLanguage Generalization
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	*	127	2	45,000
64182B95-EB72-4E2B-BE77-8050B71498CE	*	228	2	45,000
DB22A4D9-7E4D-485C-916A-9CD1386507B	*	315	2	45,000

Anonymization techniques enabling such complicated transformation methods cannot be implemented by just searching for a perfect approach all over the area of any potential output datasets; in most instances, the available search spaces are excessively enormous. This resulted in the creation of several hashing, as well as complicated data encryption, algorithms. However, we emphasize how important it is to keep in mind the conceptual representation of data anonymization processes that employ specific combinations of exposure, utility, and transformation models. For instance, the actual use of earlier algorithms is extremely restricted, as they frequently only work with a particular combination of defined models. Therefore, there are quite a few open-source technologies that are available to everyone. It is widely accepted that context—which involves the dimensionality, volume, and statistical features of data—has an important effect on the effectiveness of certain anonymization methods. The kind of analyses or applications that will be performed on the data, if any restrictions on access will be implemented, whether or not the data are going to become accessible to everyone, and whether or not the data are tabular, linear, or operational, are additional important decisions that have to be adopted. Due to this, various algorithms and methods for modifying data and measuring changes in utility need to be provided for anonymization software, to be utilized in a range of application scenarios.

In summary, generalization, removing information, suppression, and bottom coding are fundamental to anonymizing data and protecting privacy while maintaining utility. Each technique has advantages and weaknesses, and they are often combined to balance the trade-off between privacy and usability. The present paper explains the various anonymization techniques and how they should be applied to ensure the maximum privacy protection of sensitive information, without interfering with the usability of the data.

4. Enhancements to ShinyAnonymizer Hashing Techniques

4.1. Implementation of Advanced Techniques such as Hashing

The second method we will use in our methodology and system is that data will be converted utilizing a secure method of hashing into a fixed-length character sequence known as a digest or hash number. Because this process cannot be reversed, it is ideal for securely storing secret data, including passwords, and ensuring data integrity. Hashing is utilized to verify data accuracy without exposing the initial data, in contrast with encryption, which allows the material to be extracted back to its initial form. Notable hashing algorithms with various computation effectiveness and security levels are MD5, SHA512, CRC32, and XXHASH64 [18]

4.1.1. MD5

a. Description: MD5 is a common cryptographic hash algorithm that creates a 128-bit hash value, which is often expressed as a 32-character hexadecimal integer. It was intended to be a function with one direction, so it is computationally difficult to go through the steps and recover the initial data from the hash value.
b. Application: MD5 is frequently utilized as a checksum for confirming data integrity. But because of weaknesses allowing for hash collisions (in which two distinct inputs exchange the same hash), MD5 is rarely suggested for secure applications demanding cryptographic power [19].

4.1.2. SHA512

a. Description: This method is implemented when the dataset is anonymized using unique or specific data. For instance, to delete all identities, account numbers, and other private data from a database.
b. Application: SHA-512 builds on the SHA-2 family, generating a hash value of 512 bits. This shows up as a hexadecimal integer with 128 characters. It aims to be more dependable than MD5, offering better protection against collisions and standing up better to attacks on its cryptography [20].

4.1.3. CRC32

a. Description: CRC32 is also an example of a non-cryptographic hash algorithm that is 32-bit large. It stands between the most commonly utilized algorithms for error detection in transmission and retention and for checking accidental changes in raw data.
b. Application: CRC32 is utilized to guarantee integrity during file transfer and in any other action involving data storage. Though it is useful in detecting errors, it is not safe for cryptographic purposes, because of its low collision tolerance and loss of security against beneficial tampering

4.1.4. XXHASH64

a. Description: Bottom coding is a way of separating data into larger groups, especially if factors have a broad range of results. Income levels, for instance, may be categorized as “low”, “medium”, or “high”, without offering details.
b. Application: XXHASH64 is a fast, non-cryptographic hash function that determines a 64-bit hash. It was built for speedy processing and has been revised to immediately hash huge datasets.

Figure 2 illustrates briefly how our hashing process works: Data are initially extracted from the database, and in addition to identifying the significance of the data and any privacy issues, a complete risk evaluation follows. This analysis helps in selecting the most suitable hashing method. For example, despite its recognized weaknesses, due to its speed, MD5 can be chosen for more significant data. However, SHA-256 or SHA-512 may be selected for more sensitive data due to their more considerable level of security. Salting is often employed, together with hashing, to offer a higher level of security, while minimizing the risk of hash collisions—where two different sources produce a single hash value. The

data are hashed if the risk examination has ended and the most suitable hashing method has been established. To achieve this, every piece of data must be turned into a hash code, while maintaining its trustworthiness and usefulness for research [21]. To show the various levels of sensitivity demanded for security, patient genders in a dataset can be hashed utilizing MD5, while patient races can be hashed using SHA-512. After being transferred for utilization, the hashed data are thoroughly examined to verify their precision and accuracy, reducing security risks during transmission or storage. While hashing has several security and computing productivity benefits, it also has limitations. The key benefit is the fact that it can rapidly and accurately validate data, which is demanded for maintaining data integrity and secure procedures for authentication. The selection of hashing algorithm is important, however; while powerful algorithms, including SHA-512, provide greater protection at the cost of greater processing requests, quicker methods, including MD5, are simpler but more open to attacks. Also, when trying to ensure that each input generates a unique hash, methods including salting have to be utilized, owing to the issue of hash collisions. Organizations may properly enhance their data privacy practices and maintain private data security, while keeping the data's usability, by carefully selecting and utilizing the relevant hashing algorithms.

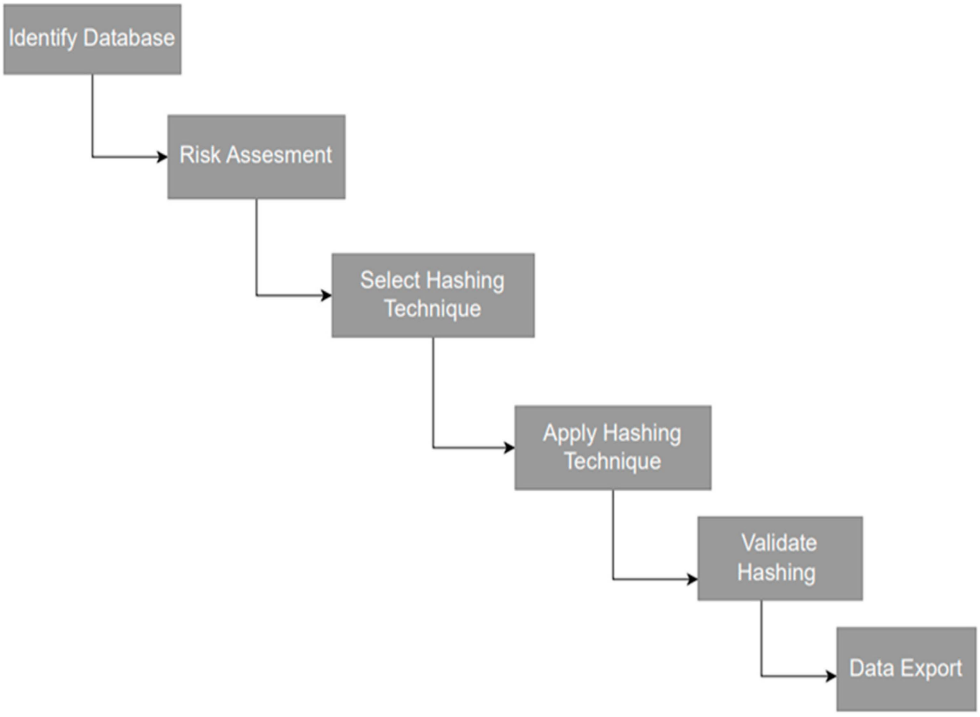

Figure 2. A visual representation of the hashing process.

Utilizing hashing as an advanced transformation method provides advantages as well as drawbacks. Its strength is in its computational effectiveness, which renders it suitable for the fast integrity of data audits and the authentication of passwords. However, adopting an appropriate hashing algorithm is essential, as smaller algorithms could expose problems. In addition, hash functions are vulnerable to hash collisions, in which multiple inputs produce an identical hash value. Salting, which introduces arbitrary data to every input before hashing, attempts to reduce this problem [22].

Table 3 shows an initial dataset, including sensitive patient data, in the context of healthcare. Together with data including gender, ethnicity, marital status, and preferred language, every record has an exclusive identification (PatientId). For example, the dataset includes FB2ABB23-C9D0-4D09-8464-49B70B982F0F0, a male patient who is African-American, married, and speaks Icelandic. While these particular and identifiable data are required for internal use, exposing them could seriously compromise privacy.

The business uses hashing techniques to enhance data security, producing the dataset in Table 4. Numerous cryptographic methods are used to hash sensitive features, including PatientGender, PatientRace, PatientMaritalStatus, and PatientLanguage. PatientLanguage is hashed with XXHASH64, PatientMaritalStatus is hashed with CRC32, PatientGender is hashed with MD5, and PatientRace is hashed with SHA512. These changes sustain the initial sensitive data's utility for certain kinds of analysis, while turning them into hashed values that make them hard to understand and detect.

Table 3. Original Dataset.

PatientId	PatientGender	PatientRace	PatientMaritalStatus	PatientLanguage
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	Male	Unknown	Married	Icelandic
64182B95-EB72-4E2B-BE77-8050B71498CE	Male	African-American	Separated	English
DB22A4D9-7E4D-485C-916A-9CD1386507B	Female	Asian	Married	English

Table 4. Hashed Dataset.

PatientId	PatientGender (MD5)	PatientRace (SHA512)	PatientMaritalStatus (CRC32)	PatientLanguage (XXHASH64)
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	$1$00IZNdMg$3Qyw6 QCoTt25vG.MSOZI.	b90e54bb3e16afad4067bf7777 16387559083b84a3a794788f824 od4d808fa3760f9d6f7d39673b 17779868eb4f6f628f8e76101256a 71476fe5a5587bb3c22	03cd7b91	ce023865e6969833
64182B95-EB72-4E2B-BE77-8050B71498CE	1UpOHW8F$Rbiks1 HN9IYoHYKd3c6gehN/	C10988cfc79f1174893cf32c9c52 66fb9a4984b05bb24170d502932 62cfdb62be73b2994a87do705548 a72367bbc04b2937e382a5ec743a 90375da30a0ef56d3	d4296c7c	f7eefde29f3e16f2
DB22A4D9-7E4D-485C-916A-9CD1386507B	1Z0Y.V1HS$y1W6Tu hjoMpu7YZQD6qp0	C33ea9d0ba47653638e278a8d3 cab55856c064a148112d377b55fbe 89b5cb17f6bfcfd709e9e989f32b245 eca33640f2ecf857253d09b4ece807 884f8aef35c051	03cd7b91	f7eefde29f3e16f2

We highlighted how important it is to select the proper algorithms based on the amount and sensitivity of the data. Regarding its high level of security, SHA-256 is often used in blockchain systems; however, this comes at some price when it comes to computational speed. We suggest that SHA-512, which provides a higher security level with a potentially quicker computing process for certain applications, could be utilized to conquer this cost of computation. Table 4's lack of jointly signed or stated hash values still creates issues concerning the system's willingness to use them. A process similar to blockchain mining can be utilized to validate these hash values, and every hash is verified with a set of criteria and signed digitally to verify its integrity and authenticity. By guaranteeing that the system acknowledges and embraces the hash values, this method enhances the entire data security. Every transaction block in blockchain mining is hashed and must fulfil several specifications, so an extensive amount of processing power is required. To find a hash value that fits the testing criteria, participants may continuously hash the block data with small changes. After that, the network obtains this proper hash, which additional nodes utilize to verify it is correct, by hashing the respective block levels and contrasting the result to the hash that was sent. Each block is integrated into the blockchain, confirming its authenticity and reliability if the hash meets all of the criteria and can be verified by a majority of nodes. Similarly, a verification process can be used in this research, in which numerous nodes or organizations sign digitally and confirm the calculated hash values [23]. This ensures that they are valid and recognized by the system, also complying with security standards and maintaining data integrity. A digital signature method, where each hash value is signed

by a trustworthy authority, may be utilized to enhance the validity of the hash functions utilized in this research. This signature confirms that the value of the hash was generated and verified in line with the security needs of the system. Also, the confirmation task could be spread between many nodes or controlled using a decentralized validation process similar to blockchain mining. This distributed method enhances security by reducing the risk of one point of failure. It increases system trust by ensuring that multiple parties verify and accept hash numbers separately.

Finally, the advanced hashing techniques implemented in ShinyAnonymizer significantly enhance data security and integrity. MD5 provides speed, while SHA-512 is powerful in cryptographic security. CRC32 is helpful in error detection, and XXHASH64 assures high performance. Last but not least, the problem of possible vulnerabilities is addressed by introducing salting, so that the hashed data stays safe, reliable, and suitable for a wide range of applications. The proper selection of a hashing method concerning data sensitivity and application requirements assures security and functionality.

5. Enhancements to ShinyAnonymizer Data Encryption Techniques

One of the most crucial elements of current data security methods is complex transformation technologies, such as data encryption. Using algorithms, data encryption converts plaintext data into ciphertext, an unreadable form. Encryption keys are used to ensure that the data are only exposed by those with the right decryption keys. To eliminate the possibility of unauthorized access, encryption is required to secure sensitive data during transfer, retrieval, or analysis [24]. Selecting the right encryption algorithm is a significant challenge when establishing data encryption. DES, XDES, Blowfish and Advanced Encryption Standard (AES512) are common symmetric encryption methods. The form of secured data, calculating speed, and unique security requirements impact the method selection. Because of its outstanding security and productivity, AES512 is commonly used to secure data in transmission or at rest [25].

5.1. Technical Details of Encryption Techniques

5.1.1. DES

a. Description: DES is a symmetric block cipher that encrypts data in 64-bit blocks with a 56-bit key. Being one of the most often used encryption algorithms, it has nowadays grown insecure because of the vulnerability to brute-force attacks.
b. Application: However, the DES previously working in the encryption of sensitive data, due to the inappropriate minimal length of keys and sensitivity to current computing power, has turned out to be useless.

5.1.2. XDES

a. Description: One of the processes involved in removing direct identifiers or specific identifiable pieces of data in a dataset. For example, scrubbing a database would involve removing names, social security numbers, and other such direct identifiers.
b. Application: This is basically an extension of DES. It deploys multiple phases of DES encryption with different keys, hence improving on the key size and security.

5.1.3. BLOWFISH

a. Description: Blowfish is a symmetric-key block cipher. Data is encrypted in blocks of 64 bits under keys ranging from 32 to 448 bits. It was generated to replace DES, as it is faster and more reliable; moreover, it exudes flexibility and efficiency.
b. Application: Blowfish allows a user to conduct the encryption of data in numerous application types. Among other important uses, these include virtual private networks, CDs, and files. Its many uses and strong safety features make it pretty suitable for most such applications.

5.1.4. AES512

a. Description: For AES-512, in this methodology, this approach uses a 512-bit input and key size that provide more strength against cryptanalysis; however, the drawback of this is an extended authorized area.
b. Application: AES-512 can find an application in any domain that needs strong security and a fast speed but has only restricted chip areas available for the block cipher. This would, therefore, apply to entertainment applications and even in satellite network communications.

Figure 3 illustrates briefly how our data encryption process works: Selecting the proper database and then setting up the connection to it is the initial step in the data encryption method. The initial step ensures that the proper data repository is utilized for any following operations. The needed cryptographic abilities, like the pgcrypto module in PostgreSQL, will be enabled once linked. After all, has been set up, the next crucial step is to identify the sensitive data stored in the database. In-depth research takes place to assess the degree of sensitivity of the data that needs to be secured and the possible effects of disclosure. After gathering sensitive data, a detailed risk evaluation is carried out to identify the possible risks and weaknesses linked to the data. This can also help select the most effective encryption technique. After the risk evaluation finishes, proper encryption methods are selected. Concerning its high level of security and productivity, the Advanced Encryption Standard (AES512) is a suitable selection. According to the requirements, other methods, including RSA, Blowfish, or Triple DES, can also be utilized. After the algorithms have been selected, they are utilized to transform sensitive data into ciphertext, making them unreadable to other people. A confirmation step is executed after the procedure of encryption, to verify that the data have been securely encrypted without losing their authenticity or utility. In the end, the encrypted data are ready for safe transfer or storage, guaranteeing their protection for every moment of its presence.

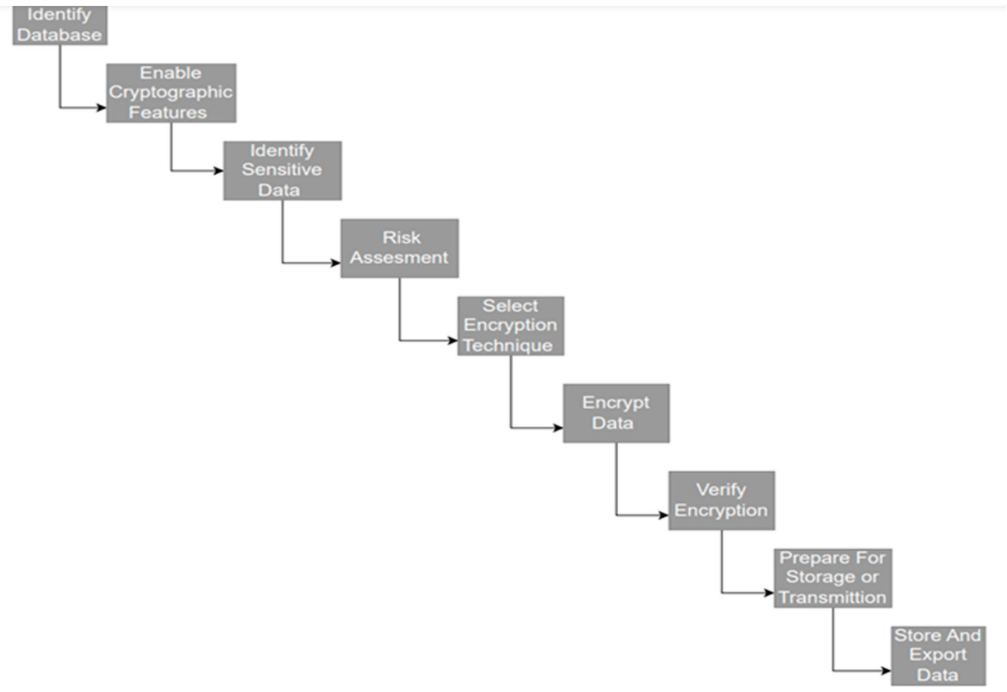

Figure 3. A visual representation of the encryption process.

Important patient data in the healthcare industry must be controlled properly to maintain privacy and comply with rules like GDPR and HIPAA. Consider a job at a hospital that keeps a large database with patient data, like PatientId, PatientGender, PatientRace, PatientMaritalStatus, and PatientLanguage. Clinical studies, medical analysis, and epi-

demiological research all require this dataset. However, before these facts can be used for additional causes, they need to be anonymized, owing to the risk of data breaches. This example shows how the hospital removes these data utilizing the enhanced ShinyAnonymizer tool, preserving patient privacy while retaining the data's analytical utility. Three patients' details were included in the original dataset (Table 5):

- Patient 1 is a married man of unclear race who talks Icelandic and can be tracked only by his PatientId;
- Patient 2: An African-American man who lives separately and understands English can also be properly identified by his PatientId;
- Patient 3: An Asian woman who has a husband and is skilled in English.

Table 5. Original Dataset.

PatientId	PatientGender	PatientRace	PatientMaritalStatus	PatientLanguage
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	Male	Unknown	Married	Icelandic
64182B95-EB72-4E2B-BE77-8050B71498CE	Male	African-American	Separated	English
DB22A4D9-7E4D-485C-916A-9CD1386507B	Female	Asian	Married	English

A scenario of integrating several encryption models is provided in Table 6. Due to the sensitivity of such data points, the hospital utilizes the enhanced ShinyAnonymizer to alter every field employing strong encryption techniques. PatientId, for instance, is encrypted with the Data Encryption Standard (DES), generating a secure, anonymous word. Gender information is encrypted with Extended Data Encryption Standard (XDES), racial data utilizing Blowfish encryption, and marital status and preferences for languages with Advanced Encryption Standard (AES512).

Table 6. Encrypted Dataset.

PatientId	PatientGender (DES)	PatientRace (XDES)	PatientMaritalStatus (BLOWFISH)	PatientLanguage (AES512)
FB2ABB23-C9D0-4D09-8464-49B70B982F0F0	J9..raL7sarlpgZHu3M	N7ez8ptklHSN2	$2a$06$dX.ox5Kw5UTea.PI1fpejeo/jxaVaKbQIq.9OrVnPHwb61NsI5Nis	\303\015\004\007\003\002\215\247\001\230\252:\270\\d\3228\001\003M/\
64182B95-EB72-4E2B-BE77-8050B71498CE	_J9..cJLyoOKPqPnzXro	uA/swfmosamBU	$2a$06$TTudssu1SSODVb9N6gWWO.wNydggPaPirel6UUFBF.feeA0ZNPM8u	\303\015\004\007\003\002\256\223\215\230\304\265\263\211z\01~L&\315
DB22A4D9-7E4D-485C-916A-9CD1386507B	_J9..hqKDaOxM7uNNmWU	\cumibnjFeBjF2	$2a$06$86GkCFeqVPctzovZQYYkS.vYP0qcu5QX3rdIP9f/Jj33ZaHO6uGzy	\303\015\004\007\003\002\302\\\304\341T\264\010\244n\3228\001\353

The enhanced dataset maintains the necessary format and statistics for study while deleting identifiable data, ensuring that one of the patients is unable to be re-identified utilizing the anonymized dataset. By applying those encryption techniques, a medical facility can securely transmit and evaluate data. Scientists can use them in an exploratory data analysis, receive perspectives on patient profiles and health trends, and help lead to advancements in medicine while respecting patient privacy. The enhanced ShinyAnonymizer finds an essential compromise between value for data and privacy, enabling medical institutions to profit from their data resources while adhering to strict privacy laws.

ShinyAnonymizer uses advanced encryption techniques for the latest and most secure data protection and confidentiality. Some suitable algorithms and approaches are DES, XDES, Blowfish, and AES512, going to enormous detail in applying the very process of encryption itself to guarantee that sensitive data are shielded from unauthorized access

but retaining the analytical value of that data. This approach ensures that all the privacy concerns are met, thus providing room for secure data management and research.

5.2. Complete Data Security: Combining Hashing, Encryption, and Anonymization

In the digital age, data create new ideas and knowledge, so providing complete data protection is important. Adjusting rules, issues with privacy, and incidents of data loss need a holistic strategy. This talk analyzes the utilization of data anonymization, hashing, and encryption to avoid exposing important data, ensure data integrity, and covers the ethical and legal standards of a privacy-conscious society in general.

Data anonymization is the very first weapon in response, with an emphasis on maintaining separate identities inside databases. Taking away or changing private data (PII) requires using techniques such as perturbation, suppression, and generalization. A complicated technique must be used for employing anonymization together with hashing. Hashed identification codes with sensitive features are utilized for generating pseudonyms. Without revealing the primary names of the individuals in question, this relationship enables the connecting of significant data between numerous datasets while preserving individual privacy protection. By protecting essential features before the data get transmitted, merging anonymization with encryption minimizes the risk of re-identification across the anonymization procedure.

Hashing is an essential process for ensuring the accuracy of data, because it converts data into hashes, which are unrecoverable types. Outside its usual role, hashing plays an important part in combining multiple methods. Hashed values, which are used as passwords or credentials for encryption methods, make the encryption process tougher [26]. Furthermore, when utilized in tandem with data anonymization, encrypted IDs enable safe data linkage without revealing real names. Integrating hashing, encryption, and data anonymization leads to an extensive approach that covers the entire data lifetime. This integrated strategy could guarantee data security from the time of collecting until all stages of processing, collaborating, and storage. It minimizes the risk of re-identification and unauthorized access to data. Also, the analytical significance of the data is retained by this method. Appropriate studies and evaluations continue to benefit from associated and anonymized datasets as soon as they obtain an appropriate balance between privacy and utility.

In short, hashing, encryption, and data anonymization collaborate to give a flexible and effective platform for general data security. This collaborative method maintains sensitive data, while resolving the challenges imposed by a constantly evolving data environment. As technology advances and worries about privacy expand, both of these strategies are going to stay at the top of data security, promoting trust and suitable handling techniques in a rapidly expanding digital world.

5.3. Integration of Data Analysis Visualization and Statistics Paradigms

Because ShinyAnonymizer integrates data analysis, visualization, and statistical paradigms, scientists can gain useful knowledge from anonymized datasets in a constantly shifting setting. Using the tool, customers could use statistical approaches to anonymize their health data and perform a fast exploratory data analysis (EDA). Description statistics could assist researchers in recognizing the delivery, main trend, and variability of the anonymized data, which might assist them in discovering trends and patterns.

Also, the tool provides scientists with a simple interface through which they can combine data visualization paradigms to examine the anonymized data. Some instances of methods of visualization that are useful to identify correlations and trends in data include scatter plots, box plots, and heatmaps. Since interactive dashboards allow for real-time enhancements and immediate feedback on the anonymized information, consumers can interact with these visualizations.

The importance of effectively sharing the results from studies is acknowledged by ShinyAnonymizer. Therefore, it simplifies the task of developing visually appealing and

useful graphics that are easy to share with colleagues or customers. The tool integrates dynamic visualizations and statistical analysis to create a link between complex approaches and simple presentation. This increases the user experience in general and facilitates a more informed analysis of the health data that was recently anonymized. In short, ShinyAnonymizer protects health data confidentiality via scientific guidelines, data analysis, and visualization, which allows scientists to carry out deep statistical and graphic examinations. Because it facilitates a greater understanding of the transformed datasets, this holistic approach boosts extensive informed analysis.

6. The Proposed Enhanced ShinyAnonymizer Tool

6.1. Workflow and System Architecture

Business experts who want to anonymize data with a computer science professional's help are an ideal audience for ShinyAnonymizer. Figure 4 shows each of the phases that form the procedure. The anonymization process is shown in the illustration. We aim to expand the popularity of data anonymization to a broader variety of consumers who might not have expertise in anonymization methods. We chose to utilize many methods that are very easy to set up and simple to use for non-expert users.

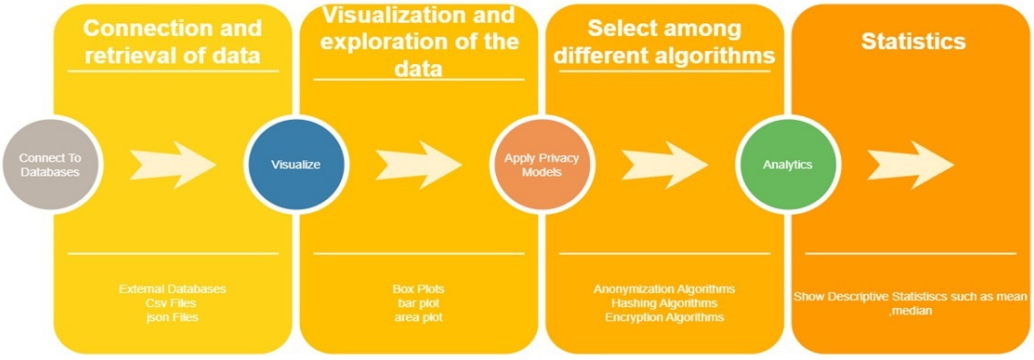

Figure 4. System workflow.

The first step in this method is connecting, gathering, and showing the data that exists already in other databases, Excel and CSV files, and other forms. Utilizing several kinds of R scripts, ShinyAnonymizer allows access to a wide range of relational databases and file formats presently in use. To access and analyze the data that are available, ShinyAnonymizer can connect to the appropriate data provider in every scenario by only setting the correct connection configurations. It must first be visualized in the following phase to allow the active discovery of publicly available data. Clients may play with the different privacy models and select what data will be processed later. The third phase lets clients choose from a broad range of hashing, anonymization, and encryption methods. In the last phase, participants can conduct multiple kinds of analysis, utilizing various privacy models to confirm the visible data. Also, in Figure 5, a three-layered architecture consisting of the Data Source Application Programming Interfaces (APIs), privacy models, and graphical user interface (GUI) has been created and formed to perform the previously defined workflow.

It is a web-based application that allows users to learn about data and understand how various privacy models impact that data through graphs and visualization. Utilizing the Shiny R package, HTML, JavaScript, and CSS were the elements in developing this layer. The data have been transformed and stored in an internal PostgreSQL database after establishing the connection to an external data provider. The user will then see a tabular view of the data, letting them take an active role in each data field to decide which must be hashed, encrypted, or anonymized.

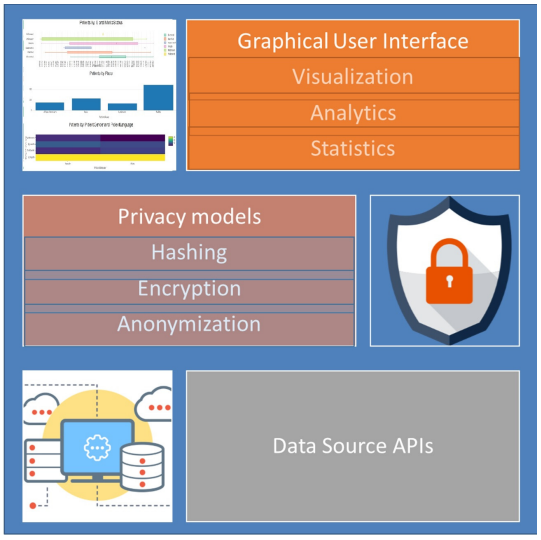

Figure 5. System architecture.

6.2. The GUI

Because of the goal of maintaining privacy in the dynamic area of health data management, the recommendation for an enhanced ShinyAnonymizer version is the perfect solution. The complex task of managing health data currently makes seeking new techniques to improve individual data and maintain usability more critical. The ShinyAnonymizer Enhanced Version resolves this issue with several enhancements and changes that fix the flaws in the initial version. The necessity of privacy in managing health data is hard for anybody to underestimate. The public's confidence in healthcare institutions is established and upheld significantly through it, and it is not merely required by legislation or ethical behavior. As there are fewer electronic medical records, there will be a higher likelihood of information leaks, which puts people's confidentiality at risk and reduces public trust. Combined with anonymizing health data, the ShinyAnonymizer Enhanced Version is an efficient solution to the broader safety concerns impacting the healthcare industry. This tool's creation indicates an attitude toward finding the right equilibrium between user confidentiality and public data availability.

Knowing the origins of ShinyAnonymizer and its place in the medical industry is essential for understanding the application's growth. As conditions evolved, it was clear that a better version had to be developed, even though its initial objective was only to anonymize health data. This page describes the ShinyAnonymizer Enhanced Version, which incorporates more advanced features and solves problems from the initial release, to go beyond simple anonymization. The objective is to present ShinyAnonymizer as a modern and innovative comprehensive solution for handling health data while ensuring privacy. The study's research targets highlight a twofold emphasis: showcasing the development of an enhanced version of ShinyAnonymizer and recognizing and correcting issues in the healthcare business's privacy-preserving technologies currently in use. The study contributes to the present debate on privacy in health data management by achieving these objectives. It provides researchers and healthcare professionals with an effective strategy for balancing patient privacy and data availability.

As seen in Figure 6, the previously mentioned application provides a graphical tool with a simple user interface that allows data importation, various ways of anonymizing data, statistics, and visualizations.

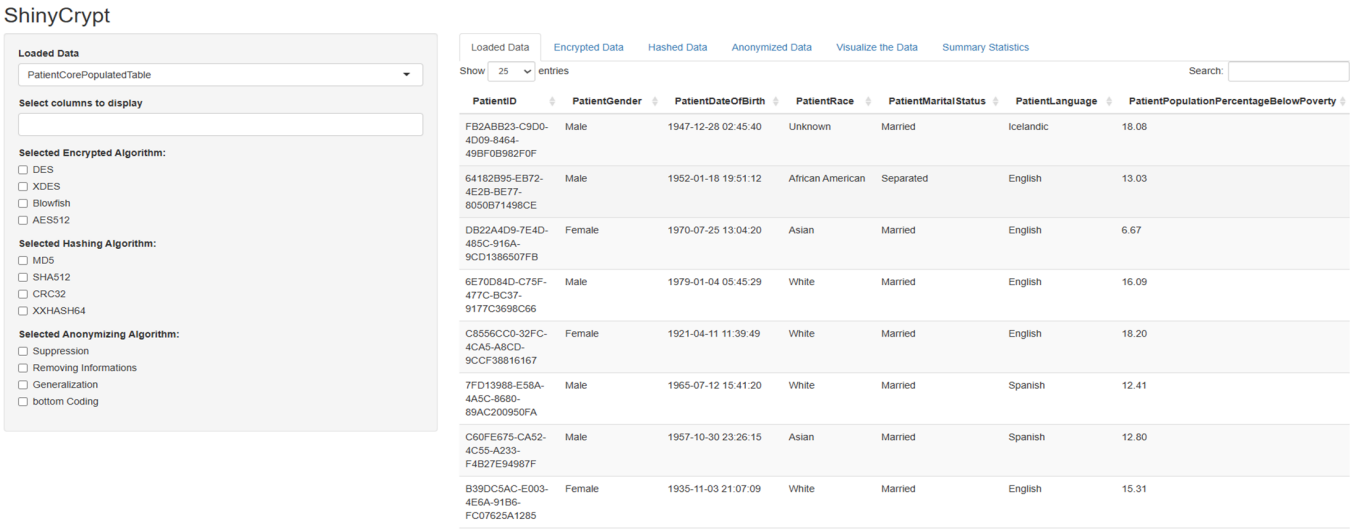

Figure 6. ShinyAnonymizer interface.

In conclusion, the enhanced ShinyAnonymizer offers an outstanding example of privacy-preserving technology, delivering an effective and user-friendly method for managing health data responsibly and securely.

7. Real World-Use Cases

The effectiveness of the ShinyAnonymizer Enhanced Version in maintaining patient privacy has been validated by its work in several real-life situations.

7.1. Clinical Research

ShinyAnonymizer is currently helpful in clinical research in anonymizing patient data. Without risking patient privacy, researchers can incorporate electronic health records, use modern anonymization techniques, and conduct comprehensive analyses. The tool's ability to work with various forms of data ensures that it can be utilized for multiple experimental scenarios.

7.2. Healthcare Analytics

ShinyAnonymizer is extremely useful to healthcare businesses that utilize data analytics for cost-effectiveness and improvements. The flexible anonymization features of the enhanced version facilitate the secure handling of big datasets, while ensuring compliance with privacy laws. Healthcare analytics teams can gain helpful knowledge, while protecting sensitive patient data by preserving data integrity and confidentiality [27].

7.3. Public Health Studies

Gathering data from multiple locations is a usual procedure in healthcare research. ShinyAnonymizer is utilized by health professionals due to its broad anonymization functions and adaptability to different data types. The tool secures users' identities, while helping the fast analysis and critique of the extracted data collection.

7.4. Collaborative Research Projects

Using ShinyAnonymizer, secure interaction occurs faster. Better verification is essential in collaborative studies when sharing data. Anonymized datasets may be shared between academics without worry about re-identification. Data analysis and statistics enable teamwork, to help scientists learn about anonymized data.

7.5. Environments for Learning

Regarding possible educational uses, ShinyAnonymizer could be a valuable resource for students with a talent for manipulating data and health analytics. It links students to real-world health data and helps them learn critical management and anonymization skills. This application benefits people by giving them more real-world experience and boosting their chances of success in healthcare data management. Ultimately, the ShinyAnonymizer Enhanced Version shows its significance when improving healthcare data management. Its ability to be utilized in various scenarios shows how flexible and effective it is at ensuring privacy, while delivering intelligent data analysis. It seems helpful in this kind of scenario for scientists, businesses, and healthcare professionals who have trouble managing sensitive patient data.

Managing all kinds of health data is an essential part of its use. Due to various mechanisms, types of data, and guidelines, electronic medical data can be complex. The ShinyAnonymizer Enhanced Version's usage in the current healthcare sector is increased by its ability to manage and anonymize organized and unorganized data. A notable feature of its use is the tool's ability to manage multiple kinds of health data. Due to their distinctive features, particular data types and standards in health data systems can be complex. Lastly, the fully interactive and simple UI of the ShinyAnonymizer Enhanced Version contributes to its greater usefulness in real-world scenarios. Scientists and healthcare professionals taking part in exploring data, statistical evaluation, and clear visualizations might encourage educated, informed decision-making. It helps effectively communicate study results by creating attractive and practical visuals.

8. Discussion and Implications in the Contents of Privacy Regulations

The real-world applications of the ShinyAnonymizer Enhanced Version may significantly affect privacy rules for individuals in managing health data storage and use. The tool allows adherence to major confidentiality laws by fully anonymizing patient data, such as the General Data Protection Regulation (GDPR) in the European Union and the Health Insurance Portability and Accountability Act (HIPAA) in the United States. They include strong protection processes and the careful storage of sensitive patient data. The upgraded ShinyAnonymizer offers an easy method for anonymizing health data, limiting legal issues and concerns. These uses include business, academia, medical studies, and scientific fields, showing the worth of privacy and confidentiality. Regulations additionally require medical facilities to establish security processes for handling modifications that impact personal data. The ShinyAnonymizer Enhanced Version helps businesses establish guidelines that prevent illicit conduct and guarantee secure data processing in every circumstance. Adherence to these rules restricts risk while enhancing trust. The enhanced version conforms with privacy rules by integrating modern privacy-preserving methods such as hashing, encryption, and anonymization. It conforms with HIPAA and GDPR criteria, covering the requirement for comprehensive security against data re-identification. It highlights the tool's essential role in ensuring privacy rules while maintaining program accuracy, using a complete security plan that utilizes numerous methods.

The tool's user-friendly design lets investigators, healthcare facilities, and educational institutions collaborate rapidly with anonymized data, while adhering to ethical and privacy requirements. The implications of the ShinyAnonymizer Enhanced Version extend through the data lifecycle, impacting data security practices from collection to elimination. Privacy legislation frequently defines how data should be collected, handled, used, and adequately eliminated at every stage of their duration. Businesses may limit the chance of legal issues or negative reputations due to data breaches via health data security measures. The enhanced version has a simple graphical user interface encouraging responsible and ethical data management. It permits greater data-processing adaptability and indicates businesses' responsibility for maintaining privacy rights. The tool's user-friendly design lets investigators, healthcare facilities, and educational institutions collaborate rapidly with anonymized data, while adhering to ethical and privacy requirements.

9. Results and Experiments

With its modern visual appearance, the ShinyAnonymizer Enhanced Version offers a strong basis for testing data and genuinely analyzing the results. Relationships, patterns, and trends could be illuminated by researchers using a variety of visualization tools that allow for a deep understanding to be drawn from anonymized datasets. The concept can alter data structures, unveil new possibilities, and articulate research findings through dynamic dashboards and easy-to-access graphs. Data models can undergo continuous review, which can lead to the improved significance and clarity of the results from time to time. Furthermore, the features in the dataset are comprehensively assessed through estimation and statistical data presentations such as mean, variance, and median, alongside graphical depictions. Regulations also need security processes for healthcare facilities to modify personal data. In all situations, the ShinyAnonymizer Enhanced Version allows businesses to formulate rules that prevent illicit conduct and ensure the safe processing of data. Thus, following these regulations helps reduce risks and build trust. The advanced version complies with data protection laws and is integrated with advanced privacy-preserving techniques like hashing, encrypting, and data anonymization methods. It satisfies HIPAA and GDPR requirements, including comprehensive security against re-identifying stolen information.

9.1. Case Study: Clinical Study on Patient Attributes

This clinical study aims to see how PatientGender, PatientRace, PatientMaritalStatus, and PatientLanguage relate to each other through different patient attributes. To examine these relationships, we created three kinds of visualizations as follows:

1. *Box plot:* The patient age (used from PatientId) is displayed for different marital status categories in this visual. Age numbers are displayed on the y-axis, and marital statuses (including single, married, and divorced) are shown on the x-axis. Descriptive statistics, comprising the mean and median ages, are given for every marital status group, to highlight the overall trends. The age range for every marital status group is examined utilizing the following statistical measures (1), (2):
   - Mean (m):
   $$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$
   where N is the number of patients in a marital status group, and x_i is the age of the i-th patient.
   - Median (M):
   $$M = \begin{cases} {}^{x}\left(\frac{N+1}{2}\right), & if\ N\ is\ odd \\ \frac{{}^{x}\left(\frac{N}{2}\right)+{}^{x}\left(\frac{N}{2}+1\right)}{2}, & if\ N\ is\ oven \end{cases} \tag{2}$$
   where $x(i)$ shows the i-th patient order statistic (i.e., the lowest value in the sorted data) [28].
2. *Bar plot:* This demonstrates how patients are divided by gender. Gender classifications (male and female) appear on the x-axis, while the number of patients is displayed on the y-axis. Every line shows the number of patients in every gender group. Distribution by gender can be obvious with the help of other statistics, like the percentage of male and female patients. The overall percentage indicates an average gender distribution, while the median number shows the central trend [29,30]. The total number of patients by gender is determined as follows (3):
   - Percentage (P):
   $$P = \left(\frac{n_g}{N}\right) \times 100 \tag{3}$$

3. *Area plot:* It displays how patient language accessibility has evolved. The x-axis shows different periods (years, for example), as the y-axis represents the number of patients understanding each language. The plot's sections each indicate one language, while the interconnected parts show modifications to the distribution of every language. Those are the mean and median number of patients speaking each language for every observation period [31]. The transformation in language distribution is examined utilizing Formulas (4) and (5):
   - The mean number of patients understanding a language (μL) during a period:

$$\mu_L = \frac{1}{T}\sum_{t=1}^{T} n_{L,t} \tag{4}$$

   where T is the total number of periods and $n_{L,t}$ is the number of patients who understand the language L at time t.
   - The median number of patients understanding a language:

$$M_L = \begin{cases} n_{L,}\left(\frac{T+1}{2}\right), & if\ T\ is\ odd \\ \frac{n_{L,}\left(\frac{T}{2}\right)+n_{L,}\left(\frac{T}{2}+1\right)}{2}, & if\ T\ is\ oven \end{cases} \tag{5}$$

   where $n_{L,(t)}$ displays the number of patients who understand language L at the t-th highest statistic over a specific period.

By integrating descriptive data with visual aids, we can identify patterns of language choice throughout time and understand modifications to language distribution at different times. The visuals illustrate changes in speech patterns and reveal prevalent language choices among the patients. Figures 7 and 8 demonstrate a dashboard that includes these illustrations and explanations.

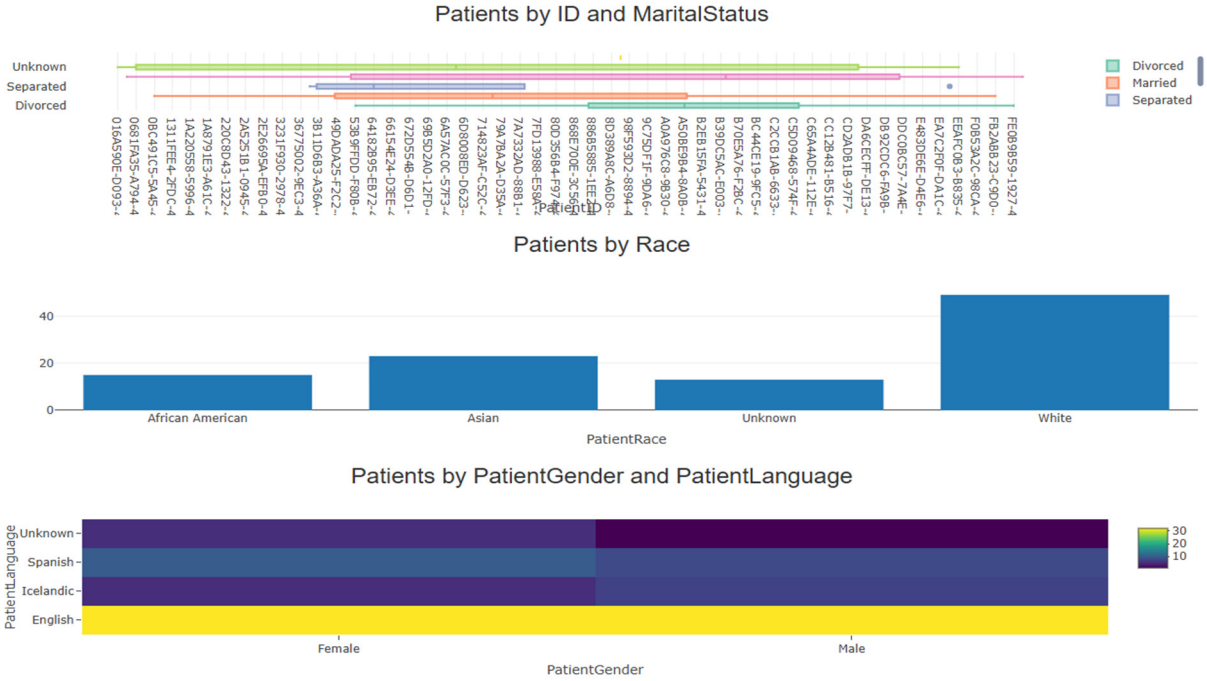

Figure 7. Graphical representation of patients.

```
 PatientID         PatientGender      PatientDateOfBirth                PatientRace        PatientMaritalStatus PatientLanguage
Length:100         Length:100         Min.   :1921-03-26 14:38:51.7999   Length:100         Length:100           Length:100
Class :character   Class :character   1st Qu.:1944-11-02 13:55:33.7189   Class :character   Class :character     Class :character
Mode  :character   Mode  :character   Median :1958-03-04 19:43:17.3429   Mode  :character   Mode  :character     Mode  :character
                                      Mean   :1957-03-31 23:35:33.6269
                                      3rd Qu.:1969-12-16 07:14:43.6652
                                      Max.   :1988-11-25 02:59:36.3730
PatientPopulationPercentageBelowPoverty
Min.   : 1.70
1st Qu.:12.70
Median :15.44
Mean   :22.10
3rd Qu.:18.69
Max.   :98.40
```

Figure 8. Descriptive statistics of patients.

Furthermore, taking advantage of three different visualization types—a box plot, a bar plot, and an area plot—and descriptive statistics, including the mean and median, we aim to investigate the data in-depth. Due to the proposed visualizations and descriptive statistics, we can identify interesting developments and trends in the data, including the age distribution of patients throughout marital status groups, the gender distribution of patients, and the change in language support over time. We might discover specific trends or challenges affecting how products and offerings are applied in the healthcare sector using these statistical methods. Also, by training elected officials, academics, and doctors, this research lets them make better choices and boosts patients' health.

9.2. Comparison with Other Works

Several options, such as open-source software, provide data anonymization. SdcMicro, Anonymizer, and ARX are among the most popular ones, so far as we are concerned. SdcMicro attempts to protect data from illegal access by focusing on data anonymization, which can then be utilized to create anonymized (micro) data or documents for general and research usage. It provides several risk evaluation methods and uses interconnected statistics to analyze the loss of data and its impact on constant variables. The SdcMicro package presents anonymization algorithms and expansive risk and productivity evaluations. It additionally provides tools for calculating, evaluating, and displaying risk and utility at different stages of the anonymization process. Still, SdcMicro lacks user-friendly interfaces, making it challenging for non-experts to use well. Its poor flexibility and scalability could become restrictive barriers for applications on a large scale [32].

Anonymizer is an additional anonymization tool which enables users to rapidly anonymize data, especially personally identifiable information (PII), using "convenience functions" or a combination of salting and hashing. While it presents an easy method for anonymization, it fails to offer complex encryption or risk evaluation methods. It was created for fast and simple anonymization tasks but fell down poorly when handling more data or being required to offer thorough analyses and reporting [27].

ARX aims for both versatility and accessibility in data anonymization. It enables several anonymization methods, quality of data analysis methods, and re-identification methods for reducing risks. Also, ARX utilizes standard anonymization methods like differential privacy, l-diversity, t-closeness, and k-anonymity. ARX offers reasonable speed and flexibility, so it fits various tasks. Still, its complex interface is an issue for people with no technical expertise, and it can demand considerable effort to install and use it effectively [26]. Table 7 mentions the comparison of those characteristics concerning the given aspect with ✔ while ✘ indicates the characteristic with its absence. As an example, ShinyAnonymizer-enhanced version and Anonymizer support "Data Encryption" and "Hashing", but SdcMicro and ARX do not. Similarly, in terms of "User Interface" being user-friendly or complex other features such as flexibility, scalability, performance is rated accordingly.

Table 7. Comparison with other works.

Aspect	ShinyAnonymization Enhanced Version	SdcMicro	Anonymizer	ARX
Data Anonymization	High	Moderate	Moderate	High
Data Encryption	✔	✘	✔	✘
Hashing	✔	✘	✔	✘
User Interface	User Friendly	Complex	User Friendly	Complex
Flexibility	High	Low	Moderate	✔
Scalability	High	Low	Moderate	✔
Performance	High	Low	Moderate	✔
Compliance with Laws	✔	✔	✔	✔

The enhanced ShinyAnonymizer shines thanks to its complex functions, such as data anonymization, encryption, and hashing, in addition to its user-friendly interface and conformity to privacy laws. With SdcMicro, ShinyAnonymizer is very scalable and adaptable, making it suitable for an extensive range of uses, from a smaller-scale examination to massive healthcare analytics. Its simple User Interface enables those without technical expertise to discover and utilize its features, an enormous advantage over ARX's complex setup. ShinyAnonymizer's powerful encryption methods guarantee that data remain protected during the anonymization procedure, a critical need that Anonymizer fails to meet. This tool additionally exceeds their speed, successfully managing big datasets while preserving a high precision. ShinyAnonymizer helps individuals understand and analyze data via extensive data analysis charts and visualization options, resulting in more reliable choices and discoveries. The ShinyAnonymizer Enhanced Version generally distinguishes itself from its rivals by integrating powerful anonymization algorithms, encryption, user-friendly interfaces, and speed and flexibility. This makes it an excellent choice for businesses that seek a comprehensive and productive solution for dealing with sensitive data while conforming to privacy requirements.

10. Beyond ShinyAnonymizer: Future Directions

In addition to what ShinyAnonymizer is capable of right now, several exciting potential routes could potentially enhance privacy-preserving technology in managing sensitive data:

10.1. Machine Learning and AI

Initially, the growth of more complicated anonymization techniques is made achievable by improvements in machine learning and artificial intelligence (AI). Future anonymization technologies will use AI algorithms to precisely and effectively locate and cover sensitive data in large datasets. This would assist in minimizing privacy breaches, while maintaining the data's analytic utility. The effectiveness of AI in data anonymization can be inferred by the following Formula (6) for data utility (U):

$$U = 1 - \frac{L_{anonymized} - L_{original}}{L_{original}} \tag{6}$$

where $L_{original}$ is a missing function for the initial data, and $L_{anonymized}$ is the missing function for the anonymized data. This formula determines how effectively the anonymized data maintains the significance of the original data.

10.2. Blockchain Technology

In addition, blockchain technology could significantly improve privacy and secure health information in medical settings. In this way, organizations using blockchain-based methods can set up open and unchangeable verification systems for receiving and changing health-related data. Adopting a hybrid method using blockchain technology and ShinyAnonymizer may add extra protection against unauthorized access or the modifica-

tion of private details [33]. The following are several representations that symbolize the integrity of a blockchain (7):

$$I = \sum_{i=1}^{N} \left(\frac{1}{H_i} \right) \tag{7}$$

where N is the total number of blocks and H_i is the hash of the i-th block that promises the integrity of every block.

10.3. Differential Privacy

Incorporating noise into the responses to inquiries ensures that individual pieces of data are rendered unreadable, while still allowing the extraction of pertinent insights. The anonymization systems of the future may consist of differential privacy models that provide more enhanced forms of user privacy assurance [34]. The mathematical expression for differential privacy can be represented as follows (8):

$$M(D) = f(D) + Lap\left(\frac{\Delta f}{\in} \right) \tag{8}$$

where $M(D)$ is the private differential value for dataset D, $f(D)$ is the name of the query function, Δf is the function's sensitivity level, and ϵ is the privacy constraint.

10.4. Federated Learning

Federated learning presents a novel way of analyzing data, while retaining privacy. Machine learning models are developed across distributed machines or computers in federated learning systems without transmitting data. This enables businesses to work collaboratively on data analyses, while preserving the confidentiality and localization of sensitive data. Later versions of anonymization software might include federated learning characteristics that promote collaboration studies without compromising personal privacy [35]. The federated learning Formula (9) can be defined as follows:

$$\min_{w} = \left(\frac{1}{N} \sum_{i=1}^{N} L_i(w) \right) \tag{9}$$

where N is the total quantity of clients, $L_i(w)$ is the local loss rate for client i, and w is the set of variables.

10.5. Homomorphic Encryption

Homomorphic encryption preserves confidentiality while allowing functional analysis, because it will enable computations on encrypted data. For example, ShinyAnonymizer employs homomorphic encryption and anonymization tools to allow businesses to analyze sensitive data without exposing them to the risk of leakage and unauthorized access [36]. Abbreviations for homomorphic encryption can be defined as in (10):

$$E(f(x)) = f(E(x)) \tag{10}$$

where E is the encryption function, f is the calculation function, and x is the plaintext data. This ensures that calculations can be carried out utilizing encrypted data.

In conclusion, there is an enormous demand for creativity and development in privacy-preserving ideas for handling data. Future versions of anonymization tools like ShinyAnonymizer might provide more robust promises of privacy protection, while permitting critical data analysis and cooperation in healthcare as well as beyond, by going after methods like artificial intelligence (AI), blockchain, differential privacy, federated learning, and homomorphic encryption.

11. Conclusions

Creating and deploying privacy-preserving technologies like ShinyAnonymizer is an essential first step in legally managing sensitive data, especially in healthcare. Our investigation examined ShinyAnonymizer's primary features, enhanced study methods, technical limitations, and possible uses. We also reported how privacy-preserving technology will grow in the years to come and examined how this impacts privacy rules. With strong hashing, encryption, and anonymization algorithms, ShinyAnonymizer addresses a critical need to maintain patient privacy, while enabling intelligent data analysis and cooperation, and balancing the utility of data and security. Its beneficial utilization in epidemiology, clinical trials, group research, healthcare analysis, and educational institutions shows flexibility. Establishing respect for security laws like GDPR and HIPAA further builds trust in data-handling practices. In future years, the options for ethical data management and collaboration will probably rise with advances such as homomorphic encryption, blockchain, artificial intelligence, federated studying, and differential privacy. Since ShinyAnonymizer ensures that private data is dealt with in a manner that shows attention to detail, truthfulness, and respect, its continual growth and usage will be necessary in affecting the development of healthcare and its associated domains. Ultimately, ShinyAnonymizer is a giant leap forward in creating privacy-preserving data management tools essential for advances in the healthcare industry and beyond.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

1. Shojaei, P.; Vlahu-Gjorgievska, E.; Chow, Y.-W. Security and privacy of technologies in health information systems: A systematic literature review. *Computers* **2024**, *13*, 41. [CrossRef]
2. Almulihi, A.H.; Alassery, F.; Khan, A.I.; Shukla, S.; Gupta, B.K.; Kumar, R. Analyzing the Implications of Healthcare Data Breaches through Computational Technique. *Intell. Autom. Soft Comput.* **2022**, *32*, 1763–1779. [CrossRef]
3. Kondylakis, H.; Despoina, M.; Glykokokalos, G.; Kalykakis, E.; Karapiperakis, M.; Lasithiotakis, M.-A.; Makridis, J.; Moraitis, P.; Panteri, A.; Plevraki, M.; et al. EvoRDF: A framework for exploring ontology evolution. In *The Semantic Web: ESWC 2017 Satellite Events, Proceedings of the ESWC 2017 Satellite Events, Portorož, Slovenia, 28 May–1 June 2017*; Revised Selected Papers 14; Springer International Publishing: Cham, Switzerland, 2017. [CrossRef]
4. Adeniyi, A.O.; Arowoogun, J.O.; Okolo, C.A.; Chidi, R.; Babawarun, O. Ethical considerations in healthcare IT: A review of data privacy and patient consent issues. *World J. Adv. Res. Rev.* **2024**, *21*, 1660–1668. [CrossRef]
5. Bakare, S.S.; Adeniyi, A.O.; Akpuokwe, C.U.; Eneh, N.E. Data privacy laws and compliance: A comparative review of the EU GDPR and USA regulations. *Comput. Sci. IT Res. J.* **2024**, *5*, 528–543. [CrossRef]
6. Guide, A. *Cybersecurity Resource. Implementing the Health Insurance Portability and Accountability Act (HIPAA) Security Rule*; National Institute of Standards and Technology: Gaithersburg, MA, USA, 2024. [CrossRef]
7. Williamson, S.M.; Prybutok, V. Balancing privacy and progress: A review of privacy challenges, systemic oversight, and patient perceptions in ai-driven healthcare. *Appl. Sci.* **2024**, *14*, 675. [CrossRef]
8. Clayton, E.W.; Embí, P.J.; Malin, B.A. Dobbs and the future of health data privacy for patients and healthcare organizations. *J. Am. Med. Inform. Assoc.* **2023**, *30*, 155–160. [CrossRef] [PubMed]

9. Vardalachakis, M.; Kondylakis, H.; Koumakis, L.; Kouroubali, A.; Katehakis, D. ShinyAnonymizer: A Tool for Anonymizing Health Data. In Proceedings of the 5th International Conference on Information and Communication Technologies for Ageing Well and e-Health, Crete, Greece, 2–4 May 2019; pp. 325–332.
10. Paul, M.; Maglaras, L.; Ferrag, M.A.; Almomani, I. Digitization of healthcare sector: A study on privacy and security concerns. *ICT Express* **2023**, *9*, 571–588. [CrossRef]
11. Dhirani, L.L.; Mukhtiar, N.; Chowdhry, B.S.; Newe, T. Ethical dilemmas and privacy issues in emerging technologies: A review. *Sensors* **2023**, *23*, 1151. [CrossRef] [PubMed]
12. Hossain, M.T. Privacy and Security for Trustworthy AI/ML in Multi-Agent Critical Infrastructures: An Analysis of Adversarial Dynamics and Protective Strategies. Ph.D. Dissertation, University of Nevada, Reno, NV, USA, 2024. Available online: https://scholarwolf.unr.edu/home (accessed on 4 August 2024).
13. Paramesthi, P.; Jati, S.P.; Suryoputro, A. The use of Electronic Medical Record (EMR) in hospitals during the COVID-19 pandemic in Indonesia: A systematic literature review. *BKM Public Health Community Med.* **2024**, *40*, e11727. [CrossRef]
14. Vardalachakis, M.; Kondylakis, H.; Tampouratzis, M.; Papadakis, N.; Mastorakis, N. Anonymization, Hashing and Data Encryption Techniques: A Comparative Case Study. In Proceedings of the 2023 International Conference on Applied Mathematics & Computer Science (ICAMCS), Lefkada Island, Greece, 8–10 August 2023; pp. 129–135.
15. Marques, J.; Bernardino, J. Analysis of Data Anonymization Techniques. In Proceedings of the 12th International Conference on Knowledge Engineering and Ontology Development, Budapest, Hungary, 2–4 November 2020; pp. 235–241.
16. Turgay, S.; İlter, I. Perturbation methods for protecting data privacy: A review of techniques and applications. *Autom. Mach. Learn.* **2023**, *4*, 31–41. [CrossRef]
17. Vovk, O.; Piho, G.; Ross, P. Methods and tools for healthcare data anonymization: A literature review. *Int. J. Gen. Syst.* **2023**, *52*, 326–342. [CrossRef]
18. Hasan, H.A.; Al-Layla, H.F.; Ibraheem, F.N. A Review of Hash Function Types and their Applications. *Wasit J. Comput. Math. Sci.* **2022**, *1*, 75–88. [CrossRef]
19. Kumar, K.K.; Ramaraj, E.; Srikanth, B.; Rao, A.S.; Prasad, P. Role of MD5 Message-Digest Algorithm for Providing Security to Low-Power Devices. In Proceedings of the 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 25–27 May 2022; pp. 352–358.
20. Al-Shareefi, F.; Al-Barmani, Z. Comparing two cryptographic hash algorithms: SHA-512 and whirlpool—A case study on file integrity monitoring. *BIO Web Conf.* **2024**, *97*, 00093. [CrossRef]
21. Yusuf, A.D.; Abdullahi, S.; Boukar, M.M.; Yusuf, S.I. Collision resolution techniques in hash table: A review. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 757–762. [CrossRef]
22. Joshua, T. A Secure Model for Student Results Verification Using Salted Hash Functions. Master's Thesis, Kampala International University, Kampala, Uganda, 2023. Available online: https://irbackend.kiu.ac.ug/server/api/core/bitstreams/c74b76c4-3f75-4393-990a-2ce5638b2ad0/content (accessed on 4 August 2024).
23. Kuznetsov, O.; Peliukh, O.; Poluyanenko, N.; Bohucharskyi, S.; Kolovanova, I. Comparative Analysis of Cryptographic Hash Functions in Blockchain Systems. In Proceedings of the CPITS-2023-II: Cybersecurity Providing in Information and Telecommunication Systems, Kyiv, Ukraine, 26 October 2023; pp. 81–94. Available online: https://ceurspt.wikidata.dbis.rwth-aachen.de/Vol-3550/paper7.html (accessed on 4 August 2024).
24. Atadoga, A.; Farayola, O.A.; Ayinla, B.S.; Amoo, O.O.; Abrahams, T.O.; Osasona, F. A Comparative Review of Data Encryption Methods in the USA and Europe. *Comput. Sci. IT Res. J.* **2024**, *5*, 447–460. [CrossRef]
25. Rameel, M.; Asif, Z. Fortifying Information Security: A Comparative Analysis of AES, DES, 3DES, RSA, and Blowfish Algorithm. *Communications* **2024**, *2*, 5.
26. Prasser, F.; Eicher, J.; Spengler, H.; Bild, R.; Kuhn, K.A. Flexible data anonymization using ARX—Current status and challenges ahead. *Softw. Pract. Exp.* **2020**, *50*, 1277–1304. [CrossRef]
27. Hendricks, P. Anonymizer: Anonymize Data Containing Personally Identifiable Information. R Package Version 0.2.0. 2015. Available online: https://github.com/paulhendricks/anonymizer (accessed on 4 August 2024).
28. Newbold, P.; Miller, W.L.; Thorne, R. *Statistics for Business and Economics*, 9th ed.; Pearson Education Limited: Harlow, UK, 2013.
29. Tukey, J.W. Exploratory Data Analysis. Addison-Wesley Publishing Company Reading, Mass.—Menlo Park, Cal., London, Amsterdam, Don Mills, Ontario, Sydney 1977, XVI, 688 S. *Biom. J.* **1981**, *23*, 413–414. [CrossRef]
30. Healy, K. *Data Visualization: A Practical Introduction*; Princeton University Press: Princeton, NJ, USA, 2018.
31. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*, 4th ed.; Springer Nature: Dordrecht, The Netherlands, 2017.
32. Templ, M.; Kowarik, A.; Meindl, B. Statistical disclosure control for micro-data using the R package sdcMicro. *J. Stat. Softw.* **2015**, *67*, 1–36. [CrossRef]
33. Vaigandla, K.K.; Karne, R.; Siluveru, M.; Kesoju, M. Review on blockchain technology: Architecture, characteristics, benefits, algorithms, challenges and applications. *Mesopotamian J. CyberSecurity* **2023**, *2023*, 73–85. [CrossRef]
34. Vasa, J.; Thakkar, A. Deep learning: Differential privacy preservation in the era of big data. *J. Comput. Inf. Syst.* **2023**, *63*, 608–631. [CrossRef]

35. Wen, J.; Zhang, Z.; Lan, Y.; Cui, Z.; Cai, J.; Zhang, W. A survey on federated learning: Challenges and applications. *Int. J. Mach. Learn. Cybern.* **2023**, *14*, 513–535. [CrossRef] [PubMed]
36. Munjal, K.; Bhatia, R. A systematic review of homomorphic encryption and its contributions in healthcare industry. *Complex Intell. Syst.* **2023**, *9*, 3759–3786. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.