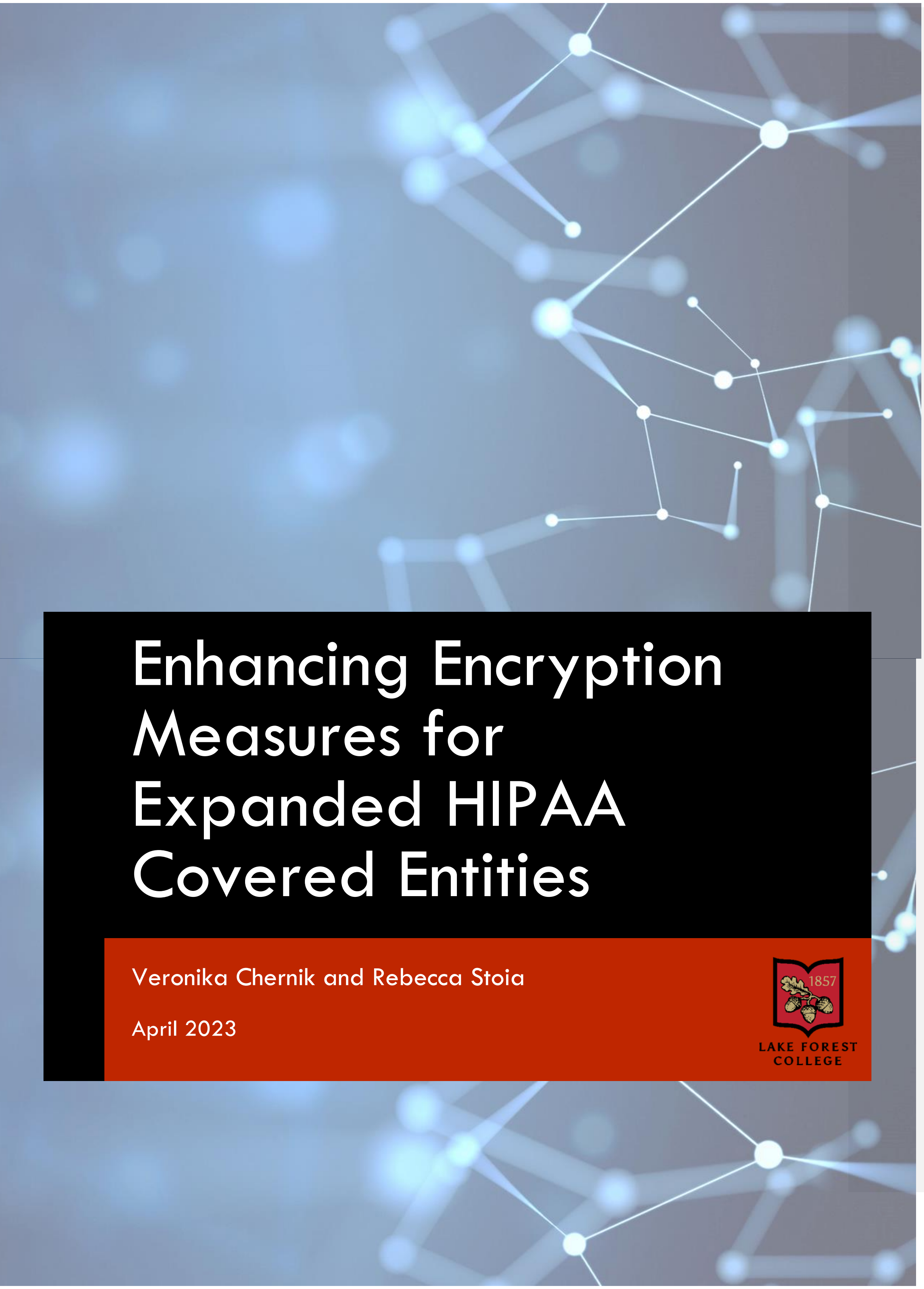


# Enhancing Encryption Measures for Expanded HIPAA Covered Entities

Veronika Chernik and Rebecca Stoia

April 2023

1857
LAKE FOREST
COLLEGE

*Thank you to our mentor, Jeffrey Sundberg for challenging us with difficult perspectives and always supporting us with your thorough feedback. Thank you for your immense support and guidance, we are privileged to have learned so much from your contributions.*

# CONTENTS

# FORWARD

Establishing standards for protecting individuals' health information will provide a safer and more secure way for people to access medical care.

**Veronika Chernik**
Lake Forest College '23

**Rebecca Stoia**
Lake Forest College '24

As technology becomes a more prominent tool in accessing healthcare, it is crucial to consider HIPAA's role in data privacy. Neglecting to address the major loopholes in this regulation will continue to allow disparities in who receives quality and confidential healthcare. Additionally, we risk limiting the implementation of a more secure way to store valuable and personal data. Such oversights hinder progress and pose an imminent threat to data privacy and protection.

An update of HIPAA should be a high priority to remain relevant and effective in protecting sensitive information while also benefiting the greater community through supported medical research, fluid data transmission, and limited data collection from predatory third-party sources.

Technology continues to advance, and regulation is severely behind. As a result, patient safety is thoroughly neglected when online searches become more common than doctor visits, and data breaches frequently occur due to outdated security measures.

Mobile devices and wearable systems are used to collect, store, and transmit health data in individual and medical settings. In addition, data stored on cloud-based servers face issues with data security, ownership, and control, posing the question of "Who *really* has rights to 'my' data?" Lastly, the rise of AI and machine learning algorithms raises concerns about the proper security of historically de-identified data.

# EXECUTIVE SUMMARY

The Health Insurance Portability and Accountability Act (HIPAA), signed in 1996, does not adequately protect individuals' sensitive health information. With the rise of healthcare offerings online, redefining what organizations must be covered by the act to ensure patient data security is essential. Apps, websites, and other aggregate health data sources online are not covered under HIPAA and have no comprehensive standardized model for data security or privacy protections. Additionally, once patient data has been subject to de-identification (as defined by HIPAA), it is no longer protected under the regulation. This is a problem as current anonymizing standards are insufficient in withstanding advancing re-identification methods, leaving personal information vulnerable to data breaches and fraud.

It is crucial to broaden HIPAA's reach in regulation. Today, HIPPA only covers the traditional health sector and a few organizations that manage direct patient data. Over the years, there have only been minor changes in adapting coverage to the growing digital age. With technology often used as the first line of support for families who cannot afford conventional healthcare services, these resources must be included in the privacy rights and protections promised by HIPAA.

Further, the sophistication level of machine learning algorithms outpaces current regulations. As a result, individuals and their personal identifiers can quickly be re-identified with exceptionally high accuracy, even with limited data available. While this information is frequently utilized in its anonymized state, after conversion, it falls outside the scope of HIPAA's coverage that aims to protect individual privacy rights. Therefore, it is essential to address the current anonymizing process and establish unwavering data protections and transparency throughout the lifetime of identifiable patient data.

A digital data ledger is proposed to implement an improved data collection and storage standard. It will house patient data from virtually infinite health-related inputs to construct a thorough and secure build of individual patient information. Each user has a personalized key that provides access to their comprehensive data. Each organization has a modified version of that key that only unencrypts data they collected or additional data with verified consumer consent. Besides the hospital setting, researchers, verified by a central audit team, can request a key to unencrypt specific data points to construct aggregate datasets. The Office for Civil Rights (OCR), currently responsible for managing HIPAA

compliance, will issue guidance for ledger verifications.

This ledger has dynamic learning technology that can be easily verified and synchronized to form large, encrypted data sets for medical research and optimized patient care. To increase data transparency and security, patients can track, correct, or remove an organization's access to their data. In addition, enforcement of data ownership and transfers will be easier than ever due to the technologies in place to flag and report data misuse and HIPAA violations to the OCR for investigation. This means direct enforcement of HIPAA will not require substantial supplemental resources, even with the expanded entities. Included entities (hospitals, insurance companies, third parties, etc.) can send advocates to be steering committee members where ledger rules are established.

With the dynamic data ledger in place, individuals may log in to their secure account and verify data collected on them, check for misuse, and give or restrict sharing of their data. Ultimately, this will provide a unified and more efficient way to store and share health information– the fundamental focus of HIPAA.

Upfront investments for a unified ledger are expected to be higher, yet reasonable, as almost all health organizations are already committed to digitally storing data. This collective storage will be funded by a per-access or per-individual subscription model to firms, on par with current HIPAA data storage compliance costs. Further, implementing a universal ledger system is cheaper than regulating this practice for each entity, particularly from an enforcement standpoint. Therefore, there is an expected high return on investment as administrative, enforcement, and maintenance costs will be greatly cut by this ledger. This feature also makes tracking policy successes or shortcomings easier, evaluated annually based on consumer reflection and entity compliance rates.

For consumers not interested in the ledger, there is no added burden on the individual. Yet, as awareness and concern about privacy protections grow, it is a safer and more convenient way to increase data transparency between organizations and consumers and ensure secure data transfer ability.

# 1 | INTRODUCTION

## Dangerous Loopholes in Medical Privacy Law

Almost three decades ago, the Health Insurance Portability and Accountability Act (HIPAA) was passed to help employees maintain health insurance coverage after a job change or loss. Today, HIPAA is most referred to for its consumer-forward privacy rule, regulating who can see an individual's medical data and allowing patients to consent to share it with other parties. However, HIPAA's primary objective regards portability: to protect an individual's Personal Health Information (PHI) as it is stored, transmitted, and accessed by covered institutions. Although HIPAA has improved the security and privacy of health data, the Act has two problematic shortcomings: HIPAA's privacy law only applies to the stated covered entities, and there is little protection for de-identified data security.

Patients receive a HIPAA Compliance Agreement upon entering a hospital or clinic by federal law, stating that the patient is willing to disclose their medical history to their current provider. Although signing the agreement is not required to receive care, medical institutions must document that they provided the document. Additionally, the Privacy Rule requires covered entities under HIPAA to provide access to the data recorded to the verified individual under request so that the individual may always have the right to view, copy, or personally transfer their personal information. In some cases, personal information used for quality assessment and general business decisions is not available for the patient to access.

Before addressing HIPAA's privacy law, it is important to consider the already-present issues in medical data practices. One major problem in data efficiency is the lack of organization from various sources, formats, and usefulness. There are numerous duplications of documents and inaccurate inputs. Additionally, poor communication between sources causes inefficiency in patient care, resource utilization, and quality of diagnoses.[1] Our solution will aim to address this problem as well as the targeted HIPAA privacy loopholes.

HIPAA's privacy rule regulates who can access and receive a patient's PHI, including electronic, written, and spoken statements. Data that falls under the category of PHI includes details about one's health status, history, or payments, which are individually identifiable. Since HIPAA only claims to protect data that falls under the category of PHI, any information that is not considered personally traceable, referred to as de-identified data, cannot be held accountable under this law.

Therefore, de-identified data serves as an ambiguous loophole for obligatory compliance, limiting the protection of confidential and sensitive personal data.[2]

Another clause of HIPAA refers to the policy's coverage. The Act only targets covered entities, defined as "a health plan, a health care clearinghouse, or a health care provider who transmits any health information in electronic form..."[3] Individual healthcare providers and medical establishments include doctors, psychologists, dentists, clinics, and nursing homes. Health plans encompass insurance companies, company health plans, and certain government medically related programs. Finally, healthcare clearing houses, such as organizations that convert health information to an electronic format, are also included. This description means that third-parties apps or online resources with access to personal health or medical information are not included in the regulation. In the age of more mental and physical health resources online, it begs the question: What makes these medical services less inclined to the necessary privacy protections than traditional medical settings?

## HIPAA's Regulatory Terrain in Privacy Protection

Mobile health apps such as MyFitnessPal (health and nutrition), Headspace (mental health), and Flo (period and pregnancy tracker) are popular health trackers and resources amongst expensive insurance plans and treatments. Many individuals resort to these websites and applications for faster access to information and treatment to reduce expenses and improve convenience in leading a healthier life. Utilizing a search engine to look up the symptoms of a disease or how concerning a symptom may be is a common first step before deciding to see a doctor and is a valuable guide to help understand the severity of the possible diagnosis. Although advice from virtual sources may not always be dependable, reputable institutions like WebMD or the National Institute of Health allow users to filter through available information while exploring possible home remedies. Even if that individual consistently consults with healthcare professionals within their network, a quick online search can facilitate a better understanding of their situation, allow them to read others' experiences, and discover preventative measures. This resource benefits families with limited access to conventional healthcare, either due to financial constraints or time limitations. Therefore, it should be regarded as a valuable tool for such individuals. However, similar to disclosing detailed personal symptoms at a medical clinic, utilizing this resource involves sharing personal information that may be stored and shared with third parties since these online tools are not subject to HIPAA regulations.

In the current era of technology, mobile or online-based companies have become ubiquitous sources of personal information. These apps collect highly sensitive personal information from users, including medical history, symptoms, and behavioral patterns, posing a significant risk to user privacy and security. In addition, the lack of regulation in this field leaves mobile or online-based companies free to sell or share their customers' data with third-party services, such as advertisers. This loophole in HIPAA presents a significant concern for protecting an individual's sensitive data since the only time any information from these apps is protected is when transferred to a covered entity, such as a primary physician.[3] Unfortunately, this issue is perpetuated by the fact that many individuals fail to read or fully comprehend the terms and conditions of these applications or online services.

The general terms and conditions or privacy policies of applications and websites are often long, intricate, and challenging to grasp. Most do not read these documents in their entirety as they are usually extremely impractical for consumers to read. This results in insufficient comprehension regarding the data collection, utilization, and distribution of their confidential information. Many companies also do not provide clear guidelines for the consumer's right to manage their data which severely limits data transparency between the individual and organization.[4]

According to a recent study that discovered and analyzed over 20,000 mobile health applications in the Google Play marketplace, 88% contained code that could collect user data such as location, contact information, and device identifiers.[5] This alarming figure highlights the vulnerability of personal information stored within these healthcare apps, especially since this data is not immune to data breaches. Furthermore, "28 percent of apps did not provide a privacy policy at all" which would normally outline what level of protection users could expect from the organization and their privacy rights.[5] Without a privacy policy, the application or website is not bound to how it can collect, use, or share personal information such as the user's email address, name, location, and browsing history. This absence of policy removes company transparency in where the data will go and will put users at risk of having their data sold or shared without their consent or knowledge.

Even out of the 72% of apps that did include a privacy policy in their Terms and Conditions, the burden falls on the consumer to read, understand, and decide if the level of privacy protection works for them. But if they disagree, what else is out there? Many apps are popular for a reason, whether it be from an established or reputable company, ability to connect with friends or family, or it has a particular feature that no similar product has. Even if the consumer did everything right, many applications and online services still fail to adequately follow their terms and continue selling or transmitting user data to other sources.[5]

The issue of data privacy and the selling of personal information is complex. Undoubtedly, unauthorized sales of sensitive personal data infringe on individual privacy rights; however, considering this in a broader context is necessary. The absence of federal regulations regarding collecting and selling personal data has enabled many companies to engage in these practices with little consequence. Therefore, the question arises: if these companies are prohibited from selling data, will they still be able to operate? The answer is yes, for many. Apps and websites that take advantage of the predatory data-selling approach do so because they simply can. There is no regulation on the federal level that prohibits the data collection and lucrative sales of personal data. Some online companies have found ways to monetize their services through other means, such as generic, non-user-targeted advertising. Still, the lack or insufficiency of regulatory measures means that users remain exposed to unconsented harvesting and potential misuse of their personal data.

## De-Identified Data is Not So Secure

The other key issue of the HIPAA agreement is the lack of protection for de-identified data, data where identifiable elements have been removed or hidden. Once data has gone through the de-identification process, it is no longer regulated by HIPAA, even though HIPAA's purpose is to care for individuals' PHI. This troubling loophole poses a few risks to an individual's privacy; lack of data protection neglects to trace the transfer or selling of de-identified data and cannot be penalized or controlled—when it can later be traced back to the individual. This lack of protection is an evident invasion of medical privacy and can lead to discrimination, such as higher health insurance premiums and societal harm to the individual. In addition, by neglecting to uphold secure privacy protections and confidentiality in medical data, patients will be more hesitant to provide honest and thorough disclosures of their sensitive information to their providers.[6] Distrust in the healthcare industry is incredibly unhelpful, particularly to those individuals reaching out in an already vulnerable state.

It is overly concerning that de-identified data can be easily re-identified, posing a significant risk to individuals' privacy. Shockingly, 87% of people in a de-identified dataset containing zip codes, gender, and date of birth could still be recognized.[7] Even more alarming is that 99.98% of the American population can be identified by having only fifteen demographic characteristics on hand.[8] This emphasizes the urgent need to discuss HIPAA's current de-identification standards to better understand the risk of reidentification. In the privacy rule, de-identification is deemed sufficient in two ways, through the Safe Harbor method, and by expert determination. First, according to the Safe Harbor encryption method, de-identified data must meet a list of 18 qualities. These qualities include removing names, dates, telephone, personal devices, social security numbers, emails, URLs, IP addresses, biometric identifiers, and other unique identifiers.[8] This list may sound comprehensive, but the Safe Harbor de-identification standards are not rigorous enough. For example, the method includes provisions where ZIP codes and dates can be included yet, considered encrypted. These features can be used as valuable keys in reidentification.

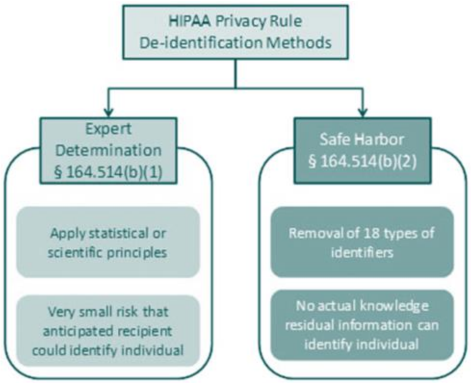

HIPAA Privacy Rule
De-identification Methods
Expert
Determination
§ 164.514(b)(1)
Safe Harbor
§ 164.514(b)(2)
Apply statistical or scientific principles
Very small risk that anticipated recipient could identify individual
Removal of 18 types of identifiers
No actual knowledge residual information can identify individual

Figure 1: HIPAA De-Identification Methods

The latter requirement, expert determination, involves the formal proclamation of encryption by a qualified member. This process includes consultation, and the application of scientific and statistical techniques to render the risk of data re-identification extremely small. The issue presented by expert determination is its lack of specificity. According to HIPPA law, an expert is defined as anyone who has experience in utilizing statistical methods of data de-identification. This individual does not need to possess any specific skills nor certifications. However, the expert may require an occasional audit by regulators.[9] Further, the method does not quantitatively define the level associated with small risk. The department of Health and Human Services (HHS) comments on this clause, reasoning that the appropriate risk level needs to be determined on a case-by-case basis.[9] The HHS thus claims that there is no defined level of security that is required across all health organizations. A final issue with expert determination, is its lack of expiry date. There exists no timeframe for which the declaration of de-identification needs to be reconsidered. This is a matter of concern, especially given the rate at which technology, artificial intelligence, and algorithms advance. Overall, the expert determination approach is too vague and can dangerously result in a myriad of de-identification dictates.

Considering HIPAA's standards for de-identification, one may wonder whether de-identification security concerns are ubiquitous. Security concerns may become more prevalent as technological expertise in artificial intelligence and algorithms continues to grow. For example, in 2019, Google was named in a lawsuit, which charged the tech company with failure to properly de-identify patient medical data.[10] In the class action complaint, Google was alleged to have uniquely identified almost every medical record due to its data mining and artificial intelligence capability. The suit called for an injunction and suggested that Google ask for consent when disclosing potentially identifiable information. Google's lawsuit indicates a more significant trend related to securing medical data. Health data is the most sensitive and intimate of all information, yet a national consensus regarding the risks of unauthorized disclosure is not set. Therefore, the HIPAA security rule must be adjusted to reflect a changing technological scene.

HIPAA's privacy rule is a vow to keep patient information private. Patient confidentiality is crucial, as it promotes honest communication, ensures quality care, and protects from discrimination, embarrassment, or economic harm.[11] The issue with de-identification is that it is imperfect, failing to protect individuals' privacy fully. De-identification further creates a free-market problem and hinders research. First, when HIPAA no longer protects data, it can be freely bought or sold. The aggregation and trade of deidentified Electronic Health Records (EHR) has become a multi-billion-dollar industry.[11] Such an industry is harmful to patients as they become the products being sold to data companies. For example, examine pharmaceutical detailing. The marketing technique uses data to target physicians who prescribe medicines, hoping to boost their purchasing and prescribing behavior. The result is an increase in drug use and prices.[11]

Further, de-identification hinders research and hinders a learning healthcare system. According to HIPAA law, patient data must be de-identified before it is used in research. The issue is that de-identification is costly, resulting in research institutions paying millions of dollars annually. The research ideal is to establish a learning healthcare system where data from medical events is aggregated and analyzed by institutions that seek to gain the knowledge necessary to improve patient care.[11] In other words, a learning healthcare system would act as a feedback loop, where past patient data is used to improve the experience of the next patient. This type of feedback would serve as a public good. Unfortunately, HIPAA's privacy rule is a barrier to creating a learning healthcare system. As a result, data is expensively de-identified, and de-identification reduces the research value. One potential solution involves changing the scope of medical data to include researchers. This idea will be elaborated on in the next section of this paper.

# Issues With Re-Identification

De-identified PHI is known for protecting patients' data while supporting research and academia's contributions to medical development and patient satisfaction. However, de-identification falls short of adequately protecting an individual's privacy. One study by the Journal of the American Medical Association (JAMA) tested the feasibility of re-identifying, de-identified PHI physical activity data. The re-identification involved using two machine learning methods: a linear support vector machine and an applied machine learning method called random forest.[12] The machine learning models took de-identified physical activity data and accurately matched the entries

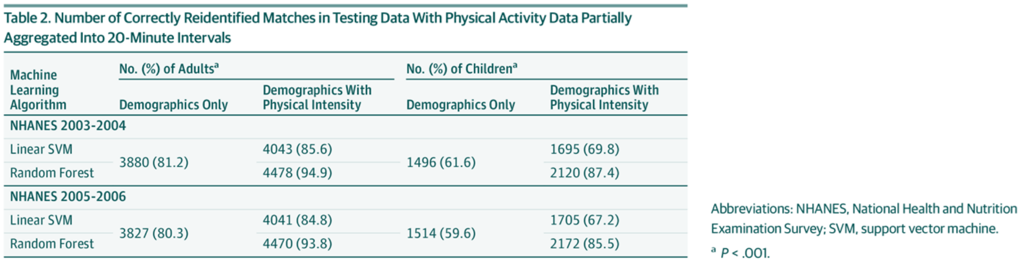


Table 2. Number of Correctly Reidentified Matches in Testing Data With Physical Activity Data Partially Aggregated Into 20-Minute Intervals

| Machine Learning Algorithm | No. (%) of Adults[a] |  | No. (%) of Children[a] |  |
|---|---|---|---|---|
|  | Demographics Only | Demographics With Physical Intensity | Demographics Only | Demographics With Physical Intensity |
| NHANES 2003-2004 |  |  |  |  |
| Linear SVM | 3880 (81.2) | 4043 (85.6) | 1496 (61.6) | 1695 (69.8) |
| Random Forest |  | 4478 (94.9) |  | 2120 (87.4) |
| NHANES 2005-2006 |  |  |  |  |
| Linear SVM | 3827 (80.3) | 4041 (84.8) | 1514 (59.6) | 1705 (67.2) |
| Random Forest |  | 4470 (93.8) |  | 2172 (85.5) |

Abbreviations: NHANES, National Health and Nutrition Examination Survey; SVM, support vector machine.

[a] $P < .001$.

Figure 2: JAMA Re-Identification Study

with their correct demographic and PHI details. The random forest algorithm correctly re-identified 94.9% of adults and 87.4% of children from the de-identified physical activity data set. The linear support vector algorithm also proved successful in re-identification, with 85.6% of adults and 69.8% of children accurately re-identified.[12] Machine learning algorithms can quickly and far too accurately match sensitive PHI details to the appropriate patient, removing the protective barrier de-identification claimed to offer.

Re-identifying patient data is rarely allowed under HIPAA. The most common exception to the rule occurs when medical providers need access to unencrypted PHI for patient treatment. If so, the data immediately becomes subject to HIPAA's intensive PHI safeguarding standards. Re-identification is a concern when attempted by cyber criminals. With modern technology, such as machine learning algorithms, malicious data re-identification increases the feasibility of compromising PHI privacy for a profit. HIPAA does not currently impose penalties on those who villainously re-identify health data since tools like machine learning are novel, and the threat is not yet on a large scale. The ease of re-identification, however, highlights the protection limitations of de-identified data. De-identification standards will always be in a race against re-identification technology.

# 2 | PROPOSED IMPLEMENTATIONS

## Ledger Technology Enhances HIPAA Compliance and Security

HIPAA has a few significant shortcomings when protecting an individual's private medical data. One of the crucial issues is its lack of data coverage beyond the currently defined covered entities and de-identified data. The provided solution seeks to address the privacy loopholes created by partial coverage and de-identification while improving the quality of data-driven research: focusing on giving the individual the power to track their data, preventing data from being traced back to the individual, and building a concise learning healthcare system that will all improve overall patient care and equity in the medical industry.

This section considers the application of dynamic data ledgers to the healthcare industry, applying a newer and more centralized method of storing data while improving portability. This ledger, by design, features increased security due to its frequently changing encryption of data. Additionally, it allows for better consumer transparency for what data is collected on them and who has access to it.

While this ledger design is like that of blockchain, there are differences that are better suited to transferring secure data that are implemented in our ledger. This ledger is also applied to expanded entities such as health mobile apps and websites, which provides a higher level of security in addition to improving patient information privacy. This section also explores the decision to unify ledger coverage versus simply keeping it as a regulation standard.

Lastly, we examine consumer and organization rights, including the ability to sell PHI and third-party protection of medical data. Overall, these proposed changes ensure an adaptable and self-regulatory environment of sensitive health information.

## What are Data Ledgers?

A data ledger captures features of a database system, adding functionality that allows the data repository to evolve overtime via updatable ledger tables.[13] Information is stored across a

Commented [CVV1]: https://learn.microsoft.com/en-us/sql/relational-databases/security/ledger/ledger-database-ledger?view=sql-server-ver16

Commented [CVV2R1]: 13

distributed network of nodes, allowing multiple parties to access and contribute to system management. The design behind data ledgers includes the purpose to create a secure and tamper proof data record. Transparency goals are achieved through the creation of immutable auditable records, where any data alterations must be verified by all entities, in a system of decentralized consensus. Data ledgers have a wide range of applications, including in industries such as finance, healthcare, and supply chain management. We believe that ledgers are especially helpful when applied to health data management and compliance fields.

## Prioritizing Secure Data Transfer and Transparency

Commented [CVV3]: Security compared to now; discuss centralization

This solution will promote a trustworthy, immutable, and encrypted digital ledger to protect individuals' sensitive health data. VDA, in addition to its valuable hashing and storage features, will aid healthcare institutions by improving patient outcomes and increasing administrative efficiency. For example, a ledger that tracks personal medical data will allow for the consolidation of relevant patient information into one source. In addition, the ledger proposed will unify all PHI data generated. Data aggregation across networks will be implemented using a distributive ledger technology (DLT) system, a software architecture built for data standardization and validation.

One DLT ledger is preferred instead of multiple ledgers, each used by a particular organization. If each entity were to develop and implement a PHI ledger, transferring and aggregating data across businesses would be a burden. Further, developing and maintaining a ledger can be expensive, particularly if it is challenging to integrate with existing healthcare systems. For example, health IT, a government healthcare management agency, estimates that purchasing and installing a ledger system costs between $15,000 and $70,000 per provider.[14] Such a cost can be prohibitive for smaller healthcare providers.

Commented [CVV4]: Was 16

An additional benefit to centralized DLP systems is the reduction of administrative burden. A centralized system will eliminate the need to transfer shared medical data and minimize the risk of data loss resulting from transmitting information across systems. Further, when health providers have consolidated access to data, treatment will be quicker and more personalized. Finally, implementing a central ledger would allow health institutions to effectively run automated queries on their data and detect data misuse.

While a VDA-inspired ledger can have incredible benefits, some hurdles restrict rapid ledger development. First, a digital ledger is only helpful once proven trustworthy. Ledgers must be immutable to gain trust, meaning there is no exception to what data is logged, unlike limited or current de-identified datasets. The subsequent downfall of a ledger-based system is that the patient may not care to read a lengthy catalog of data. However, regulation can simplify this by implementing features that summarize the data for the individual.

A PHI ledger would allow individuals to track the location and use of their sensitive health data. This ledger is critical as HIPAA permits healthcare providers to use de-identified data throughout various operations. Common healthcare operations include business planning and development, fraud detection and abuse audits, and insurance credentialing. De-identified data is used widely within the healthcare industry, yet no specified restrictions are in place to limit the operational handling of de-identified PHI. Therefore, the risk exists that data will be improperly used. Tracking data via a ledger will give patients peace of mind regarding the use of their data.

## VDA-Inspired Ledger Over Blockchain

Blockchain technology bears no formal definition. The technology is mainly referred to as a distributed ledger system, built to handle data shared by multiple parties. Distributive ledger technology (DLT) is a network structure that aggregates records of transactions from various sources simultaneously. Essentially, blocks of data are linked, or chained, together by encryption methods to facilitate secure data storage and verification[15]. DLT, in theory, is not centralized, which is the main distinction between blockchain and our proposed ledger. However, our ledger is crafted to handle extremely sensitive and personal PHI data. Thus, centralization and verification by a trusted authority is critical until consumer trust in DLT technology is established.

DLT verification is additionally highly costly in terms of time and space complexity. For Bitcoin, a popular applied DLT structure, verification is completed in a process called mining. Miners apply elaborate mathematical algorithms using specialized software and hardware to validate transactions and append new blocks to the chain. When a miner solves an algorithm, they successfully validate the block, and their efforts are rewarded with a Bitcoin. This validate-reward system allows verification to be decentralized and self-funding. However, due to the competitive nature of mining, the difficulty in solving mathematical verification algorithms has been continuously increasing. Greater complexity results in a higher resource drain, affecting hardware and software memory space.

The promised success of Bitcoin has occupied financial technology news since the early 2010s. The applied blockchain has been predicted to replace banks and industries relying on trusted third parties' verification functionality.[16] DLT systems, like Bitcoin, can be extremely impactful in the future; however, two prohibitive hurdles

**Commented [CVV5]:** done

**Commented [CVV6]:** Source: https://www.ibm.com/topics/blockchain

**Commented [CVV7R6]:** Now 15

**Commented [CVV8]:** "Will We Realize Blockchain's Promise of Decentralization?" *Harvard Business Review*, 30 Aug. 2021, https://hbr.org/2019/09/will-we-realize-blockchains-promise-of-decentralization.

**Commented [CVV9R8]:** Now 16

restrict prompt DLT implementation. First, the cost of mining is unaffordable to scale: energy and computation resources must be used in every instance of DLT verification. Secondly, once costs are managed, the decentralized system must gain the public's trust. Bitcoin developers claim their technology framework should be trusted by the public because “protocol and software are published openly, and any developer around the world can review the code or make their own modified version of the Bitcoin software.”[17] However, for Bitcoin’s acclaimed openness to be a convincing argument for trust, the public must either have a thorough understanding of the code or trust the developers who review the Bitcoin chain.

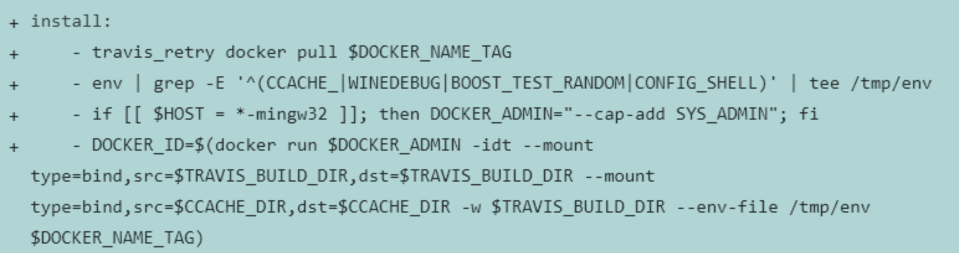
**Commented [CVV10]:** Source: *Can We Trust Blockchain? | American Enterprise Institute - AEI*. https://www.aei.org/technology-and-innovation/can-we-trust-blockchain/.

**Commented [CVV11R10]:** Now 17

Consider a code block from Bitcoin. The block is difficult to interpret, even for those who have a background in computer science:

**Commented [CVV12]:** Add whatever number we're on to GitHub source thing: https://github.com/bitcoin/bitcoin/commit/566f826902cf1a1df18dba83d5302cf173b64e1d

**Commented [CVV13R12]:** 18



```
+ install:
+     - travis_retry docker pull $DOCKER_NAME_TAG
+     - env | grep -E '^(CCACHE_|WINEDEBUG|BOOST_TEST_RANDOM|CONFIG_SHELL)' | tee /tmp/env
+     - if [[ $HOST = *-mingw32 ]]; then DOCKER_ADMIN="--cap-add SYS_ADMIN"; fi
+     - DOCKER_ID=$(docker run $DOCKER_ADMIN -idt --mount
  type=bind,src=$TRAVIS_BUILD_DIR,dst=$TRAVIS_BUILD_DIR --mount
  type=bind,src=$CCACHE_DIR,dst=$CCACHE_DIR -w $TRAVIS_BUILD_DIR --env-file /tmp/env
  $DOCKER_NAME_TAG)
```

Source: GitHub[18]

Bitcoin has gained popularity due to its promise of immutability and decentralized verification. However, decentralization relies on trust, which comes from personal validation of the open-source Bitcoin code or reliance on developers. Thus, there are better approaches than a blockchain-inspired DLT system when building a ledger for personal health data storage. Our proposed ledger relies on centralized verification by the Office of Civil Rights (OCR) to enforce HIPAA compliance.

## Expanding Defined Covered Entities

**Commented [CVV14]:** done

The suggested ledger is proposed to increase data security by implementing technologically adept encryption tools and improving PHI transparency. The covered entities under HIPAA include individual healthcare providers, health plans, and organizations that convert medical information to an online format. Other entities subject to HIPAA regulations include organizations that manage PHI, such as schools, some government agencies, research institutions, and business associates of the traditionally covered entities.[19] However, this only takes care of the conventional medical industry and its extensions, failing to acknowledge the need for equitable care and protection to individuals seeking guidance from websites, apps, or third-party providers. Thus, another possible expansion of the proposed ledger solution involves allowing its appending access to expanded entities.

By expanding such a measure, we allow more cohesive data to be stored and allow for data to be reconfigured at a faster rate, increasing data robustness. Further, this would provide a much-needed standardized model for organizations not currently under HIPAA regulations to store data safely. This approach could lead to a more comprehensive and inclusive healthcare data ecosystem, benefiting individuals seeking care and the healthcare industry.

Using data ledgers in healthcare data is not a novel method as it provides a secure way to store data, maintain accurate medical data through immutable records, and, most importantly, allow for more seamless data sharing. Dynamic data encryption can also expand this ability to safely transfer data between providers to include researchers and other organizations under one overarching bubble. This would boost the collection of better-suited datasets, providing a deeper understanding of population health trends and allowing for the development of more effective treatments. In addition, expanding covered entities and using dynamic ledgers with a mix of de-identifiable data in healthcare data could benefit individuals seeking equitable and comprehensive help.[20] Consolidating health data into one encrypted set could enhance security by reducing the likelihood of data breaches and unauthorized access. Nevertheless, this approach may incur significant expenses as it requires extensive technology and infrastructure investments. Moreover, merging all health information in a single data set might pose privacy concerns due to the increased risk of re-identification and other types of misuse.

Commented [CVV15]: Was 21

Commented [CVV16]: 22

## Considering a National Standard vs Unified Ledger

Commented [CVV17]: done

Data ledgers are a potentially costly method. Instead of establishing a national baseline standard for how companies should or should not protect their data or automatically encompass health data under one bubble, regulation with stricter penalties for companies that misuse or sell their data to other providers or brokers can be enforced. One of these solutions is for firms to state and routinely update how they protect their data, including methods, research, and future improvement plans. If the process is not deemed safe, a fine will be given based on the total data they hold to discourage risky protections. The potential benefits of implementing such regulations could be the prevention of data breaches, increased transparency and accountability from companies, and a greater sense of trust between individuals and healthcare organizations. Along with a lower government cost, it will allow each firm to explore how to best secure their data on a case-by-case basis, offering more customized and adaptable solutions to their specific needs.[21]

The issue with this is the cost to enforce proper compliance since each medical or fitness app, website, and health organization must be monitored to validate their protection measures. Data security audits must be done often to remain competitive against the tech industry, potentially requiring additional funding or forcing third-party audits, which may not be as accurate due to aligning interests. Overall, this will be expensive for businesses to implement and may discourage firms from working together to identify the best way to secure data, using up more resources in the long term. It may be the case that natural competition may drive more secure and consumer-focused firms to the top, creating a more innovative and developed environment, but that is no way to guarantee a higher and safer standard for all health data organizations.[22]

Implementing a unified data ledger would have the same benefits as the data ledger standard and some other key improvements. First, organizations would be able to specialize in their products or services rather than financing proper HIPAA compliance measures. For example, hospitals or clinics can better redirect their resources to more and improved health technology, training programs, and better patient care than investing funds into less-comparable data logging, patient data transparency portals, and storage systems.

Second, security for individuals' data storage will improve with more covered entities. Because of the nature of the data ledger, each time new data is appended, the individual's "key" has changed, re-encrypting the personal ledger and hindering the success of hacking attacks on data. For most, primary care doctor visits are only an annual occurrence, and specialized or emergency care is even less frequent. Due to this, instating a regulation for each entity would limit the potential security of the ledger. Adding entities like wearable health devices such as Fitbit would add more noise to the data, encrypting the data frequently and thus improving information safety.

This would also significantly improve convenience for consumers and facilitate better transparency as data for each individual would be stored under a single ledger. When accessing the data, consumers only need to check one data source instead of logging in to multiple accounts for each provider. They can filter through data and see who has access all in one space, making it more difficult for organizations to hide behind this regulation.

Lastly, HIPAA was created with the primary goal of developing safe methods of health data

**Commented [CVV18]:** 23

**Commented [CVV19]:** 24

portability. A unified dynamic data ledger simplifies the data-sharing process as all information is already centralized. Instead of transferring the data and risking a breach or attempt at hacking, the data is already in one place. Rather, with consumer consent, an organization would receive a modified patient key, allowing them to un-encrypt specific data points from another organization. For example, instead of exporting or transferring weight loss progress to a primary care doctor, that access can be easily shared directly through the ledger, bypassing external safety leaks and risks.

## Consumer Ownership of PHI

Patients may seek more autonomy, such as owning and selling their health data. The costs and benefits of expansion must be considered to decide whether to expand ledger functionality and allow data ownership for sale. Selling PHI is primarily motivated by the revenue generated. To understand the commercial value of PHI, consider the incentive of criminals who breach medical networks and compromise institutional health records. After the acquisition, hackers turn to the dark web to sell stolen medical data, where they can sell for $1 to $1000, depending on the record's completion and the number of records included in the sale.[23] A breach of an entire network can result in a substantial profit. While criminals are motivated to sell multiple records for a profit, individuals would be less motivated to sell their single record– where a dollar is a generous estimate of the individual's gain. While the sale of PHI may have little personal value, the social benefit may differ. In its anonymous form, PHI can provide extensive functionality to health technology applications and researchers studying diseases and making diagnoses. However, there are ways to improve data-driven research without the privacy consequences of PHI sale.

The transaction costs of selling PHI are minimal, and the popularity of e-commerce applications make transactions easier. Selling personal health records on websites like Craigslist and eBay is entirely legal and the cost to sale is a small fee.[24] It is common for e-commerce sites to charge sellers fees that range from 8% to 45% of profit, with an average charge of 15%.[25]

The main costs associated with PHI sales are related to its consequences. Price discrimination is a concerning byproduct. Insurers, employers, and health providers could utilize PHI, and charge customers higher rates if their medical history defines them as less healthy and, therefore, more costly to treat. On the other hand, allowing PHI sales could improve insurance efficiency. If PHI is accessible in an unencrypted form, insurers can offer personalized plans. Insurance buyers will then pay the price that reflects their level of risk. This can incentivize healthier individuals who would otherwise remain uninsured to purchase coverage. As healthier individuals are included in

Commented [CVV21]: 18

Commented [CVV22]: 19

Commented [CVV20]: 17

an insurer's plan, the firm's risk portfolio is diversified, and all parties– the healthy, unhealthy, and the insurance agency, will benefit.

There additionally exists the risk of inaccurate PHI being sold. Inaccurate data includes data entry errors, missing information, outdated information, or erroneous data deliberately appended by cyber criminals. Errors may occur when medical personnel record incorrect information in a patient's electronic health record (EHR). For example, a patient's physical attributes may be recorded incorrectly. Outdated data involves patient information that is no longer relevant. Examples include failure to update patient information after an annual physical or negligence to update demographic information, such as a change in a patient's home address or phone number.

When considering the entry of inaccurate data, the most severe repercussions arise from the malicious addition of erroneous data. Any digital structure is susceptible to cyberattack. Cyber criminals could gain unauthorized access to the ledger, and then add, modify, or distort data. The incentive for criminals to access the health ledger includes the financial gain of mass selling PHI or the monetary benefit of ransomware attacks against health institutions. Additional motivations include espionage and terrorism. To contextualize the repercussions of erroneous health data, consider the 2016 political election: fake electronic health records for the Democratic candidate Hillary Clinton were generated and made public. These fictitious records caused the voting public to be hesitant about Clinton's health and position as a presidential candidate.[26] In summary, any successful health ledger should contain data integrity checks and a recommendation to update data regularly. The ledger should additionally be built with strong access controls, such as multi-factor authentication, to prevent access by malicious actors. In consideration of the option of allowing individual PHI sale, a conclusion from the analysis of costs and benefits suggests that the costs far outweigh the social benefits. Thus, the policy solution does not seek to give individuals the accessibility to sell personal health data.

Commented [CVV23]: 20

## Online Privacy and Third-Party Protection

Another aspect revolves around eliminating non-necessary cookies– text files that websites store on a device to remember user information. Like most online platforms, medical or fitness websites use cookies to personalize an individual's experience and collect data for analytics or advertising.[27] For this solution, Google or other search engines can include a checkbox alongside the search bar for the user to signify that their search is medically related and will signal to the following sites that trackers cannot be used. Some cookies are helpful; first-party cookies set by the website can help users remember login information or website history. However, even

Commented [CVV24]: 25

though it's easier for the user, is allowing that worth it? Many well-known bank websites and apps purposefully don't save users' login credentials and implement two-factor authentication to protect sensitive information from leaking.[28] Health-related digital platforms should be held to the same higher standard. Although it would be slightly more costly for the user regarding login time, it would be a small cost compared to the increased risk of sensitive data getting into the wrong hands.

While HIPAA may already regulate cookies on medical sites, not all health-related ones are. Most websites that are unregulated by HIPAA use trackers like Google Analytics, Facebook Pixel, third-party targeted advertisements, and social media widgets.[29] By removing or limiting these from medical and health-related websites and apps, individuals can feel more secure when browsing online for nearby specialized doctors, prescriptions, or concerning symptoms without worrying about their personal information being collected, tracked, or sold. They can wear a fitness watch or use a period tracker without fear of their data being used against them, ultimately leading to a more trustworthy and ethical digital health industry.[30] By doing this, we allow users to signal to websites that they do not want to be tracked by "opting in" to a more anonymous search. It also sets a limit on what these websites can collect by restricting the use of third-party cookies. The costs involve implementing and maintaining a blocker and may be expensive for search engines, especially small ones, to develop and regularly update the tool. It will also require an additional cost on the compliance side to monitor. Although compared to the costs of privacy violations and legal fees, this change is worthwhile.

There has been a growing push for regulation regarding the use of cookies. For example, the new European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) require companies to obtain user consent before collecting and using their data. Companies like Google have responded to these policies with little pushback, adapting their products and services to fit the new regulations. However, even though tech companies like Facebook and Google were able to make changes to adjust for the CCPA, more needs to be done to continue to protect consumers' privacy.[31] However, there continue to be positive changes. For example, with the introduction of GDPR, some US tech companies are incorporating privacy practices in regions outside of Europe, such as data portability and increased transparency due to consumer demand and the desire to gain a competitive advantage.[32] These changes must continue to promote a safer and more secure online environment.

**Commented [CVV25]:** 26

**Commented [CVV26]:** 27

**Commented [CVV28]:** 29

**Commented [CVV27]:** 28

**Commented [CVV29]:** 30

## Consumer Rights and Ledger Features

The proposed data ledger aims to provide consumers with rights to their data, partially inspired by the GDPR. The features consumers will have on their data consist of Access, Correction, and Deletion—The right to view their data and who has access to it, the right to correct any misinformation, and the right to restrict data sharing or revoke consent to future sharing.

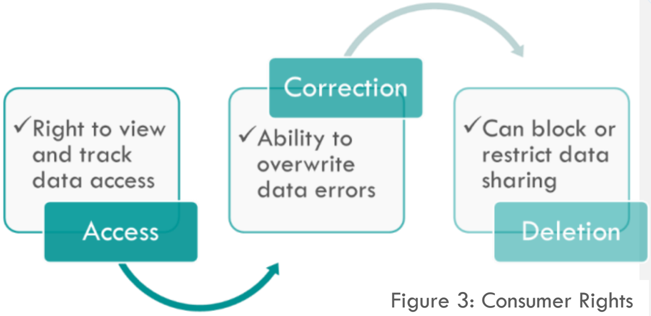

Correction
✓Right to view and track data access
✓Ability to overwrite data errors
✓Can block or restrict data sharing
Access
Deletion

Figure 3: Consumer Rights

Access represents the individual's ability to log in and view what data was recently appended to their ledger. Consumers can also track what organizations have access to their data. If there are organizations that they have not allowed access to, they can report the organization for investigation and block access to their data for that entity. However, there is certain data that patients do not have access to. This data includes information that helps the day-to-day functions of the organization, such as aggregate survey data, or medical records not disclosed to the patient, such as a psychologist's notes on the patient.

Correction refers to the individual's right to append corrected data or overwrite missing information. This is particularly helpful for demographic mistakes. Data directly appended by a current provider may only be corrected by that provider. For example, test results for blood pressure have been appended to the ledger. The patient does not have the right to override these results directly, but if they suspect an error, they can contact the provider for the correct data to be re-appended.[33]

Lastly, while deletion in the traditional sense is not possible with a dynamic data ledger as it is append-only, consumers will have the right to block or restrict access to their data. If a consumer is no longer using a provider, app, or service, they can choose to revoke their consent  for future data access. This means that an individual will have the ability to "delete" an organization's rights to access any data on them in the future. Additionally, the consumer can also completely restrict or block access of the organization on their data. This is done by re-administering a restricted access key on the patient to the organization.

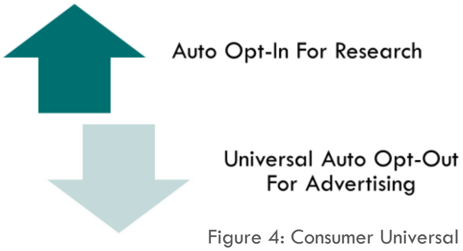


Figure 4: Consumer Universal Opt-in/Opt-out

Commented [CVV30]: https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/

Commented [CVV31R30]: 33

## Firm Access and Ledger Cost

The firms' access to data within the proposed ledger consists of modified rights to collect, share, and sell information.

Organizations have the ability to access any data they have collected on the consumer without additional consent from the user. Likewise, this information is universally opt-in for research purposes. However, if an organization wants access to other information, they must ask the individual for consent to have that data shared with them.

While there is no flat fee for an organization to use this data ledger, access will be granted by a subscription model comparable to the price of current HIPAA compliance. For hospitals, clinics, and researchers, access will be given per verified individual to not restrict data use. For example, a hospital or research institution may pay for a particular number of doctors, nurses, or researchers to have access to patient data on the ledger.

For other organizations, the subscription model will feature a cost per access. In other words, each time a data point was accessed to conduct in-house research or gathered for the purpose of the company, a small fee will be charged. There is no cost to the consumer to see or track their own data.

In terms of sharing, an organization may have the right to consolidate information with another party. However, this must be done by notifying the consumer or with their consent. To sell collected, the organization must have received the opt-in verification for targeted advertising to proceed. Otherwise, data sales are not permitted for collected data.

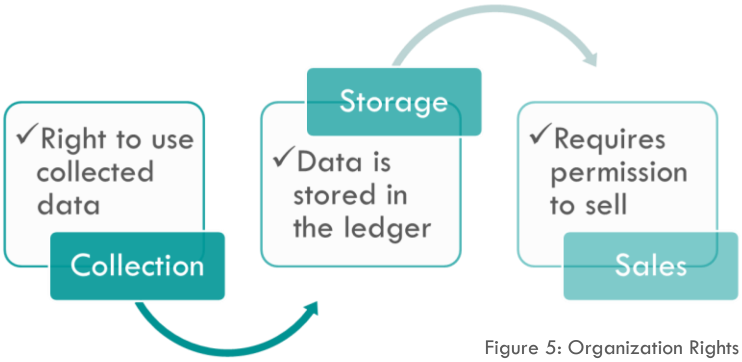

✓Right to use collected data
Collection
Storage
✓Data is stored in the ledger
✓Requires permission to sell
Sales

Figure 5: Organization Rights

Source: Rebecca Stoia

# 3 | PATH FORWARD

## Current Standards and Overview

This proposal suggests a digital ledger to store PHI. PHI appended to this ledger is data collected from any source handling PHI, whether it be hospitals, insurance companies, or mobile health applications.

Before discussing how a distributed PHI ledger can be implemented, let us consider the current healthcare data storage system. In 2008, 9% of hospitals collected their health records in digital databases, compared to 2019, where 96% of hospitals were digitized.[34] The shift to electronic record-keeping provided healthcare institutions with unique benefits, including standardization, reduced operational errors, improved patient satisfaction, and fraud detection capabilities. These improvements provide a modern foundation for next-generation digitization technologies like the enclosed ledger technology. The proposed ledger will solve a major downfall of the database system: disconnection. Currently, each hospital stores health records individually and no central unified data repository exists.

As a response to the lack of unification in traditional database systems, the cloud emerged. Cloud technology's computing capabilities allow for centralizing large volumes of electronic records. Records are in-sourced from various hospitals and made accessible to connected parties. Cloud utilizes a network of remote servers to manage and process data, eliminating the need and challenge that bespoke on-site solutions provide. The result is increased data accessibility and organizational efficiency. Given recent advancements in centralization and access features, a staggering 98% of healthcare organizations have adopted or made explicit plans to adopt cloud technology by 2023.[35]

Cloud technology partially targets the decentralization issue present in database systems. However, centralization benefits are limited within the network of healthcare organizations that implement cloud technology due to organizational silos. PHI aggregated by other entities, such as insurance providers, do not gain access to the shared ledger. Further, the data produced by third-party applications is not mentioned in healthcare cloud solutions. The ledger promoted in this proposal resembles a shared cloud SaaS (Software as a Service) offering by promoting standardized data audit, centralization, and structured remote access. However, the dynamic PHI ledger critically focuses on PHI rather than the general EHR data. PHI is the most sensitive of all health data

**Commented [CVV33]:** 32

**Commented [CVV32]:** 31

collected and merits unique protection. The proposed ledger is crafted specifically to meet PHI protection requirements. This includes a focus on PHI aggregated from all sources: hospitals, the broader covered entities, and third-party applications.

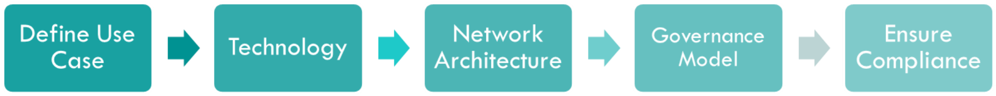
Healthcare organizations have rapidly digitalized, and technology has become omnipresent within hospital operations. A PHI ledger will thus integrate conveniently into existing business models. Additionally, the ledger has features analogous to both cloud and database systems, which will further increase convenience in implementation.

## Legislative 5-Step Plan

Executing the data ledger will require planning, coordination, and collaboration among the traditionally covered entities and third-party organizations. Implementing this technology involves taking five components into account: defining the use case, selecting the technology, considering the network architecture, developing a governance model, and ensuring compliance with regulators.

This 5-Step process will ensure dynamic enforcement that can quickly adapt to necessary changes in the healthcare industry. To encourage development of the ledger system, necessary audits and storage maintenance are reduced, leaving resources available for the continuous development and success of this model.

Figure 6: 5-Step Legislative Plan



## Improving Communication and Expectations

First, the proposed ledger is purposed to offer comprehensive reidentification protection to all PHI generated. This includes data from traditionally defined HIPAA-covered entities and the new third-party applications. Comprehensive anonymity will be established by implementing technology features that increase data transparency.

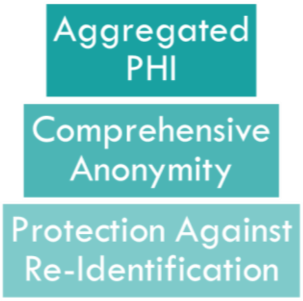


Figure 7: Use Case Goals

## Implementing Technology

The technology selected is the distributed ledger, built upon the VDA framework. Within the ledger, a range of authorized entities can aggregate their de-identified data. Patients will have streamlined access to track their individual PHI, and misuse detection will be implemented based on the synchronization and verification functionality of Merkle trees.

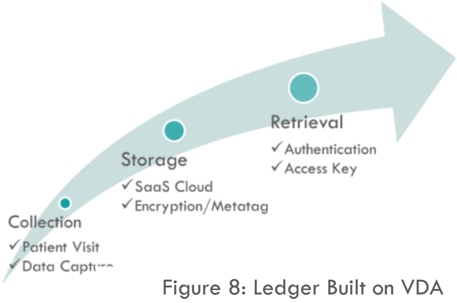


Figure 8: Ledger Built on VDA Framework

## Network Architecture Design

Network architecture involves making decisions around the required level of resilience, implementing protocols for achieving consensus, and maintaining data integrity. We recommend setting a standard level of resilience and establishing consensus models via governance.

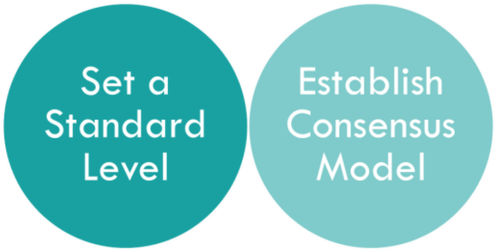


Figure 9: Functioning Network

## Establishing a Governance Model

Any governance model should ensure all involved parties are equally represented and able to contribute to system management. We recommend verification be issued by The Office for Civil Rights (OCR), the agency that currently is responsible for enforcing HIPAA. Further, a steering committee should be established, where interested parties can send a member to advocate for desired ledger rules.

Figure 10: Ensure Parties Are Represented

## Ensuring Future Compliance

Finally, the proposed ledger is fully compliant with HIPAA regulations. Data de-identification standards are not only met but held to a higher standard than before protecting the confidentiality, integrity, and availability of patient health data. Additionally, strong verification processes are in place to detect HIPAA violations, reducing the risk of breached health data incidents.

Lastly, after the five implementation criteria are established, the proposed ledger will be piloted with a small group of healthcare providers. Defects will be identified and revised before the technology is scaled. Successful ledger implementation will offer measurable improvements in security, transparency, and efficiency.

The success of a PHI ledger depends on the technology's ability to provide secure, private, interoperable, and user-friendly storage of sensitive health data. The main factors considered when evaluating PHI ledger success include speed of adoption, interoperability, and consumer sentiment.

The first two criteria of the evaluation plan are inherently intertwined: the more interoperable the ledger, the faster the expected speed of adoption. Interoperability involves the nature of system implementation within current organizational platforms. A successful ledger should be implemented seamlessly within current organizational platforms. Considering that most hospitals are collecting data digitally and storing electronic records in the cloud, our proposed ledger is highly interoperable. High interoperability will increase the expected speed of adoption; however, adoption additionally depends on the realization of tangible benefits and patient usability.

The final component of success evaluation involves a consumer sentiment analysis. This analysis will analyze attitudes, options, and emotions expressed by patients, healthcare providers, and organizations, towards the digital ledger. The objectives of such an analysis are to identify the strengths and weaknesses of the PHI ledger, while simultaneously considering consumer preferences. After outlining objectives, the next step involves data collection. Data can be aggregated from social media, online reviews, surveys, and hospital reports. Then, data will be examined using natural language processing and sentiment analysis to classify consumer perspective of the ledger. A positive sentiment and strong insight trends signal a successful public response.

# CONCLUSION

## There are evident inefficiencies in how medical data is stored and used. It is time to close these data protection loopholes.

Businesses have proven to have suboptimal privacy policies in mind when creating and running their organization. From unclear and lengthy terms and conditions to predatory profit-centered behavior, most take advantage of HIPAA's loopholes rather than take an honest consumer-forward approach.

**Commented [CVV34]:** Let's also revamp this

**Commented [CVV35R34]:** Add patient application; benefits of using this ledger, etc. how would this appy to a day-to-day routine/how would it change the lives of patients

On the other hand, hospitals face inefficiencies with data collection and storage. Older and underfunded hospitals and clinics still heavily rely on non-digital data collection, which reduces efficiency and makes unifying data challenging to implement. Other hospitals, however, have begun installing database systems or cloud technology. These technologies have improved operations, yet centralization is still lacking due to the high quantity of PHI processed. The ledger solution proposed provides a way to unify data aggregated across networks, while maintaining data integrity.

This is why amending the regulation is essential. The proposed data ledger will make data storage and accessibility easier and cheaper for institutions. In addition, this will establish a unified database with standardized expectations for the data collected and stored, providing much-needed accessibility and transparency to individuals.

De-identification is not sufficient as current standards cannot keep up in the race against re-identification algorithms. Our proposed ledger gives the patient the ability to track and access their data using personal authentication keys.

Including a broader scope of covered entities will also help expand data privacy to all consumers, regardless of their preferred method of getting medical resources. These privacy protections should not be restricted to those who can afford to pay for or take a day off to consult with a traditional in-house doctor about their concerns.

This is an issue with the equal right to privacy just as much as it is the safety and security of private medical data. Current regulations must be diligent in addressing all these concerns.

# ENDNOTES

[1] Khanna, Ajay. “4 Health Care Data Challenges and How to Overcome Them.” Corporate Compliance Insights, 28 June 2018, https://www.corporatecomplianceinsights.com/4-health-care-data-challenges-overcome/.

[2] “Guess What? HIPAA Isn't a Medical Privacy Law.” MSN, 10 August 2022, https://www.msn.com/enus/news/technology/guess-what-hipaa-isn-e2-80-99t-a-medical-privacy-law/ar-AAYpK9i.

[3] “HIPAA BUSINESS ASSOCIATE COMPLIANCE AGREEMENT CLAUSE.” DC Attorney General, 2019, https://oag.dc.gov/sites/default/files/2019-09/Attachment-J9-RFP-DCCB-2019-R-0013.pdf.

[4] Lomas, Natasha, and Romain Dillet. “Terms And Conditions Are The Biggest Lie Of Our Industry.” TechCrunch, 21 August 2015, https://techcrunch.com/2015/08/21/agree-to-disagree/?guccounter=1.

[5] Hales, Maggie. “Mobile Health Apps and HIPAA.” The HIPAA E-Tool, 29 June 2021, https://thehipaaetool.com/mobile-health-apps-and-hipaa/.

[6] National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/books/NBK9579/.

[7] Talby, David. “Medical Data De-Identification Is Under Attack.” Forbes, Forbes Magazine, 27 August 2019, https://www.forbes.com/sites/forbestechcouncil/2019/08/27/medical-data-de-identification-is-under-attack/.

[8] “Methods for De-identification of PHI.” HHS.gov, Office for Civil Rights (OCR), 26 October 2022, https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#safeharborguidance.

[9] “Methods for De-identification of PHI.” HHS.gov, Office for Civil Rights (OCR), 26 October 2022, https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#safeharborguidance.

[10] “Google, University of Chicago named in suit charging misuse of patient data.” Healthcare IT News, 1 July 2019, https://www.healthcareitnews.com/news/google-university-chicago-named-suit-charging-misuse-patient-data.

[11] Miller, Katharine. “De-Identifying Medical Patient Data Doesn't Protect Our Privacy.” Stanford HAI, 19 July 2021, https://hai.stanford.edu/news/de-identifying-medical-patient-data-doesnt-protect-our-privacy.

[12] Liangyuan Na, BA. “Feasibility of Reidentifying Individuals by Their Protected Health Information.” JAMA Network Open, JAMA Network, Dec. 2018, https://es.jamanetwork.com/journals/jamanetworkopen/fullarticle/2719130.

[13] “Database ledger - SQL Server.” *Microsoft Learn*, 3 March 2023, https://learn.microsoft.com/en-us/sql/relational-databases/security/ledger/ledger-database-ledger?view=sql-server-ver16.

[14] “How Much Is This Going to Cost Me?” How Much Is This Going to Cost Me? | HealthIT.gov, 12 Nov. 2014, https://www.healthit.gov/faq/how-much-going-cost-me#:~:text=Several%20studies%20estimate%20the%20cost%20of%20purchasing%20and,select%20on-site%20EHR%20deployment%20or%20web-based%20EHR%20deployment

15 “What is Blockchain Technology? - IBM Blockchain.” *IBM*, https://www.ibm.com/topics/blockchain.

16 “Will We Realize Blockchain's Promise of Decentralization?” *Harvard Business Review*, 30 Aug. 2021, https://hbr.org/2019/09/will-we-realize-blockchains-promise-of-decentralization.

*17 Jamison, Mark, and Christine Wilson's. “Can We Trust Blockchain?” American Enterprise Institute, 9 August 2018, https://www.aei.org/technology-and-innovation/can-we-trust-blockchain/.*

*18 GitHub*, 2 August 2018, https://github.com/bitcoin/bitcoin/commit/566f826902cf1a1df18dba83d5302cf173b64e1d.

19 Alder, Steve. “What Does HIPAA Cover? Updated for 2023.” HIPAA Journal, 1 January 2023, https://www.hipaajournal.com/what-does-hipaa-cover/.

20 Snell, Elizabeth. “Benefits, Challenges of Secure Healthcare Data Sharing.” HealthITSecurity, 20 October 2017, https://healthitsecurity.com/features/benefits-challenges-of-secure-healthcare-data-sharing.

21 Rahnama, Hossein. “The New Rules of Data Privacy.” Harvard Business Review, 25 February 2022, https://hbr.org/2022/02/the-new-rules-of-data-privacy.

22 Boushey, Heather, and Helen Knudsen. “The Importance of Competition for the American Economy | CEA.” The White House, 9 July 2021, https://www.whitehouse.gov/cea/written-materials/2021/07/09/the-importance-of-competition-for-the-american-economy/.

23 Stack, Brian. “Here's How Much Your Personal Information Is Selling for on the Dark Web.” Experian, Experian, 11 Mar. 2019, https://www.experian.com/blogs/ask-experian/heres-how-much-your-personal-information-is-selling-for-on-the-dark-web/.

24 “You Can Now Make Money Selling Your Own Health Data, but Should You?” Fast Company, 27 Sept. 2019, https://www.fastcompany.com/90409942/would-you-sell-your-own-health-data-theres-a-market-for-it-but-ethical-concerns-remain.

25 Brophy, Meaghan. “Amazon Seller Fees: Cost of Selling on Amazon in 2023.” Fit Small Business, Fit Small Business, 13 Feb. 2023, https://fitsmallbusiness.com/amazon-seller-fees/.

26 Yao, Mariya. “Your Electronic Medical Records Could Be Worth $1000 to Hackers.” Forbes, Forbes Magazine, 18 Apr. 2017, https://www.forbes.com/sites/mariyayao/2017/04/14/your-electronic-medical-records-can-be-worth-1000-to-hackers/

27 “Cookie Text Explained: What Is It and How To Use It?” CookieYes, 19 August 2022, https://www.cookieyes.com/blog/cookie-text/.

28 Sabau, Codina. “5 Ways in Which Banks Secure Their Data.” Endpoint Protector, 8 June 2022, https://www.endpointprotector.com/blog/ways-banks-secure-data/.

29 Lubowicka, Karolina, and Małgorzata Poddębniak. “Is Google Analytics HIPAA-compliant?” Piwik PRO, 20 February 2023, https://piwik.pro/blog/is-google-analytics-hipaa-compliant/.

[30] Germain, Thomas. “Guess What? HIPAA Isn’t a Medical Privacy Law.” Consumer Reports, 13 June 2022, https://www.consumerreports.org/health-privacy/guess-what-hipaa-isnt-a-medical-privacy-law-a2469399940/.

[31] Singer, Natasha. “What Does California's New Data Privacy Law Mean? Nobody Agrees (Published 2019).” The New York Times, 29 December 2019, https://www.nytimes.com/2019/12/29/technology/california-privacy-law.html.

[32] Edinger, Julia. “Tech Companies Embrace Some GDPR Privacy Practices Outside of Europe.” Government Technology, https://www.govtech.com/policy/tech-companies-embrace-some-gdpr-privacy-practices-outside-of-europe.html.

[33] Bellamy, Fredric D. “U.S. data privacy laws to enter new era in 2023.” *Reuters*, 12 January 2023, https://www.reuters.com/legal/legalindustry/us-data-privacy-laws-enter-new-era-2023-2023-01-12/.

[34] NetspectiveMedia. “The Pros and Cons of Healthcare Database Systems.” *The Healthcare Guys*, 26 Aug. 2021, https://www.healthcareguys.com/2019/10/27/the-pros-and-cons-of-healthcare-database-systems/#:~:text=In%20the%20U.S.%20there%20are%20many%20types%20of,sector%20is%20the%20OLTP%20%28Online%20Transaction%20Processing%29%20database.

[35] Glib, Oleksii. “Cloud Computing in Healthcare Explained.” *Bespoke Software Development Company Acropolium*, Acropoliumhttps://Acropolium.com/Img/Base/Logo.svg, 7 Apr. 2023, https://acropolium.com/blog/cloud-computinghealthcare/#:~:text=With%20cloud%20computing%2C%20healthcare%20organizations%20can%20use%20remote,cloud-based%20architecture%20cost%20less%20than%20with%20on-premise%20servers.