[[1]](#footnote-1)

Practical Framework for a Universal Health Insurance Database

Meredith Ludlow and Javier Saldaña Jr.

*Abstract*—A database is only as good as the data it collects. The health insurance industry like many other industries, is fragmented and only able to perform data analytics to the depth their database allows. There are legal and practical constraints which complicate the health insurance industry to strive for a collective database like the one privy to the property and casualty insurance industry. A framework for a collective relational database to be shared by health carriers is introduced. Along with the framework and design of the collective database, a roadmap for overcoming the constraints currently in place are also presented.

*Index Terms*—business, claim, database, insurance, universal.

# INTRODUCTION

O

NCE considered a competitive advantage, data analytics are now mandatory for a business to be able to survive in today’s market. The insurance industry, which has been hoarding data on its consumer base for decades, is starting to implement the use of data analytics into other facets of the business outside of their underwriting practices. However, the databases used by the carriers are only as powerful as the data they contain within themselves. This creates an information asymmetry which gives the large carriers a competitive advantage over smaller carriers, thus allowing a borderline oligopoly to remain. The problem with an oligopoly in the health insurance industry is the same with all oligopolies, the market is manipulated by few and controlled to serve the shareholders of those companies in order to enhance their wealth.

This problem of information asymmetry can be solved through the development of a collective relational database that would be

was once considered a luxury and competitive advantage. Today, the use of data analytics is seen as a requirement across various industries. Although the insurance industry has only recently began to leverage their late to the game, the insurance industry has been slow to embrace the big data strategies many other industries (such as retail and marketing) have grown to become synonymous with. However, the slow adoption of big data strategies doesn’t mean the data isn’t being collected. In 2011, the “U.S. health care data alone reached 150 exabytes,” which is the equivalent of 161,061,273,600 gigabytes (iHT2, 2013). The data collected by health carriers may range from a wide spectrum of facets pertaining to an individuals’ life.

Consider a scenario in which a customer has an offspring added to his/her health insurance policy. Assuming the child is insured through the same carrier until he/she is old enough to be on their own policy, every medical record and/or visit of that child will be present in the health carrier’s database. In every visit, the physician must obtain health performance data in order to determine which medicine may be administered, etc. Should the same child, now a grown adult, choose to remain with the same insurance carrier, the data collected on that individual will continue until the policy and/or the customer expire. Considering the sensitivity of the health information collected, health carriers (along with anyone else who collects and stores health information in the United States) must comply with federal regulations in order to ensure the entity is practicing great care and responsibility of the data.

The three pieces of legislation that govern the health carriers are The Gramm-Leach-Bliley Act (GLBA), the Health Insurance Profitability and Accountability Act (HIPPA), and the Fair Credit Reporting Act (Nunn, 2007).

As previously mentioned, collecting and storing data is something the health carriers are already efficient in and have been doing for years. Whether it is through traditional means such as mail correspondence or digital delivery, insurance laws across the United States have bestowed the burden of proof on the claimants when seeking coverage under a policy. In other words, if a claimant files a claim for a hospital visit, it is ultimately up to them to provide the health carrier the necessary medical records in order to meet the burden of proof.

In order to comply with the federal and local regulations, carriers have developed standard processes in which employees are expected to act accordingly in order to preserve the integrity of the data but also protect the data from external threats. However, this has led to a fragmentation of databases as each carrier strives to build their own universal health care database. The fragmented databases have spurred new challenges and hurdles that carriers are now facing today.

As a result, third-party companies have taken the initiative to attempt to create collaboration between carriers and their data, which may be used to report and query prior claim information for a claimant regardless of who the carrier was at the time. However, the subscription fees tend to be very expensive and many small insurance companies are not members of such services, which creates inaccurate and/or incomplete records. Other services provide access to the verification of assets owned and prior insurance information but have the same pitfalls as the prior service. Since most of these services are membership driven, they are home to the major insurance carriers but neglect the value of the small insurance carriers. The information inaccuracies result in inefficiencies within major corporations and increase in claim cost, which results in a higher premium penalty for the customers.

In this paper, we will propose a solution to the fragmented databases by presenting the framework for a universal database with an efficient input/query system. A solution to this problem would be the development of a universal insurance database. A universal insurance database would host claim, sales, underwriting, and much more data for the insurance carriers to be used for cost-saving strategies in an effort to reduce premiums for the customers. While we believe a universal insurance database would result in quicker claim resolutions and a substantial decrease in expenses, we also believe a single database housing the information would be much more efficient and practical for all parties involved. Yet, we are also aware of the challenges and costs associated with this proposal and will also address them here along with creative solutions we believe may help overcome those challenges.

# Centralize Database Examples

## National Crime Information Center (NCIC)

One of the most successful examples of a centralized database is the National Crime Information Center, which was developed by the Federal Bureau of Investigation and launched in January 27, 1967. Originally just a small network between 15 state and city computers, the network only had 350,000 records over 5 different categories (FBI, 2017). The centralized database was improved over the years in order to expand to additional categories and new technological additions. By 1975, details for missing persons cases were added, which were then followed by the addition of Canadian arrest warrants in 1980 (FBI, 2017). By 1999, images and additional categories were added along with structural improvements which allowed for near instant query results (FBI, 2017). Today, the NCIC database contains 21 different categories and 12 million active records, while handling 14 million transactions per day and serving 90,000 criminal justice and law enforcement agencies (FBI, 2017).

The NCIC has been an instrumental tool in criminal investigations all over the country. However, it was once in a similar situation as was the current health care system today. Prior to the use of computers, law enforcement agencies kept physical criminal records and the most common way to query an individual’s criminal background was by asking neighboring law enforcement. This meant that an individual could easily move locales and/or states should they wish to start a new life without the burden of a criminal background following them. To query every law enforcement agency at the time was unheard of. Yet, the emergence of technology gave law enforcement the capabilities to add value to their record keeping. Director J. Edgar Hoover proposed the NCIC in 1965 as a centralized crime information center for all participating agencies, which would require them contribute to the system to be able to partake (FBI, 2017). What originally started in Washington, DC and surrounding metropolitan areas, soon spread to other regions and by 1971 included all 50 states.

The law enforcement saw the added value of a centralized crime database and has seen it pay dividends since its inception. In December 2016, a Tennessee state trooper queried the NCIC for licensed plates after spotting a suspicious vehicle at a rest stop. The results showed the vehicle was stolen in Kentucky and the Illinois-native who was driving the vehicle was a suspect in a Kentucky murder (FBI, 2017). The National Crime Information Center is a key example of member recognizing the data is worth more as a collective than a fragmented database.

## Insurance Services Office (ISO)

There are voids the public sector is unable to fill and it is there where the private sector is contributing. The Insurance Services Office is a subsidiary of Verisk Analytics, which is headquartered in Jersey City, New Jersey. While Verisk offers a collection of analytical services, it also has one core purposes which is similar to that of the NCIC. The Insurance Services Office servers mostly the property and casualty insurance sector, which is typically auto and property (personal and commercial). An insurance carrier/agent/broker is able to subscribe to ISO through its services, in which ISO then grants the subscriber access to the large collective data (including its analytical services). The subscriber provides its claim data to ISO, who then adds it to the collective database and sales access to it as a research claims tool for others to use. This strategy is beneficial for small companies who don’t have the data depth or infrastructure the major carriers have in order to be able to conduct cutting-edge analysis. For larger companies (and small ones as well), this helps them cut down on claim cost by ensuring the industry as a whole is evaluating claims similarly, which helps reduce the claim cost to its true value.

ISO began as a collective non-profit initiative by insurers as an effort to be able to share data from each other’s book of business and more accurately rate customers’ risk level in 1971. As the needs of the insurers evolved, so did ISO. Eventually, it became a for-profit organization and a subsidiary of Verisk Analytics. The evolution to a for-profit company allowed ISO to be able to provide analytical services to startup insurance carriers while leveraging the data of the major carriers.

However, even ISO has its limitations and shortcomings. Federal regulations prohibit ISO from obtaining the medical records from the carriers regarding claims. Instead, it circumvents this regulation by only collecting the injury type and settlements instead of the medical record itself.

# Proposed Framework

The proposed framework for the universal healthcare database is show in Figure 1 at the end of this paper. This section explains the though processes and assumptions that went in to creating this framework.

# Challenges & Regulations

The main challenges of creating a centralized database for healthcare insurance and claim information are federal regulations and privacy concerns. There are many federal regulations that dictate how health care data is protected and shared. This paper will focus on the Gramm-Leach-Bailey Act and HIPPA.

The Gramm-Leach-Bailey Act requires that financial institutions, including insurance companies, must share their data sharing practices with their customers and obligates an insurance company to protect its customer data. Giving this data to a universal healthcare database could conflict with staying in compliance with this rule. Since insurance companies must share with customers their data sharing practices, customers would also have the option to opt out of sharing with the universal database. This could lead to a lack of data in the database if large amounts of people choose to opt out.

The second major regulation is HIPPA, or the Health Insurance Profitability and Accountability Act. HIPPA has quite a few parts that involve healthcare insurance practices, but the relevant section for this paper is Title II: Administrative Simplification. This section of HIPPA has a privacy rule that requires health insurance carriers to protect patient health information, specifically, information that makes a patient identifiable. A proposed universal healthcare database could potentially violate the HIPPA privacy clause depending on what data was shared with the universal database. If the data given to the database was such that it didn’t violate a customer’s privacy, it might be so sparse that it could lessen the benefits of even having a universal database at all.

Another major challenge that arises in building a universal database is the danger of having all healthcare data in one place. While fragmented databases, each piece with a different carrier, are less beneficial for analytics, the data ends up being safer. If a carrier’s data where to be hacked, the hacker would only have access to a portion of the nation’s healthcare information. If a universal database where hacked, the magnitude of people whose information has the potential to be stolen increases. There is a balanced that needs to be found in the benefits of analytical ability and data privacy.

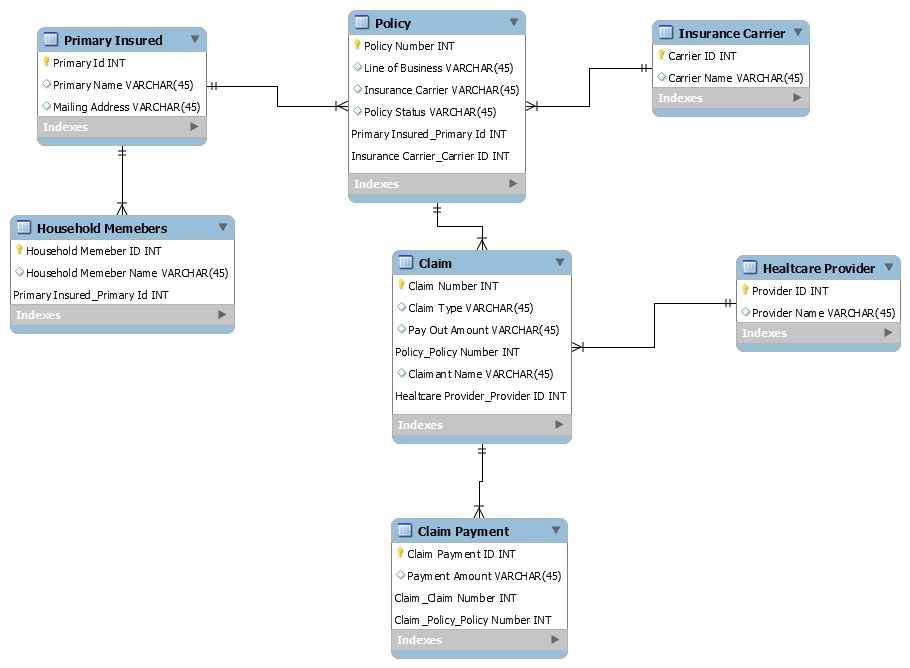
# Conclusion

References

1. iHT2 (2013). *Transforming Health Care Through Big Data: Strategies for leveraging big data in the health care industry*. Retrieved from: <http://c4fd63cb482ce6861463-bc6183f1c18e748a49b87a25911a0555.r93.cf2.rackcdn.com/iHT2_BigData_2013.pdf>
2. Nunn, T. (2007, March 26). *Consumer Data Protection*. Retrieved from Insurance Journal: <https://www.insurancejournal.com/magazines/mag-legalbeat/2007/03/26/78291.htm>
3. Federal Bureau of Investigation (2017, January 27). *NCIC Turns 50: Centralized Database Continues to Prove Its Value in Fighting Crime.* Retrieved from: <https://www.fbi.gov/news/stories/ncic-turns-50>
4. “Gramm-Leach-Bliley Act.” Federal Trade Commission, [www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act](http://www.ftc.gov/tips-advice/business-center/privacy-and-security/gramm-leach-bliley-act)
5. “What Is HIPAA (Health Insurance Portability and Accountability Act) ? - Definition from WhatIs.com.” SearchHealthIT, searchhealthit.techtarget.com/definition/HIPAA

APPENDIX

Figure 1



1. This rough draft was submitted for review on Mar. 17, 2019.

   M. Ludlow is a graduate student at Southern Methodist University in Dallas, TX. She is currently employed by \_\_\_\_ Insurance Agency in \_\_\_, TX as a Licensed Insurance Agent (e-mail: [mludlow@mail.smu.edu](mailto:mludlow@mail.smu.edu)).

   J. Saldaña is a graduate student at Southern Methodist University in Dallas, TX. He is employed by Allstate Insurance in Irving, TX as a Senior Data Consultant (e-mail: [saldanaj@smu.edu](mailto:saldanaj@smu.edu)). [↑](#footnote-ref-1)