



# Data Ownerships of Big Data Usages in Healthcare Industry

By Stephen Li, MBA Candidate, Class of 2017

May 2016



## Big Data in Healthcare

The goal of value-based healthcare is providing high-quality healthcare service while reducing the cost. In a pay-for-service system, payers need to pay for every single admission (including readmission), drug, and medical device. That brings high cost without increasing value to patients. The ACA transforms healthcare from pay-for-service to value-based care models, the transaction dedicated in providing better service with lower cost. Big data, by virtue of huge information, can help us predict disease trends and prevent and treat disease<sup>1</sup>. For example, as early as 2001, several research group developed models to predict foot-and-mouth disease in UK<sup>1</sup>. Also, UnitedHealth Group is using big data to analysis the health status of people, and suggest them to make appointments with doctors when UHG see risks<sup>2,3,4,5</sup>. Big data are also more commonly used in precision medicine and cancer treatment<sup>6</sup>.

## Data Sources and Challenges

When we try to build up a big data application for healthcare, we have to first figure out where the data comes from. Of all healthcare data, 10% happens in the hospital, and 90% happens in our daily life. Considering different part of data, our sources include electronic healthcare records (EHRs), claims, data collected by medical devices, and data in our daily life.

### Incompatible Healthcare Systems

Since all healthcare systems are started independently and the data structure and code are heterogeneous, it is very hard for different systems to communicate with each other.

For example, EHR systems are not compatible with each other. EHR systems started in the 70s and there are currently five large companies (Semmens, General Electric, Epic, Cerner, and Meditech) competing in this area. Incompatibility caused two results: First, once a provider chooses an EHR system, it is extremely hard to change to another. Second, it is hard to integrate these data due to different structures.

Major payers, on the other hand, build in-house systems rather than purchase them, which brings in a bigger variety of data structures. Medical device companies face a similar situation.

### Various Daily Life Data Sources

For daily life data, some are collected from the internet of things (IoT, e.g. Fitbit records), some are shared in social network applications (such as Twitter), and some are stored in Isolated systems (e.g. gym visiting records). The variety of these data, especially without identification information among systems, makes data integration almost impossible.

### Unwillingness to Share

Because of the privacy demands of HIPPA and HITECH, organizations that exchange or share data need either to get patients' approval of data usage or de-identify data before sharing. Although following these acts avoids the privacy issue, there are still some reasons that organizations are not willing to share their data. First, organizations see data as their assets and they don't want other companies to take advantage of their data. Second, providers want to keep patients within their business rather than share their information to other providers. The un-standardized EHR systems somewhat helped them build barriers.



## Big Data Usages and Data Ownerships

Even though there is some reluctance in the marketplace, other models based on different types of data ownership have been introduced.

### Private Sector Led Model - Optum Labs

Optum Labs is a not-for-profit organization under Optum, which is the information technology company under UnitedHealth Group<sup>3,4</sup>. The lab was started by Optum and Mayo Clinic, and is partnered with more than 20 other organizations. Optum Labs provides data mining and analytics service for partners and customers.

Optum Labs creates partnerships in several different ways, one of which is called “strategic data contributors,” who provide data to help partners tackle big problems in healthcare<sup>2</sup>. All data are de-identified prior to being used at Optum Labs. The open-ended, ease-of-access features allow Optum Labs to attract more partners. In its first year, Optum Labs has attracted 15 total partners<sup>5</sup>. Collectively, the partnership can view insurance claims from more than 150 million people and electronic medical records of 40 million people<sup>5</sup>.

By using these data, Optum Labs helped to finish more than 20 research projects, including estimating the population of diabetics in the U.S. By September 2015, Optum Labs was involved in 70 other projects<sup>6</sup>. A doctor said the biggest benefit of the collaboration was that “we called these larger studies constellations— because they bring together a variety of elements to create a bigger picture. You do a study that helps you design the next study.”<sup>7</sup>

### Public Sector Lead Model - MN E-health

MN E-health is a Healthcare Information Exchange (HIE) Center run by the Office of Health Information Technology (OHIT) at the Minnesota Department of Health<sup>8</sup>. The Purpose of E-health is the adoption and effective use of electronic health record systems and other health information technology to improve health care quality, increase patient safety, reduce health care costs, and improve public health<sup>9</sup>.

E-health encourages organizations to share data by giving them financial support. In 2014, E-health granted \$3.8 million to support community data exchanges under the Minnesota State Innovation Model Project, and a nearly \$500 million federal incentive payment was given to providers for meeting the criteria of adoption and meaningful use of EHR. E-health is facing not only a technology problem, but also a collaboration problem –it needs to establish a governance system and gain public trust and engagement. In 2004, the Minnesota E-health Initiative was established to “work as a public-private collaboration to pursue strong policies and practices to accelerate e-health with a focus on achieving interoperability (the ability to share information seamlessly) across the continuum of care.” A Minnesota e-Health Advisory Committee of 25 members from different healthcare organizations helped these organizations build trust and work smoothly together. With the help of the Office of Health Information Technology, it is easier to find synergy among organizations, and provide expertise in practice.

As a result, E-health uniformly collects and routinely shares the results statewide, helps measure Minnesota’s status on achieving state or national goals, and helps organizations adjust their decisions and policies by using research results.



### Sharing model - Clinically Integrated Network

A Clinically Integrated Network (CIN) is a network of health providers and other service providers collaborating to improve and maintain the health of the people in their communities. Information systems are the key for successful CINs, because collaboration requires health information technologies that communicate, aggregate, integrate, and analyze data.

Different CINs tend to choose different data sharing strategies<sup>10</sup>. Most CINs are choosing to develop a single health information system with integrated inpatient and outpatient information, and merge old data into the new system. In this way, all providers need to share their data into one single system, with the data owned by the CIN. An opposite way is a requirement-based information exchange. Some CINs build their network based on requirement—the information exchange only happens when providers require it. If the CIN choose this option, each provider's data are exclusively its own.

As the integration enlarges data size, CINs could help their providers and physicians be more accountable for higher quality and more efficient care delivered at a lower cost<sup>11</sup>.

### New Trends of Healthcare Data Usages

In recent years, some new trends have great potential to contribute to data sharing. The most significant trends that differentiate from traditional ways are crowdsourcing and data donation.

#### Crowdsourcing - Twitter

Twitter provided various APIs for developers to access both public and protected data sets. Public streams API is an API which allows users to obtain high-throughput, near-real-time access to various subsets of public and protected data. Together with sophisticated linguistic algorithms, Twitter's data can be used for customized purposes. For example, at the Pediatric Academic Societies' 2016 meeting, a study showed that using Twitter data could predict and possibly prevent severe asthma attacks in a community<sup>12</sup>. Representative, real-time, and free-of-use features are crucial to epidemic research.

#### Data Donation – Data Donors

In the healthcare industry, patients—or healthy people—started to donate data for research, and some organizations established channels to accept donations. One example is DataDonors<sup>13</sup>, a website owned by the non-profit WikiLife Foundation. DataDonors' mission is to provide the global health community a database that can be used by researchers, scientists, and physicians around the world to discover new research behind healthy living. With donors' agreements, the website is collecting data on every aspect of daily life. Sources include manual input (e.g. education information, work experience), social network (e.g. Facebook, LinkedIn), wearable devices (e.g. fitbit, nike+), special service providers (e.g. 23 and me), and more. With this information, the website can statistically show data on health, nutrition, and other areas. Even though data donation is not well-known yet, it is worth paying attention to.



Figure 1. Sample Statistics from DataDonors

## Conclusion

As a speaker in Ted Med said, "Data is a non-depletable resource, and the more we make use of it, the more valuable it becomes. We need to give this data a second life." Data are more and more important in the healthcare industry. Organizations (providers, payers, manufacturers, analytic companies, and research institutes) need to choose different strategies to manage their data source. If organizations have built up trust between each other, a sharing-model is a great way to organize data; if there are some dominant forces, they can choose a public/private sector model based on the nature of the organization. Beside mature models, we also have to pay attention to new trends, such as using data from social networks, e-commercial records, and IoT devices, as well as from data donation. Efficient data usage can help healthcare organizations better achieve their "Triple Aim" goals: improved patient experience, improved quality of care, and decreased cost of care.

## Citation

1. Dawson PM, Werkman M, Brooks-Pollock E, Tildesley MJ. 2015 Epidemic predictions in an imperfect world: modelling disease spread with partial data. Proc. R. Soc. B 282: 20150205.
2. Paul J. Wallace, Nilay D. Shah, Taylor Dennen, Paul A. Bleicher, and William H. Crown. Optum Labs: Building A Novel Node in The Learning Health Care System. Health Affairs 33, no.7 (2014):1187-1194
3. Optum, Mayo Clinic Partner to Launch Optum Labs: An Open, Collaborative Research and Innovation Facility Focused on Better Care for Patients.  
<http://www.unitedhealthgroup.com/newsroom/articles/news/optum/2013/0115optumlabs.aspx>
4. UnitedHealth Group: [https://en.wikipedia.org/wiki/UnitedHealth\\_Group#The\\_role\\_of\\_Optum](https://en.wikipedia.org/wiki/UnitedHealth_Group#The_role_of_Optum)
5. Melanie Evans. Healthcare data mining: OptumLabs collaborative begins to offer insight about what works. <http://www.modernhealthcare.com/article/20150926/MAGAZINE/309269979>



6. Am. J. Epidemiol. Planning for the Future of Epidemiology in the Era of Big Data and Precision Medicine. (2015) 182 (12): 977-979. doi: 10.1093/aje/kwv228 First published online: December 1, 2015
7. Dan Verel. Optum Labs making progress in huge data collaboration – 40 million lives worth. <http://medcitynews.com/2014/07/optum-labs-making-progress-huge-data-collaboration/?rf=1>
8. Minnesota Department of Health. Minnesota e-Health Brief
9. Minnesota Department of Health. Minnesota e-Health Initiative. Report to the Minnesota Legislature 2015.
10. athenahealth, Inc. Clinical Integration: 7 Myths and a Blueprint for Success. <http://www.athenahealth.com/whitepapers/clinical-integration-model/>
11. Richard Noffsinger. Building an IT infrastructure that supports clinical integration. <http://www.healthcareitnews.com/blog/building-it-infrastructure-supports-clinical-integration>
12. AMERICAN ACADEMY OF PEDIATRICS. Study shows asthma-related Twitter posts can predict rise in hospital visits. PUBLIC RELEASE: 30-APR-2016. [http://www.eurekalert.org/pub\\_releases/2016-04/aaop-ssa042216.php](http://www.eurekalert.org/pub_releases/2016-04/aaop-ssa042216.php)
13. DataDonors. <http://datadonors.org/>