Mergers, acquisitions, and partnerships are a dominant force in today’s healthcare system. These new organizations, however, will not be functionally integrated until their data is integrated. The de facto solution to integrating data is to rip and replace the workflow information systems, such as enterprise resource planning systems (ERPs) and EHRs, with common vendor products. I would argue that this an incredibly expensive, complex, and low-value option; and that a better option in today’s world is to integrate data content first, at a fraction of the cost, using a data warehouse—or as described later, a data operating system.

Data has a profound impact on healthcare organizations’ pre- and post-merger activities, yet is easily and often overlooked by those who should be the most aware. C-suite executives need to immerse themselves in the technical conversation—the language of IT—because, for better or worse, organizations run at the speed of software. This is true in finance, education, manufacturing, and, especially, healthcare. If U.S. healthcare is ever going to get better, faster, more affordable, and more accessible, we must develop better software and make it available to all parts of the industry.

Everything healthcare organizations need to do, strategically, is either helped or hindered by software and data. Therefore, all C-suite leaders need to see themselves as partial chief information officer (CIO) and chief digital officer (CDO). This is the nature of the world in which healthcare operates now. Delegating this knowledge to others means also delegating the future of the company to others. While trust is a significant component of good leadership, good leaders must understand the domain in which they operate. And software drives the healthcare domain.
The goals of this report are threefold:

1. To comprehend how fully digitized our world has become and understand healthcare’s standing in this world.
2. To position the importance of IT strategy and data convergence in the growing healthcare merger and acquisition (M&A) landscape.
3. To introduce the Health Catalyst Data Operating System (DOS™) platform and its role in facilitating IT strategy for an unprecedented future in healthcare.

Digital Disruption Is Everywhere

We don’t have to look very hard to see how digitized the world has become. As Thornton and Lambert explain in their piece, Digital Transformation or Digital Disruption, the largest taxi company doesn’t own taxis (Uber). The most popular media company doesn’t own content (Facebook). The largest lodging company doesn’t own property (Airbnb). And the largest software vendors don’t write the apps (Apple and Google). It’s a very different world (digitally and economically) than we have ever seen before. The rules have been turned upside down about what makes great products and great economic models.

In healthcare’s digitized future, the world’s largest, most successful healthcare and health management companies won’t own hospitals. This will be true, for the most part, by 2025.

As we look at the need to digitize healthcare, it’s interesting to note some parallels between preventive healthcare and car maintenance. It’s often said that an ounce of prevention is worth a pound of cure. Every 10 hours, Tesla collects the equivalent of one million miles of driving data. This telemetry and performance data—the same kind of data, categorically, that healthcare needs to collect about patients—amounts to 25GB of data per car per hour. Not counting digital images, healthcare collects only 100MB of data per patient per year, on average.

Digitizing any industry boils down to two fundamental requirements: digitizing the assets of the industry and digitizing the processes that affect and surround those assets. For example, in commercial aviation, the aircraft, as the assets, would be digitized and baggage handling, ticketing, maintenance, etc., as the processes, would also need to be digitized. Similarly, in healthcare, the patient is the asset and registration, scheduling, diagnosis, orders, etc., are the processes requiring digitization. Figure 1 summarizes these concepts.
We’ve Barely Begun to Digitize Healthcare

Unfortunately, healthcare has a long way to go to truly digitize its assets (patients) and its processes. Health systems have invested a lot of time and billions of dollars into EHRs, which digitized the process of clinical encounters and creating billable events; however, EHRs do not digitize the patient. That burden falls to the clinicians to do during a clinical encounter by typing and clicking through today’s EHRs.

According to a 2010 model developed by the University of Wisconsin Population Health Institute in collaboration with the Robert Wood Johnson Foundation, 80 percent of the factors affecting health outcomes fall outside the traditional delivery of healthcare (e.g., diet and exercise, income, education, etc.). This leaves us with at most about 20 percent of the necessary data, which means we have at most a 20 percent influence on length and quality of life (in the areas of access to care and quality of care).

To capture the significant breadth of the human health data ecosystem and fully digitize the patient, health systems need to expand their skills and data ecosystems up and down the patient information continuum; this means collecting data about behaviors, social and economic factors, and the physical environments of our patients, and combining those factors with the clinical care we provide today, as seen in Figure 2.
Figure 2: 80 percent of the factors affecting health outcomes fall outside the traditional delivery of healthcare

The only way we can fully digitize healthcare is by prioritizing IT strategy and data integration. Solid practices in both are not only imperative for the future of healthcare delivery, but will be a survival asset as the industry is increasingly defined by M&As and partnerships, and consequent integration demands.

**Digitization Means IT Strategy and Data Integration Are Critical in an M&A Environment**

IT strategy is increasingly important in a highly active M&A healthcare environment, especially given the growing digitization of healthcare. At the end of 2017, PwC reported 13 straight quarters of more than 200 M&A deals in healthcare, totaling $175.2 billion.

The top-performing M&As across all industries focus first on data integration and have a plan to do so within six months post-merger. A company is not successfully integrated until its data is integrated. Additionally, 40 percent of the M&A value in healthcare can be tied directly to IT strategy. Without a focus on IT integration, that 40 percent is at high risk.

In today's healthcare, M&A strategy is more about data acquisition than bricks and mortar (i.e., hospitals and clinics). Healthcare leaders must think about the data they're acquiring in an M&A, and how it's going to help them deliver better, faster care to the broader scope of patients. These leaders must develop a data integration strategy to optimize that data.
Healthcare Can’t Fully Digitize with Single-Vendor Solutions

The entire ecosystem of the IT infrastructure needs an integration strategy focused on data. Ripping and replacing EMR and ERP systems with a single common vendor tends to be a default strategy in healthcare, especially among C-suites; but I firmly believe this is not an affordable or timely strategy, and it doesn’t enable IT integration or support data as a strategic asset.

In the 1990s, healthcare was characterized by a best-of-breed architecture (non-integrated systems that were superior in a specialized area, but limited to that area). This architecture used HL7 engines and messages to tie together disparate systems, such as EMRs, lab systems, radiology, pharmacy systems, registration, scheduling, and billing. Best-of-breed architecture had flexibility, but it was fragile and expensive to maintain.

As healthcare IT leaders decided it wasn’t worthwhile to maintain those HL7 interfaces, they migrated towards monolithic single-vendor models. Inadequate solutions, the single-vendor models are difficult to upgrade due to the tight coupling between all these systems (versus loose coupling with HL7). Figure 3 shows the evolution of healthcare IT from best-of-breed to single-vendor solutions.

![Evolution of Healthcare IT](Image)

**Figure 3: The evolution of healthcare IT**

Single-vendor systems are also less flexible, as users can’t easily move products and vendors in and out of this architecture.

A Data Integration-First Strategy Is a Better Choice than Rip and Replace

Dale Sanders estimates the rip-and-replace process takes three to four years to complete, which translates to 74 percent over schedule, 59 percent over budget, and 56 percent less value than predicted. In the meantime, your risks as an organization in a fast-moving industry are going up, and your margins are going down.
Today’s healthcare systems contain dozens of applications. Based on Dale’s personally calculated numbers, drawn from firsthand experience, ripping and replacing an EHR is going to cost a minimum of $13,000 per employee (with some organizations having 8,000 to 10,000 employees). At a minimum, organizations will spend $41,000 per physician to rip and replace an EHR.

According to a [2017 survey](#) of 1,100 healthcare officials, 61 percent of healthcare officials indicate a terrible or poor ROI on EHRs. If we largely consider investment in EHRs as terrible or poor, it doesn’t make sense that our default solution to integration is to rip and replace EHRs and invest more in them.

With a good data integration strategy (versus rip and replace), new technology and evolution can extend the lifecycle and value of existing EHRs. Organizations can survive by reinventing the EHR midpoint in its maturity lifecycle, before it’s too late in the lifecycle and the product exits the market.

Normally, you could extend the product lifecycle by refreshing the technology, but most EHRs are based on older technology and software applications, making this harder to do. Unless they reinvent themselves now with updated technology, all EHRs risk being commoditized and replaced by more innovative products.

We can, however, stretch the lifecycle and the value of EHRs with a data operating system (discussed more in depth later) and open application programming interfaces (APIs). If the EHR vendors open up their APIs, we can actually reinvent them and increase their value, similar to what Microsoft has done with its products.

**Healthcare Mergers and Acquisitions Require an Update to the Traditional Technology Stack**

One of our main challenges in M&A data optimization is that much of healthcare runs on a traditional technology stack (Figure 4) that doesn’t support IT integration.

![IT Integration vs. The Technology Stack](image-url)

*Figure 4: IT integration versus the traditional technology stack*
Technology Stack Layer #1: Computing Infrastructure

The lowest layer of the IT stack in Figure 4 contains computing infrastructure (servers, networks, data centers, and storage). The most important part of the IT strategy at the lowest level is to plot a strategy towards the cloud (versus organizational data centers). The strategy needs to involve the large cloud vendors (Amazon, Google, and Microsoft). Healthcare CIOs need to learn to leverage the computing resources and infrastructure management that’s available in the cloud instead of investing more in their own data centers. These data centers aren’t scalable anymore; they’re also a lot less secure than what Google, Microsoft, and Amazon can provide.

Technology Stack Layer #2: Databases and Operating Systems

The next layer in the stack is the databases and operating systems (e.g., Oracle, SQL, MySQL, Hadoop, Spark, iOS, Android, Linux, and Windows). Activity in the second level of a technology stack starts to affect the applications above it, making it harder to integrate at this level without disturbing the applications that reside on top of them.

Technology Stack Layer #3: Data Content

By starting IT integration strategy at the data content layer, the M&A achieves value faster. Current healthcare software applications are difficult to integrate and consolidate because consolidation is vertical and impacts the entire stack (sometimes down to the computing infrastructure and certainly through the databases and operating systems). To achieve consolidation at the data content layer, the current healthcare software environment requires integration of layers one, two, and three, which is expensive. With the introduction of the data operating system to this layer (more on this later), users will be able to peel data content from data models, populate new applications with it, and reuse and repurpose the data without vertically affecting the other three layers of the stack.

Technology stack layer #4: Software Applications

Healthcare tends to start M&A IT integration at the software applications (e.g., EHR, HR, finance, email, web, etc.) layer, which doesn’t work because, in the traditional technology stack, application developers at this layer can’t access the data in the other layers. More effective IT integration relies on infrastructure (a data operating system) that gives developers complete access to the data they need.

Now that we understand the barriers to data integration with single-vendor systems and the traditional technology stack and the limitations of a rip-and-replace approach, we can understand how the solution, the data operating system, optimizes the technology stack for data integration.

The Data Operating System Supports a Fully Digitized Healthcare World

Getting back to the technology stack (Figure 5), we have work to do in the second layer from the top (domain-specific, or healthcare, data content) to make it easier for software developers and application engineers to work within the data domain of their software. The data operating system focuses on this area of improvement. I believe that the missing piece of the stack is providing
domain-specific data content pre-loaded, pre-managed, and pre-organized for the benefit of software developers at the top of the stack—that also takes advantage of all the incredible capability underneath it.

**Back to the Technology Stack**

- **Amazing capability**
  - Software Development Environments, e.g., Git, Eclipse, Angular, D3, Mono, Node, Python, R

- **This is the layer of the Data Operating System**
  - Domain-Specific Data Content, e.g. Health and Healthcare

- **Amazing capability**
  - Databases and Operating Systems, e.g. Oracle, SQL, MySQL, Hadoop, Spark, iOS, Android, Linux, Windows

- **Amazing capability**
  - Cloud Computing, e.g. Amazon, Microsoft, Google

*Figure 5: The technology stack*

**Two Key Attributes of the Data Operating System**

The data operating system optimizes the technology stack for today’s data integration demands in two key ways:

**#1: The data operating system unwraps data from its applications.**

The data operating system peels away data from its applications and removes the data content layer from the technology stack and populates it in the data lake. This makes the data easily available to the health system and its partners.

Instead of restricting the data for analytics, it is now available to software developers to build workflow applications. If that data is organized and optimized for the developers who are writing applications, they simply need to put the new applications at the top of the technology stack. The other technology layers have already done the heavy lifting of aggregating, organizing, and making the data available. With this process, the new applications contribute data back into the data operating system, which is a considerable innovation because it enables health systems to add previously under-reported data (e.g., patient-reported outcomes) to the data lake.

**#2: The data operating systems supports personalization and broader data management.**

The data operating system can accommodate constantly updated raw and organized data from multiple transaction systems within a domain (e.g., healthcare). This platform enables rapid development and
changes to the software applications built upon it. Using custom tailoring concepts, the data operating system is a combination of the following:

- The transaction level functionality of an HIE.
- Record sharing functionality, including sharing single records between disparate applications.
- Clinical data repository functionality for clinical care and research.
- A read-only enterprise data warehouse (EDW) that also makes data available as a read/write environment for application developers.

The data operating system amounts to what Gartner calls a hybrid transactional/analytic processing architecture (HTAP). As a hybrid, HTAP is superior to either one of its parent architectures, in one or more traits. The Gartner report states, “Traditional data warehouse practices will be outdated by 2018, so we have to evolve toward a broader data management solution for analytics.” Evolving toward broader data management is exactly what the data operating system and HTAP are doing.

Seven Attributes of Health Catalyst’s DOS in IT Integration

Seven attributes of DOS make it a new concept:

1. **Reusable clinical and business logic:** Registries, value sets, and other data logic lie on top of the raw data, so users can access, reuse, and update the logic through open APIs; this enables third-party application development.

2. **Streaming data:** Near- or real-time data streaming from the source, from expression of that data through the data operating system, can support the transaction-level exchange of data or analytics processing.

3. **Structured and unstructured data integration:** The data operating system integrates text and structured data in the same environment.

4. **Closed-loop capability:** The methods for expressing the knowledge in the data operating system include the ability to deliver that knowledge at the point of decision making, including back into the workflow systems, such as the EHR.

5. **Microservices architecture:** Open microservices APIs exist for data operating systems, including authorization, identity management, data pipeline management, and DevOps (software development and software operation) telemetry. These microservices also enable developers to develop third-party applications on the data operating system.

6. **Machine learning:** The data operating system runs built-in machine learning models and enables rapid development and utilization of these models.

7. **Agnostic data lake:** The data lake can function with any vendor’s database management system, as some or all of the data operating system can be deployed over the top of any healthcare data lake. The reusable forms of logic must support different computation engines (e.g., SQL, Spark SQL, and SQL on Hadoop).
Rather than ripping and replacing older technologies, the data operating system can extend their lifecycle and upgrade them to handle the increasing data integration demands of an M&A environment. Starting at the bottom of the technology stack in Figure 4, we have amazing capabilities in cloud computing (e.g., Amazon, Microsoft, Google). The databases and operating systems sit on top of the incredibly capable cloud computing layer. Today, technologies such as Oracle, SQL, and MySQL are available, more affordable, easier to maintain, and incredibly effective and flexible.

One of the reasons a rip-and-replace strategy doesn’t apply today is that modern software is developed around microservices architecture (the fifth attribute above). In microservices architecture, a single application is developed as a suite of independently deployable services. Each service runs on its own process and communicates with a mechanism, such as an API.

Amazon, Facebook, and a lot of smartphone apps constantly undergo small updates. Ripping and replacing these applications is no longer necessary because they’re constantly evolving through minor updates. A lot of healthcare IT vendors face challenges now because they started their software engineering environments decades before microservices were standard. The rip-and-replace mentality (the six-month upgrades that typically required for ERP and EHR systems to stay current) is a symptom of an old, non-microservice architecture.

Figure 6 is a diagram of DOS. This architecture is the future of not just healthcare software, but also other industries (which are already using this architecture).
Starting with low-level data, DOS peels data out of the hundreds of different systems that are currently in an organization. That data is now available for software developers, not just analysts, as it would be in the traditional data warehouse. The top layer leverages all the software tools available in the microservices architecture, which allows users to build applications against the data operating system and against third-party applications.

**The Cost of the Data Operating System Is Low Risk**

Data operating systems and their recent predecessors (EDWs) cost a fraction of ripping and replacing EHRs, EPRs, and other applications. In an M&A, if you’re trying to first achieve data integration, you need to start with some version of a data operating system or, if nothing else, an EDW.

It typically costs $3 million to $5 million to install a data warehouse in a medium- to large-sized organization. The data warehouse can be deployed and operational in a few months; certainly, within the six-month window that McKinsey suggests for a successful M&A. The data warehouse then costs about $2 million to $3 million per year to operate and evolve, which works out to about $340 per employee versus $13,000 per employee to rip and replace an EHR.

Even if an organization is committed to a rip-and-replace strategy, it can still benefit from spending a small amount (a low-risk number compared to $13,000 per employee) to achieve rapid value from data integration using a data warehouse or data operating system.

**IT Integration Is King in a Fully Digitized World, Especially for Healthcare Mergers and Acquisitions**

The new organizations that result from M&As will only be fully integrated when the acquiring organization’s and target organization’s data are integrated. This makes an IT strategy that prioritizes data integration and uses a data operating system (or EDW) an essential component of an M&A—and of healthcare’s digitized future.

The wrong IT strategy (rip and replace) might haunt M&A value for decades, providing only incremental value for massive investment. The digital future of healthcare is about data that persists for a patient’s lifetime and the applications that are going to turn constantly on top of that data. The optimal IT strategy demands adaptability and freedom of choice in software. For successful M&As, acquiring organizations need to focus on integrating data first, not applications, and make sure they achieve IT integration through a data operating system or EDW.
About Health Catalyst

Health Catalyst is a next-generation data, analytics, and decision support company committed to being a catalyst for massive, sustained improvements in healthcare outcomes. We are the leaders in a new era of advanced predictive analytics for population health and value-based care, with a suite of machine learning-driven solutions, decades of outcomes-improvement expertise, and an unparalleled ability to integrate data from across the healthcare ecosystem. Our proven data warehousing and analytics platform helps improve quality, add efficiency and lower costs in support of more than 85 million patients and growing, ranging from the largest US health system to forward-thinking physician practices. Our technology and professional services can help you keep patients engaged and healthy in their homes and workplaces, and we can help you optimize care delivery to those patients when it becomes necessary. We are grateful to be recognized by Fortune, Gallup, Glassdoor, Modern Healthcare and a host of others as a Best Place to Work in technology and healthcare.

Visit www.healthcatalyst.com, and follow us on Twitter, LinkedIn, and Facebook.