

Healthcare Analytics Platform: DOS Delivers the 7 Essential Components

By Imran Qureshi

Chief Software Development Officer

Health Catalyst

Population health, value-based care, decreasing revenues, and increasing costs are just a few of the pivotal issues challenging healthcare IT leaders today. Among other requirements, these issues call for large amounts of data, real-time data access, scalable analytics platforms, and systems interoperability. Healthcare IT must address all these issues, and more, even as technology becomes obsolete at an exponentially increasing pace.

Current <u>healthcare analytics</u> platforms are struggling because they are inflexible and don't provide access to the right data, in the right place, at the right time to effectively support clinical, financial, and operational decision making. The current platform's shortcomings appear in many ways: through process-driven EMRs, expensive data lakes, and unrefined transactional data.

Healthcare desperately needs a new IT model. Understanding the solution requires further investigation into a few specific IT challenges, followed by a close examination of the new model: a data-first, analytics and application platform called the Data Operating System (DOS TM).

Healthcare's Problems Extend to the Healthcare Analytics Platform

The big issues impacting healthcare, such as rising costs and decreasing margins, revolve around healthcare IT and specifically around data.

The United States is projected to spend \$66 billion a year on healthcare IT by 2020. The U.S. healthcare system has spent over \$34.7 billion on Meaningful Use (MU) to digitize data, yet margins are declining, physician burnout is on the rise, and patient satisfaction is low. The data produced from EMRs is plentiful, but improvement through MU is a broken promise because the data is inaccessible.







Inaccessible Data

Different healthcare roles need different data for different reasons:

- IT departments need data to meet the demand for reports and apps.
- Clinicians need access to real-time insights in their workflow.
- Population health leaders need data to scale up from smaller population sizes.
- Financial leaders need data to survive in a decreasing revenue/increasing cost environment.
- Health system leaders undergoing mergers and acquisitions need data to migrate their EMRs to a single system.
- Independent software vendors need clinical data access to integrate into clinical workflows.

Though data demand is everywhere, access to the data supply remains elusive, particularly through EMRs. Clinicians often complain that they constantly put data into the EMR, but never get any data or value out. This is largely because of the current siloed and monolithic analytics environment.

A Fractured Analytics Environment

The common healthcare analytics environment today looks like a series of islands (Figure 1) and navigating them is difficult.

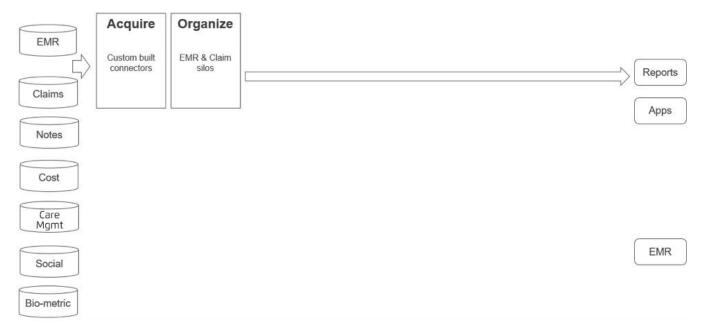


Figure 1: The current healthcare analytics environment

The analytics environment comprises many data sources, including EMR and claims data. Acquiring data from these sources requires custom-built connectors. Every new data source requires new understanding and another custom connector. Then the data is organized, but it still exists in silos. EMR and claims data may exist in one database or the same location, but navigation between the two data types isn't possible (i.e., claims data isn't accessible from the EMR record, and vice versa).







<u>Eighty percent</u> of the data from an EMR is unstructured, including physician notes, which hold significant value but sit completely outside the analytics system. Other data sources—social determinants of health, cost, care management, and biometric—contribute significantly to health outcomes, yet exist in silos.

In the current analytics environment, EMR and claims data is acquired and organized into reports and apps, but this is where content and logic can get confusing. For example, when determining a cohort of patients with diabetes, individual reports may use different logic to define which ICD-10 codes constitute diabetes.

Further complicating the environment is an isolated EMR, which is entirely separate from the analytics infrastructure. It is very hard to get insights generated by the analytics infrastructure in front of clinicians at the point of care.

The entire system is monolithic—users cannot customize it with components and they are predestined to a single configuration.

The current analytics model is the lockbox around healthcare data and the key to unlock it is DOS.

DOS Is the Key to Unlocking Data

DOS is a new data-first analytics and applications platform. Before we talk about the mechanics of DOS, it helps to understand the idea of "data first" and why EMRs cannot trace a data-first trajectory. Then high-level and magnified diagrams will help explain DOS structure and components, and the purpose of each piece in the system.

Process-First (EMR) vs. Data-First (DOS) Design

Processes have driven the current healthcare data environment. The requirement to digitize medical records has spurred massive EMR deployment. EMRs are process-first systems, designed to move data from person to person to follow a process. EMRs collect data within a specific workflow, but don't allow access to, or the import of, data outside this workflow, making it very difficult to do any meaningful analytics on transactional data. Like mainframe computers, EMRs will always be necessary, but they are not the right tool for unlocking data.

To extract value from the healthcare data, we need a data-first approach. In data-first systems, the focus is on the following:

- Oetting as much of the available data as possible.
- Understanding and mining the data for insights.
- Returning those insights back into the ecosystem.
- Presenting data to appropriate users.
- Fueling other apps to create more insights.

The iPhone is a good example of a data-first environment. Various applications built into the iPhone receive data and enter it into the ecosystem in different ways:

- Real-time (live video)
- Near real-time (stock quotes)







- Streaming (music)
- Images (camera)
- Text (notes)
- Cloud (photos and videos)

Other applications consume this data, create insights, then publish those insights into the ecosystem, where they can be picked up by yet another application. In a data-first environment like DOS, data drives the process. Instead of implementing a process and moving data from place to place (like an EMR), data is analyzed and the analysis drives the process.

Five Ways EMRs and DOS Are Fundamentally Different

EMRs were designed to produce documentation for fee-for-service (FFS) billing, which is their fundamental purpose. DOS is designed for analytics and outcomes improvement. The core of EMRs and the core of DOS differ in five other key areas (Figure 2).

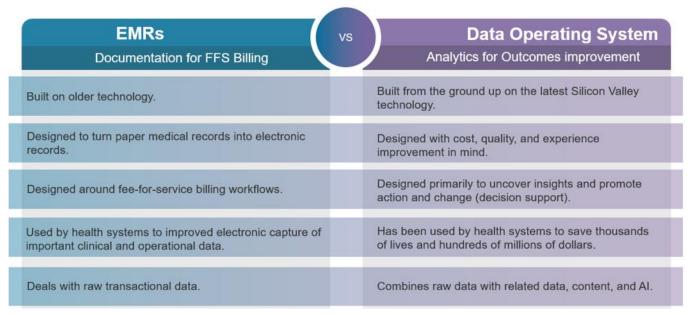


Figure 2: The core differences between EMRs and DOS

- 1. **Technology**—EMRs are built on 20-year-old technology. DOS is built on the latest technology for analytics.
- 2. Purpose—EMRs are designed to take paper medical records and convert them to electronic records. DOS is designed to improve costs, quality, patient experience, and provider experience within the organization.
- 3. Workflow—EMRs automate the FFS billing workflow, a process-first model. DOS uncovers insights and promotes action and change in the organization.
- 4. **Utility**—EMRs capture clinical and operational data. DOS leverages this same kind of data to save thousands of lives and hundreds of millions of dollars (more on results toward the end of this report).
- 5. Data Type—EMRs only deal with raw transactional data. DOS combines raw transactional data with related data, content, and AI to create data that's analytics ready.







The Seven DOS Components

Like the current analytics environment, DOS interacts with multiple data sources, reports, apps, and EMRs. The unique and comprehensive structure, missing from the current model, lives in between these elements. This structure transforms data so reports and apps don't have to, and it includes seven components.

1. Acquire

Healthcare analytics systems need data from many sources, so the cost of acquiring data from any data source needs to be low. DOS builds in connections to more than 200 data sources, including most of the common healthcare data.

2. Organize

The Health Catalyst® Analytics Platform is the organizing function of DOS. It includes all the tools to rapidly consolidate data from every source, precluding the need to custom build anything. Data analysts can create data pipelines to transform data through multiple steps, then leverage data lakes, and Hadoop and Spark infrastructures to process large amounts of data through the system.

3. Standardize

After it's organized, clinical data (e.g., patient populations, diagnosis classifications, medication groupings) can be efficiently standardized through pre-built shared data marts. These can be used out of the box, or users can create their own from built-in tooling.

4. Analyze

Using data and content microservices, the analyze component combines and enriches data, then makes predictions using AI models. This is the process of converting raw data to deep data.

5. Deliver

Once data is analyzed, DOS can deliver it directly to the EMR workflow and right to clinicians' fingertips. Data can also be exported to other workflows, putting it where data analysts are, rather than making them navigate to the data. DOS learns through a closed-loop function. As analysts interact with the system, those interactions feed back into the system and improve the data over time.

6. Orchestrate

Orchestrating all this data requires a central place for storing metadata (i.e., data about data), so people can easily find and understand the data. Orchestration also requires a central place for enforcing security, and a central place to manage the various processes happening in the analytical environment. DOS provides these tools so anyone can find, manage, and secure all the organization's data.









7. Extend

All the pieces of DOS are accessible through open APIs to ensure sustainability of the DOS environment. A marketplace allows users to choose different applications and content pieces for their own data operating system.

This is the high-level view of DOS. The following section will zoom into the operating system for greater scrutiny of each component.

A Deeper View into DOS

What does DOS offer, through each of the seven components, that really helps jumpstart a healthcare system's analytical efforts? A deeper examination reveals DOS's full capabilities (Figure 3).

Detailed DOS Data Flow

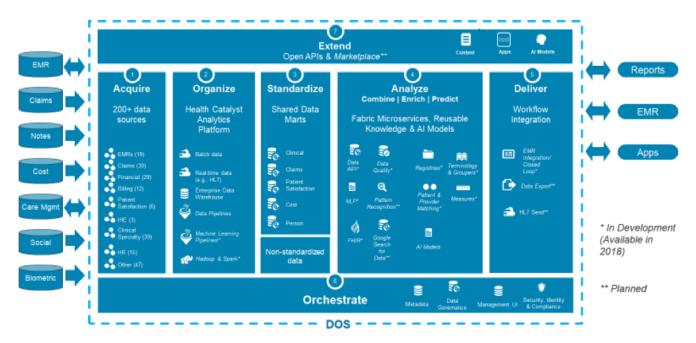


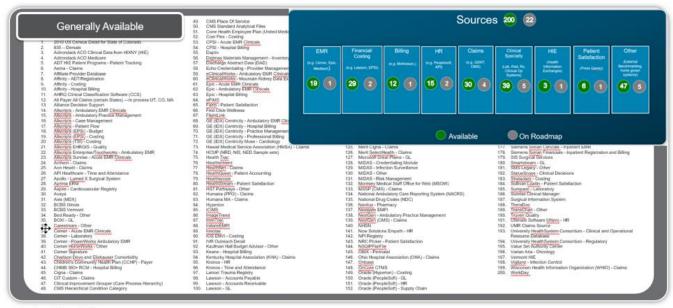
Figure 3: A detailed DOS diagram shows the functions of each component

Acquires Data

DOS builds in connectors to more than 200 data sources (Figure 4), including all the common EMRs, most claims systems, financial systems, billing systems, HIEs, and clinical and patient satisfaction systems.







As of May - 2017

Figure 4: DOS builds in more than 200 data sources, with many more coming in the future

It takes very little effort to bring in data from these systems. DOS already understands them, can map to them appropriately, and can serve up the data for analysis.

Organizes Data

To understand the value of this DOS component, consider how data lakes and data warehouses have been developed and deployed (Figure 5).

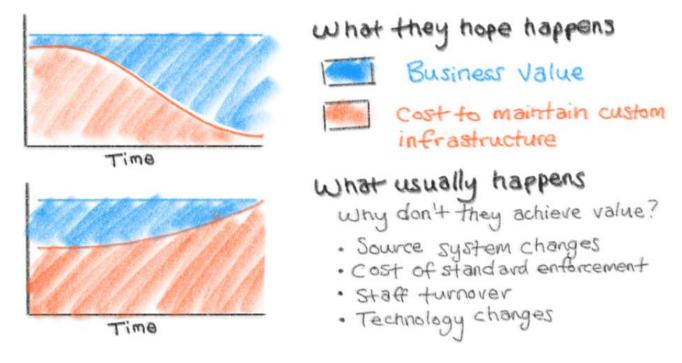


Figure 5: The value-to-cost curve of data lakes and data warehouses





At the outset, the hope is that, over time, the business value of the data lake or warehouse will increase as the costs to maintain it decrease. What ends up happening is almost the exact opposite. Source systems change. An EMR is upgraded. The warehouse needs to pull data from a new claim source. These all incur substantial costs. After a while, staff turnover interrupts continuity, then the underlying technology changes, the data warehouse can't keep up and, as a result, people stop using it.

DOS delivers the desired value-to-cost curve, with value increasing and costs decreasing over time (Figure 6).

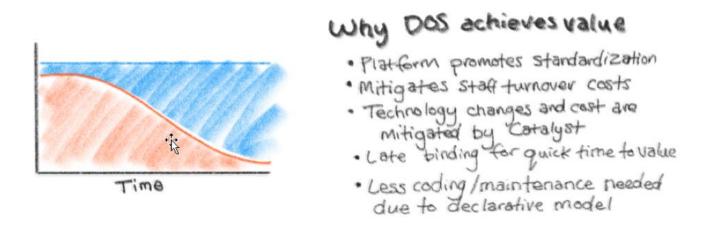


Figure 6: How DOS achieves value

DOS starts with all the necessary building blocks, so there's less programming involved. Much of the data organizing function is performed by DOS "under the hood." As new staff comes on board, they can use DOS's built-in tools (UI and other simple models) to manage and enhance the existing functionality rather than looking through piles of handwritten scripts.

Health Catalyst manages technology changes, so healthcare systems can focus on deriving business value from the data warehouse. The <u>Late-Binding™</u> technology in DOS binds only the data that's needed immediately, which generates very fast time-to-value. More data can be added over time. It's not necessary to create a dimensional model of all the data up front, a process that usually takes years before any value is derived from it.

Standardizes Data

DOS standardizes data through shared data marts (Figure 7), but there's a cost to doing this, so pursue standardization only when there's agreement on the interpretation of that data and when standardization adds value. For example, standardizing the ICD-10 codes classified as diabetes likely makes sense, but creating a grouping for medications that are only ever used by one department may not be worth it.







Figure 7: DOS ships with built-in shared data marts

Standardizing data creates faster time-to-value and allows for building multiple reports and applications without having to create multiple mappings of the same data. Standardizing creates tighter <u>data</u> <u>governance</u> through controlled data access and data definitions.

DOS has built-in shared data marts around clinical, claims, patient satisfaction, cost, person, and terminology. DOS takes the 200+ data sources mentioned earlier and maps them into these shared data marts, giving access to data that's aggregated across every data mart. This mapping effectively disintegrates silos by joining clinical and business data from multiple EMR systems and other data sources, such as claims, GL, and patient satisfaction.

Analyzes Data

The term <u>big data</u> refers to data in terms of volume, variety, and velocity. But the IT vernacular has introduced a new term, deep data, which is about understanding data. The analyze component of DOS transforms raw data into deep data through five steps:

- 1. Related data is combined with raw data.
- 2. The combined data is analyzed to reveal trends over time.
- 3. The combined data is enriched with shared content.
- 4. All or machine learning is added to predict outcomes.
- 5. The result is deep data, which is searchable, longitudinal, meaningful, and actionable.









Figure 8 shows an example of transforming raw data (from a single patient with diabetes) into deep data that predicts risks and suggests interventions.

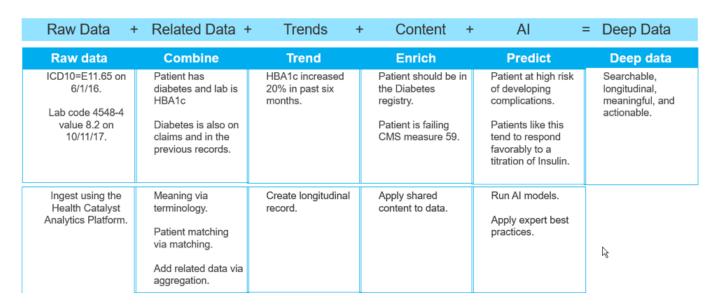


Figure 8: A healthcare example of transforming raw data into deep data

Delivers Data

DOS's delivery component gets the right data to the right place at the right time. The EMR's current model for clinical decision support (CDS) is interruptive. Pop-ups serve as alerts, but don't provide any underlying data so clinicians can understand the reasons behind the alerts. And there's no way to correct a misleading alert. DOS delivery focuses on synthesizing data, so clinicians can choose the right path, rather than conforming to an A or B approach.

The model for CDS in DOS is like the Google Maps model. Google Maps shows a recommended route, but also displays alternate routes and travel times. The user can click any route, even if it's not the fastest or shortest one. The app color codes routes, which helps explain why one route is recommended over another. Users get the benefits of Google Maps, but can still direct it according to their preferences.

This CDS model is built into DOS, as is the technology that can show previously created data and insights directly within existing workflows. Figure 9 shows a sample record of a patient at high risk of opioid abuse. It displays the patient's risk factors, history, and suggested interventions (actions): all relevant insights for the attending clinician. Any actions are directed back into the workflow of the EMR, keeping all documentation internal. This augments the EMR and delivers the power of deep data and insights to the clinician's fingertips.







Figure 9: DOS delivers deep data and insights to clinicians at the point of care

Orchestrates

What's the value of data if analysts can't find what they are looking for or can't understand where the data came from? Many organizations try to build their own orchestration tools but end up spending a significant amount of time on these tools instead of focusing on getting business value. DOS provides orchestration tools around creating and managing metadata, enforcing security, and managing all the analytical processes.

Extends

DOS can be extended through open APIs and a marketplace of content, apps, and AI models. This extend component allows customization for every healthcare enterprise.

APIs are important because they accelerate app development. With APIs, developers can customize existing tools and services to build only what they need. DOS ensures that everything that needs to be built is done so within the operating system and connected by an API fabric.

Several fabric APIs are already available in DOS:

- Metadata Services API—Read/write metadata stored in the Health Catalyst Analytics Platform, such as schemas, additional attributes, etc.
- **Data Pipeline API**—Start, stop, and monitor data pipeline jobs running in the Health Catalyst Analytics Platform.









- Service Discovery API—Register internal microservices and find other microservices.
- Identity API—Add authentication to an internal microservice using various underlying authentication providers, such as Active Directory, Okta, Azure AD, etc.
- Authorization API—Store and retrieve permissions securely from a central data store.
- Realtime API—Receive HL7 and EDI/X12 messages; subscribe and process the messages in XML form.
- Terminology API—Store, retrieve, and search terminology mappings.

Three fabric APIs are in development:

- Data Service API—Access all data in DOS via a REST API.
- EMR Closed Loop Service API—Show data and insights as worklists or in the patient chart inside EMR workflows.
- Machine Learning API—Allows client apps to run any R or Python models via a REST API without needing to install additional infrastructure.

DOS Learns Within a Closed-Loop

A huge value of DOS is its ability to learn from data within a closed-loop system. After clinicians receive data, there's the possibility they won't agree with it or that the data is non-responsive (i.e., it's read-only data). Data should be alive (i.e., read-write) and be able to transmit insights across the care spectrum. This is what happens in a closed-loop system (Figure 10).

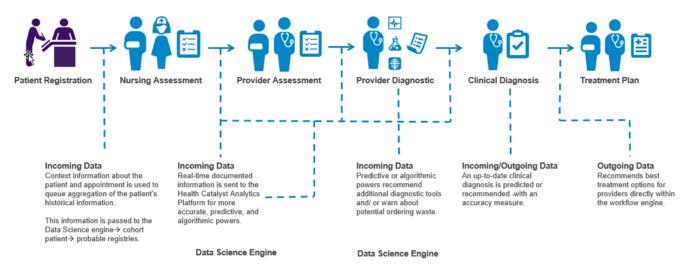


Figure 10: Data is improved through a closed-loop system inside DOS

In the example shown in Figure 10, data is generated when a patient registers. This data is analyzed to see what stands out in the patient's history. As clinicians make assessments, the patient is assigned to appropriate cohorts and <u>registries</u>. This data feeds back into the system in real time, which continues to update insights even as additional diagnoses feed into the system in real time. <u>Risk predictions</u> determine additional diagnoses or interventions, which are then presented as part of the treatment plan. This process involves different clinicians who interact with the data; the closed loop keeps bringing all the information back around so the data changes and improves over time.







Real-time Data Processing in DOS

Data consumers want their data sooner than the next day, but faster turnaround has been next to impossible. Real-time data processing follows a complex pathway that must receive messages, process them asynchronously, establish data reliability, and map the data from HL7 to a database (Figure 11).

Receiving messages

Have to setup, manage, and monitor an interface engine for receiving.

Asynchronous Processing

Need to process messages asynchronously so you don't block the sender.

Have to setup, manage, and monitor a queue service for reliable processing of real-time data.

Reliability

Set up High Availability (HA) clusters.

Handle out of sequence and duplicate messages.

Map HL7

Write code to extract data from HL7 and save to database.

Create UI for monitoring and handling errors.

Manage data consistency (message updated one entity but not the other).

Figure 11: Real-time data processing follows a complex pathway

Real-time data processing is vital to healthcare, but too difficult for healthcare systems to develop on their own. DOS provides a four-step real-time processing solution:

- 1. Point the source (e.g., EMR) HL7 feed to the DOS endpoint.
- 2. Use an HL7 content starter set to automatically map data into database entities for quick time-to-value.
- 3. Data shows up in your database, already transformed.
- 4. Edit the HL7 content starter set to customize.

Installing an interface engine isn't necessary; all the asynchronous processing is built in. High availability and deduplication are built in. Data is already converted to XML, so it is easy for applications to consume. Applications can subscribe to the queue server and receive processed messages in real time too.





DOS Attributes and Benefits

DOS is a complete analytics ecosystem that has seven key attributes:

- 1. **Reusable clinical and business logic:** Registries, value sets, and other data logic lies on top of the raw data and can be accessed, reused, and updated through open APIs, enabling third-party application development.
- Streaming data: Near- or real-time data streaming from the source all the way to the expression of that data through DOS that can support transaction-level exchange of data or analytic processing of data.
- 3. **Integrates structured and unstructured data:** Integrates text and structured data in the same environment. Eventually, it will incorporate images, too.
- 4. Closed-loop capability: The methods for expressing the knowledge in DOS include delivering that knowledge at the point of decision making; for example, back into the workflow of source systems, such as an EMR.
- 5. **Microservices architecture:** In addition to abstracted data logic, open microservices APIs exist for DOS operations, such as authorization, identity management, data pipeline management, and DevOps telemetry. These microservices also enable third-party applications to be built on DOS.
- 6. **Machine learning:** DOS natively runs machine learning models, and enables rapid development and utilization of machine learning models, embedded in all applications.
- 7. **Agnostic data lake:** Some or all of DOS 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, SQL on Hadoop, et al.).

DOS's attributes generate many benefits that impact leaders throughout the healthcare organization:

- IT leaders—Meet the increasing <u>organizational</u> demands for deep data and <u>predictive analytics</u> through the use of SQL, Big Data, and Al capabilities, without building huge infrastructures.
- Clinical leaders—Get real-time insights using all of the patient's data without leaving their workflow, enabling them to interact more with their patients.
- Population health leaders—Access deep data (claims, clinical, costing, and socioeconomic) across the entire care continuum. Data enhancements (e.g. risk scores, attribution models, patient matching capabilities, and AI) deliver more effective and financially sustainable, scalable population health programs.
- Financial leaders—Access deep data (clinical, costing, and operational) across the entire care delivery system to thrive in a world of decreasing revenue and increasing risk contracts.
- Health system leaders—Avoid having to rip and replace expensive technologies to get the deep data they need to succeed.
- Independent software vendors—Eliminate the frustration of waiting for access to clinical data and the insurmountable task of getting their tools integrated into clinical workflows.







DOS enables many applications and services, which sit on top of the operating system (Figure 12). Several cross-cutting analytic applications and vertical, use-case driven applications facilitate many healthcare programs. Health Catalyst builds some of these applications, but third parties can also build their own applications. In addition, Health Catalyst provides a line of strategic consulting services to implement the technologies and ongoing operating system improvements.

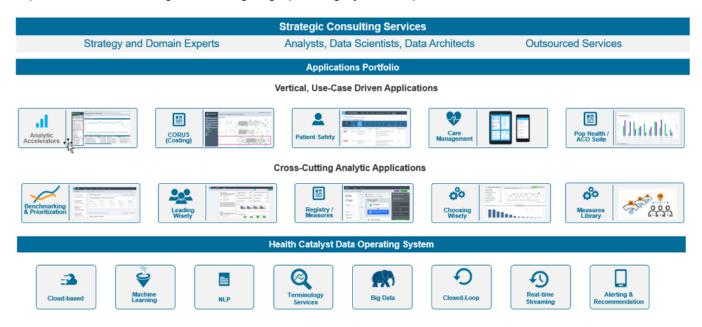


Figure 12: DOS-enabled applications and services

DOS is Built with Reliable and Proven Components

Components built over the last decade, and now available in DOS, have shown value in millions of dollars saved through process and outcomes improvement (Figure 13).

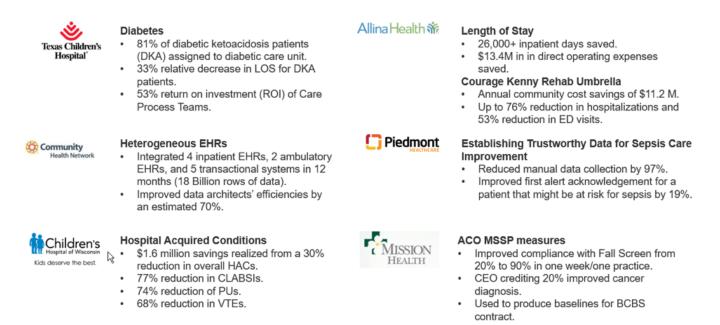


Figure 13: Components of DOS have driven outcomes improvements for many healthcare systems







Healthcare Needs DOS to Unlock Healthcare Data

Healthcare needs a new IT model to accommodate a rapidly evolving data landscape. EMRs, raw transactional data, and data lakes are useful, but no longer sufficient. The DOS approach is the only way to unlock healthcare data.

DOS easily acquires data from more than 200 data sources, organizes it with a built-in analytics platform, leverages standard and custom data models, accepts real-time data, and brings report logic into a common layer. DOS conveys data to the EMR and other workflow tools so clinicians can act on it. The closed-loop system keeps the data live and updated, and APIs extend the operating system.

Leaders across the healthcare enterprise can benefit from DOS. DOS also benefits independent software vendors, giving them access to clinical data through built-in APIs, so they can easily integrate data into clinical workflows.

Despite its huge scope and vast capabilities, DOS can get up and running quickly, delivering value within just a few months. DOS is a foundational and transformational system that has the potential to change the healthcare economy and massively improve financial, clinical, and operational outcomes.

About the Author



Imran Qureshi, Chief Software Developer Officer

Imran Qureshi is the Chief Software Development Officer at Health Catalyst where he is responsible for all software development in the company. He also leads the Engineering team building the Data Operating System (DOS). Before Health Catalyst, Imran was the Chief Technology Officer at Acupera where he led the team that built the care management platform that was successfully implemented in Ascension, Montefiore, Kaiser, and other health systems. Prior to that, Imran was VP of Engineering at CareAnyware, where he led development of the largest cloud-based EHR for Home Health and Hospice. He spent 12 years at Microsoft, including building the slideshow for PowerPoint and building the email experience for Hotmail. Imran holds several patents and has a Computer Science degree from Stanford University.







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.

3165 East Millrock Drive, Suite 400 Salt Lake City, Utah 84121 ph. 800-309-6800







