Dale: Thanks everyone for joining us today. We’ve got a lot to cover so I’m going to jump right into this. I’ll mention that we’ll push through the slides within an hour. I don’t imagine that we will get to questions and answers within an hour, but I’ll stick around after the hour for as long as there’s interest in doing so to try to answer questions. Okay, let’s move out here.

Dale: First I want to say that in these webinars I don’t normally sell Health Catalyst® for those of you that have watched previous webinars in the past. I offer advice from past experience, mostly mistakes that I’ve made or watched others make. I advocate change. In this webinar it’s a little different, I’m selling Health Catalyst® a little bit but only as evidence that we practice what we preach. In this case the development of this Data Operating System. And of course there’s also some subtle advice buried in the selling that, you know, if we’re developing and building the Data Operating System maybe other folks and vendors should think about it as well. So that’s— You’ll see a slightly different take on this webinar than you have in the past, but it’s still kind of a light sell on Health Catalyst®, it’s still more about trying to motivate change and advocate for new and different things.
Dale: So the story of today’s meeting will come in four parts. First, let’s jump right in and identify what a Data Operating System is conceptually and what makes it different than traditional data warehousing? Why do we need one now in healthcare more than ever? What are the options for implementing and can it be implemented? And is it real or just another buzz phrase? And I noticed in my little cartoon here that it says “One upon a time...” not “Once upon a time...” so bad news about my cartoons is that they can’t be spell-checked yet, but I suspect that Google is probably working on that end of that.

Dale: Okay, I also want to start off by acknowledging and giving thanks to our entire product development team for the incredible performance they’ve had in the last 18 months. I’ve either built things from scratch or been involved in turn-arounds my whole career, that’s the two flavors of my career, and I’ve never been associated with this much change and productivity in 8 months ever in my life. So many thanks to the product development team and everyone that contributed here at Health Catalyst®. I also want to call out and thank Brian Hinton, Imran Qureshi and Shawn Stohl in particular for their brainstorming and the engineering and implementation of the Data Operating System and everything they are doing in that regard. I want to thank Rus Tabet, one of our user interface and graphics experts, you’ll see his illustrations and cartoons in this slide deck and you’ll be able to tell the difference between his and mine—that’s for sure. Of course I want to thank the other artists whose work inspired some of the doodles in these slides as well.

Dale: I want to mention very firmly that I come at this from a position of humility. We are not holier-than-thou in what we are advocating here. I’m not satisfied as a patient and as a citizen with the current trajectory of digital health—at the macro level. But I can tell you right now—we’re not satisfied at Health Catalyst® with ourselves either. We are far from perfect. We have plenty of flaws and part of what we are doing with this Data Operating System is disrupting ourselves very purposefully because of that.
dissatisfaction with what we’re doing. So, you’ll hear me offer some criticisms about healthcare IT in these slides today. But it’s not from a position of being holier than thou, it’s from a position of being motivated to do something better and taking that accountability very personally ourselves.

Dale: I want to give fair warning to the executives in the audience, and I see that there are a few of those. Get ready to dive into these topics. I believe very strongly that you need to understand these technical topics I see a good percentage of our attendees today are IT, and that’s great, but, I want to remind the executives that the most expensive capital purchase you’ve ever made wasn’t a hospital, it was your EHR. And if you can’t own up to the notion that software now runs your company for better or worse, you’re probably behind the times. I would argue that healthcare CEOs who thrive going forward will understand their software technology and data. They will be the leaders that rise to the top in the next generation of US healthcare. Software runs everything today. I just offer the case in point the ransomware attack on the UK national health system last week. And that’s relatively utilitarian and mundane software and technology issues that brought them down. Had they been keeping up with their Microsoft security patches and operating system none of that would have happened. But healthcare executives of the future are going to have a firm grasp of the topics we’re going to talk about today. It’s really important.

Dale: All right, so let’s offer a definition of a data operating system. This is my first version, subject to improvement by other brains. But let’s go through this, a data operating system combines real-time, granular data, and domain-specific—UG healthcare is the domain, reusable analytic and computational
logic about that data into a single computing ecosystem for application development. So, granular data at the low-level, reusable analytic and computational logic on top of that granular data. And the data operating system can support the real-time processing and movement of data from point to point as well as batch-oriented loading and computational analytic processing on that data. So, what you see here is the merging of what amounts to a data warehouse and an HIE. Conceptually that’s pretty important.

Dale: Here is an architectural diagram of what the Health Catalyst® data operating system looks like. You can see at the lower level the data platform consisting of the catalyst analytic platform, so more or less the batch-oriented, less than real-time analytics calculations and computational services. There’s the core data services that include pattern recognition, NLP governance tools, metadata repositories, data quality tools, across subject areas, and then you have real-time data services on the far right. The ability to stream data, a reference lambda architecture. In more of a generic sense about the announcement that came out about the webinar. Lambda in this context means the ability to process real-time as well as batch-oriented data for analytic in the same ecosystem. There’s actually, emerging from lambdas, an improvement on that architecture called a Kappa architecture. We’ll talk about that later on today, but basically it’s the combination of real-time and analytics processing, and batch-oriented processing in the same environment. On top of that then is the fabric. And the fabric is that first layer of clinical and business logic and services that can be laid on top of the granular data and services beneath it. Including the notion of machine learning embedded as a part of virtually everything that we do, is a native part of the fabric. And then on top of that are the applications that Health Catalyst® builds, and we have those in our portfolio. As well as applications that can be built by hospital and clinic IT application development teams and then 3rd-party applications as well. So that’s the high-level architecture. Real-time and batch-oriented in the same granular data platform. Open API is in the fabric, and you’ll note as I mentioned earlier our emphasis will be on developing fire-based services where possible. And where not published or possible we’ll extend fire or pursue another means, but fire will be our first default for the services in that fabric layer. And then traditional software development in the apps layer. Java, JavaScript, JAX London, that kind of thing. A key point to make here is that the fabric is being developed at Health Catalyst® so that it can lay on top of a variety of topologies and data platforms. Including platforms other than Health Catalyst®, so we want the fabric to be compatible with …, IBM, Oracle, homegrown data warehouses, Epic, Cerner, be able to cross all of those so that we can lay this fabric down whether a Health Catalyst® granular data platform exists underneath it or not. That’s a really important part of the future. So, I want to summarize quickly for a minute, I’m going to go back to this slide, what’s the difference between this and a traditional data warehouse? And I think it’s encapsulated in three
You can support everything out of one ecosystem. There’s a fabric of micro-services and data bindings that can lay on top of any data system, not just Health Catalyst®, that’s a big difference. And it’s a system that’s designed from the beginning to use open API’s and make the data platform and fabric services available to 3rd-party application developers. So, we’re building for the future for extensibility, realizing that Health Catalyst® will never build all the great applications that there are in the world. Those are the three big differences between a traditional data warehouse and what we’re calling the data operating system.

---

**Seven Attributes of the Healthcare Data Operating System**

1. **Reusable clinical and business logic:** registries, value sets, and other data logic that lies on top of the raw data to access reused, and updated through open APIs, specifically enabling 3rd party application development against it. Designed from the beginning with that in mind, not retrofitting it, as we are doing right now with APIs—Open APIs in the healthcare IT environment. This is being designed from the beginning to support 3rd party application development.

2. **Streaming data:** near or real-time data streaming from the source all the way through to the expression of that data through the DOS, which means we can also support transaction-level data exchange from point to point. It integrates structured and unstructured data so text has to be a part of the healthcare data operating system. Eventually we’re going to incorporate images as well, it is a minimum that it has to include text. It has to have closed loop capability, that is methods for expressing the knowledge in the DOS. Including the ability to deliver that knowledge at the point of decision making. For example, backing the workflow of source systems such as an EHR. It has to be built around a micro-services architecture meaning the ability to update constantly, continues development and continuous release. Not the giant painful upgrades that we’re accustomed to and we’ve been desensitized to in healthcare. Micro-services updates. Things like authorization, identity management, the data pipeline management, DevOps telemetry. These micro-services will also enable 3rd parties to develop applications against the DOS, they don’t have to recreate those. An important attribute is number 6, that is machine learning. DOS has to natively run machine learning models and enable rapid development and utilization models embedded in all applications. That’s a real strength of the Big Data, a dupe ecosystem that came out of Silicon Valley. It is natively designed to support machine learning and computational analytics that traditional relational databases are not. And finally the last is that has to support an agnostic data lake. Some or all of the dots can be deployed over the top of any healthcare data lake. All those that I mentioned earlier including those listed here and we’ll be holding ourselves accountable in Health Catalyst® for our progress in our ability to meet these 7 attributes. But I think again this applies to any vendor or any locally
developed solution. These are the attributes you have to achieve to become a healthcare data operating system and meet the future needs of data in our industry.

Dale: So let’s talk about why the DOS is important now in healthcare, kind of the business reasons.

Dale: There’s a convergence of things happening in the country right now. We’ve got new Big Data technology that’s emerged as a consequence of open source collaboration in Silicon Valley. Incredibly fortunate for us to be here at this time in history to take advantage of what Facebook, Google, Amazon, Twitter and others have developed and put out for us to consume free of charge. It’s unbelievable the value that we’re going to derive from that and we are— should all be very grateful for what they’ve done for us. No surprise that healthcare expenses are continuing to go up. You know we’re going to hit 20% of GDP very soon. It’s cannibalizing our economy, so if we can’t change the trajectory through digitization we’re all going to be in trouble. We’re eating up the future of the country in health care expenses. Physicians are burned out and in large part they’re burned out because of the technology that they are now using. It’s taking time away from clinical care and human decision making. More than 50% of their time now is spent in front of a computer instead of a patient. We have to change that. PHRs have largely been a failure unfortunately. Until they’re successful we will really never put patients at the center of care. They’ve been a failure on several fronts, interoperability being one of those. But also because it’s just no one’s interest right now economically to give that data up to a PHR that’s transportable and can be moved from facility to facility. We have to change that. FHIR is emerging, I’m very optimistic about FHIR. I’ve had concerns with HL7 in the past and the message oriented architecture of HL7, but due credit to the rebels within HL7 who started FHIR. I have great hope for that and optimism for FHIR in the future and it’s a very well-founded framework. HIEs have largely been a failure. I know— I wouldn’t know— I can’t imagine anyone arguing from the position of data or logic
that HIEs have been successful on a number of levels. Economic models, technically usability of the data within the HER, they have not worked and it’s time to do something different. And the good news is we’ve adopted EHRs like never before so we’re starting the digitization process. We have data now that we’ve never had in the past. The problem is of course that those EHRs are still not well received. 54% of those polled last year were dissatisfied with their EHRs. So we can do things with the dots to make EHRs better. It’s interesting because of course all the EHRs want to get into the analytics and what amounts to the DOS business as well. So they don’t have a lot of motivation to collaborate with us but the irony is we can make EHRs more valuable with what we’re building in Health Catalyst®, and that’s what we hope to do.

Dale: Another interesting thing that we need to— I think appreciate and recognize is the content is King but the network is Kong. And what I mean by that is you look at modern businesses, data content is becoming the driving force behind a business strategy and business value. So GE, Tesla, Google, Facebook, all these companies, United Healthcare, Amazon, they understand the value of data content and they’re going after it, Optum you could add to this. But the network around that data is as important as the data content itself. So think of Metcalfe’s Law and then you know the value of the community around data versus the hub and spoke model. Sticky relationships take place and occur when you have great data content and you put a network of people around that content. Right? There’s a reason that Google Plus never took off and Facebook is still accelerating. It’s the combined content as well as the network of people that make that stick relationship with Facebook difficult or impossible to move and transport to Google Plus. So executive of the future need to understand that you’ve got to have data content and you have to build a network of people patients, healthcare providers, physicians, researchers, around that data and that’s going to create the sticky relationships for a successful business going forward.
Dale: The folks at McKinsey, every once in a while, produce some incredibly insightful information and analyses and I love their digitization index. And the— And the equation for their digitization index is described there in the sub-bullet. Data assets times data usage times skilled labor around that data all combined to create a digitization index. How many assets— How— What kind of data do you have are you using it and do you have the skilled labor to actually take advantage of that data. Really good framework and their analysis is probably no surprise to any of us. Healthcare comes out as having a very low, almost the lowest, overall digitization of any industry. But what’s also interesting is when you plot that against the y-axis of three year changes in post-tax profit margins, healthcare also suffers. So C-levels need to appreciate this— that there’s a strong correlation between your digitization index and your post-tax profit margin. And as margins get harder and harder and tighter and tighter to manage going forward the executives need to understand the importance of digitization to retain whatever competitive edge they might still have. Really important concept.

Dale: So, C-level advice here is the population health, value based care, precision medicine are all about data. You need a strategic data acquisition strategy that goes beyond bricks and mortar. What data do you need for pop health, risk contracting, and precision medicine and how do you plan to acquire it? You need a chief analytics or chief data officer, that’s critically important. You need to ask yourself if that’s gong to be your CIO or not. But in any case you need to appoint someone to fill that role, to manage this critical asset and be the executive cheerleader for it. Your physicians and nurses are over-measured, they’re undervalued and it’s in large part because they’re slaves to data entry and poor software. So you need to push all vendors to follow modern open software APIs including but not limited to FHIR. It’s critically important for your data strategy, C-levels, that you push vendors to support these open APIs. This is not something that you want to relegate to someone— It’s sort of minimized in the organization. You need to be aware of the impact that software has on your business and what these open APIs can or can’t do for you. And I would argue that you need this concept of a data operating system and you can create that by leveraging and expanding the capability of your enterprise data warehouse if you have one.
Dale: So I’ve listed out here a number of what I call DOS needs. I think there’s six or seven in the following slides here. The first is what I call a shark tank story from a business perspective. I was in the audience of some healthcare IT startups about a year ago. They were pitching great software applications and creative ideas to me about healthcare. And as brilliant as they were none of the solutions offered any appreciation for an answer for the underlying healthcare data that they needed. So they all had pretty decent demo data, but they didn’t have an answer for the massive acquisition of data and the scalability of that acquisition across an entire industry. Nor did they have an answer for both clinical and business logic that resided on top of that data. So in my head the whole time I’m thinking we’ve got to give these young folks with these great ideas and applications the data that they need. They cannot possibly afford to build the data infrastructure and skills that we have in Health Catalyst®. Not only can they not afford it, but the industry can’t afford it. It is not scalable and I’ll show you more evidence of why I believe that here shortly.

Dale: So it was interesting, right? I realized. And this is where the term the data operating system actually came out of that meeting. What we have is in computer science a great expansion of modern programming languages at the top. The things that we can build quickly with all the different libraries that exist and all the different awesome programming environments is beyond anything we’ve ever had in my career. 33 years of doing this. Not only the languages, but the DevOps tools to support the languages, meaning the ability to push out and manage and measure applications once they’re out in the field. And at the lower level we have modern databases and modern movement of technology thanks to a dupe and a Big Data Apache ecosystem and that all sits on top of modern operating systems like iOS, Android, Windows, and it looks like I left Linux off of here, it should be on there. But what we haven’t done for application development is solve the middle layer which is the raw data content bound and organized according to the domain it needs to support. So those great programmers at the top are
taking advantage of those great languages. And they’re taking advantage of the technology at the bottom, but it’s still incredibly painful and unscalable for them to recreate the kind of data content and the organized logic around that data content that exists in healthcare. It’s just not scalable, there’s no way that we can build out a dozen or more health catalyst® organizations in the industry. It’s not going to work.

Dale: So DOS need number two. Another thing I’ve watched in the last 20 years in healthcare mergers and acquisitions, I always say that a new company is not integrated until the data is integrated. So you see executives jumping into mergers and acquisitions, and then within a few months realizing they can’t even pull together basic financial reports about the new company. Much less more complicated clinical quality measures that they’re at risk for being reimbursed. HIEs are not sufficient for this kind of data integration. Not even close. They barely support rudimentary clinical integrations, but you’re never going to balance a new company’s general ledger with an HIE. Ripping and replacing EHRs with a single common vendor is not an affordable strategy. I don’t— that’s stating the obvious isn’t it? There’s just no way that you can rip and replace EHRs as a strategy for interoperability. And besides, I would argue that hybrid vigor is a good thing in this context. You should not put all of your digital eggs in one basket. That’s the reality of it. So not only is it not affordable but it’s not a good thing long term to put all of your digital and data eggs in one vendor’s platform.

Dale: So, Rip and replace is not an answer for M&A. You can keep the disparate existing source systems. So finance, supply chain, registration, scheduling A/R, EHRs are just many— one of many source systems that you have to deal with in an M&A. But you can virtually integrate all of that with the data operating system. You can share transaction level data as you would with an HIE. You can integrate data for
common metrics around finance, clinical quality, utilization, all of those things and you don’t have to replace those source systems.

Dale: DOS need number three. We need to finally enable a personal health record. So I have— I think I have now, four personal health records. Depending on where I’ve lived and my care environment. So I still have a record from Intermountain, I still have a record from Northwestern, I have a record in the Cayman Islands, I have a record with my primary care physician here in Salt Lake City, who’s on Aetna health. And it’s up to me to try and figure out how to consolidate all that data into a concise personal health record that I can move around and share as I feel appropriate. That is not putting patients at the center if health care friends. We should all be embarrassed by that. With a data operating system we can lay a fabric over all those disparate systems that I just mentioned and we can pull that together and I would advocate that going forward that we think about pulling it into what amounts to the last remaining personal health record on the planet which is Microsoft’s HealthVault. But even if it’s not helpful we can provide a better personal health record than what we have right now. And by the way, you might think that me being a health care professional and a data guy that I would be willing to enter and manually consolidate all of that personal health data in one repository, there isn’t a chance. No way do I have time to do that. That content has to be seeded by the data that exists in all the different facilities and treatment areas that I’ve participated in over the last several years.

Dale: DOS need number four. Scaling and existing homegrown data warehouses. So what we’re seeing is interesting from a— from a very commercial selfish standpoint, Health Catalyst® doesn’t have an answer in the market for anything but ripping and replacing an existing data warehouse and we realize that that’s not possible— that’s not economically possible. So we’re motivated at a very granular level at Health Catalyst® to offer a better solution to that than ripping and replacing a data warehouse. But the
interesting thing is, independent of any Health Catalyst® selfish motives, homegrown data warehouses are easy to start and build but they are expensive to evolve and maintain. And there’s a lot of them in healthcare but they’re also very hard to retire. How do you retire it, right? Intermountain is still using the data warehouse that my team and I built and designed 20 years ago, and I haven’t talked to them very much about the details. But I suspect it’s time to do something pretty significant and different with that data warehouse. It’s 20 years old now but how do you retire that? You rip and replace it with another vendor solution? I don’t think so. That’s previously the only answer that we had at Health Catalyst®. But not having an answer to this is not good for the industry and it’s not good for Health Catalyst®. We need better options. Hence, the motivation to develop this data operating system with a fabric that can be laid over the top of any granular data.

Dale: So, again, selfishly we had to solve it. The industry, that will benefit too, that’s kind of the beauty of capitalism. And we can lay our Health Catalyst® and other applications over the DOS fabric and then that can be laid over the top of homegrown data warehouses. And in fact, there’s a side of me that would like to expand this market because the value to the industry, the value to Health Catalyst®, isn’t necessarily in the aggregation of granular data. That’s becoming quickly a commodity. The value is in the logic that resides on top of it and the models that reside on top of that granular data.

Dale: DOS need number five is the human healthcare data ecosystem. As a consequence of a project I’ve been involved in up in Alberta we conclude that only 8% of the data we need for that precision medicine and population health initiative resides in today’s EHRs. And, by the way, sorry for the carriage return there, we don’t have much data on healthy patients. We’re going to need way more data than what we have right now. So the ability to have a scalable platform for ingesting data and a scalable fabric on top
of that is only going to get more and more challenging if we don’t address this. We will never achieve precision medicine and population health without something like the data operating system.

Dale: Healthcare data as I mentioned earlier is becoming a bit of a commodity at the ingestion level. Thanks to open-source technology and—late binding schema-on-read designs, the data models, we can ingest data pretty cheaply and pretty quickly right now. But what’s not a commodity is understanding the data content, the data models, insanely complicated nuances in healthcare, that’s not going to be commoditized in their lifetime. The analytic logic or data bindings that apply to that data, how to organize it into meaningful chunks, the technology and skills to deliver to the right person at the right time, and right modality. The mundane thing of keeping up with changes in the source system data, what we call change data capture, enormously more complicated than what you appreciate if you’ve never worked in healthcare data before. Data quality management and scaling all this for a single healthcare system, that’s not going to become a commodity, ingesting data is.

Dale: So for the sake of dramatic impact, let me just kind of give you an idea of the data content and sources in the Health Catalyst® library to illustrate how impossible it will be for the industry to scale this kind of thing if we don’t—if we leave it up to each individual great idea and great application to do on their own. So these are all the different source systems of data that Health Catalyst® built up and I would remind everyone that this is just the beginning of healthcare data ecosystem, what I’m about to show you. This is just the beginning, okay?
Dale: Those are the EMR data sources that Health Catalyst® understands how to access the data, we understand the data models behind those. We understand how to bind that data together in meaningful ways, how to integrate it across EHRs and EMRs.

Dale: There’s the financial costing data sources. Again, all the same kind of knowledge required in the domain we’ve had to build up around the data models to data quality in those.

Dale: There’s all the billing data sources that we’ve had to build out, like that joke by the way. Now let’s run an expensive test to see if the cheap ones we ran were accurate.
Dale: There are all the HR and ERP data sources that we’ve connected to and again, we understand the data models, we understand how to bind and use that data across the continuum of care.

Dale: There’s all the claims data sources that we’ve built against.

Dale: There are all the clinical subspecialty data sources. Typically departmental kinds of sources of data.
Dale: There’s the health information exchanges that we pull data from.

Dale: Patient satisfaction data sources that we’ve integrated and understand the data models.

Dale: And then other sources of healthcare related data. Everything from census data in the state of Colorado to UAC data proven. All these other sources of data.
Dale: Then finally there’s master reference and terminology data. Like NPI registry, Rx Norm, HRQ, the HCCs that are coming out of CMS.

Dale: So all I have to say is true right? That’s all the data we have in the US healthcare ecosystem today, but we are barely getting started. So imagine what the future data ecosystem looks like. We have to create a more scalable way for ingesting that data, organizing it, and delivering it back to the point of decision making and we are just getting started. What we offer right now for the most part with traditional data warehousing and what’s emerging from the EHRs is not going to be scalable.

Dale: DOS need number six. Providers becoming payers, right? The insurance industry in my opinion is the tail wagging the healthcare dog. I can’t imagine much—many people other than those in the insurance industry seriously believing that the current payer insurance economic model is working. I believe very strongly it’s critical to the improvement of the situation for providers to model an assumed financial risk
and compete with or completely disintermediate insurance companies. With the data operating system providers have more and better data to model and manage risk than the insurers. And being a Sagittarius I was glad that the centaur turned out like you did on the left, not on the right, by the way. But that’s the hybrid that we have to have in the future. Providers have to become payers to change this situation that’s so unhealthy for the industry.

Dale: One of the interesting things here is that we can extend the life and value of current EHR investments with the data operating system. So if you look at a kind of a plot of expectations over time, we have pretty high expectations. I think some people had high expectations about what we could do with EHRs. Now the reality of EHRs is starting to settle in. I don’t think we’ve quite reached the trough of that yet because as we start optimizing those EHRs and we start trying to make them work in a different revenue cycle and reimbursement model and pop health, I think you’re going to see how difficult it is to do that. So, with open APIs and the data operating system we can actually reduce the depth of that trough and we can also increase and achieve the expectations that we all had for EHRs way back when. So if we can get the HR vendors to participate with us in the development of the DOS and open APIs, we can actually make their products better.

Dale: So as Dr. Robert Perl, from the CEO Permanente group, mentioned the other day; healthcare is using information technology from the last century. That’s a pretty big statement from an executive position leading 9,000 physicians and 34 staffers. It’s one of the more impressive healthcare systems in the world. Give that we’ve invested thirty billion dollars in tax money, what do we do now? So we replace all this outdated last century technology? No. There’s no way anybody has the appetite to do that. We can make what we have better while new products emerge. I guarantee you in the next few years new and better products are going to finally emerge in healthcare. It’s kind of interesting. The
appetite for change right now is pretty low, but in the next few years we will have revolutionary, better technology and software. You can see it coming.

Dale: The inevitable curve that occurs for technology is indicated in this drawing and what I’m trying to send out here is that the demand for EHRs was stretched by federal incentives. So demand, being on the y-axis, it was stretched artificially by federal incentives. That’s over. The underlying software and database technologies of EHRs has been commoditized a long time ago. So if you look at those two things, that the unnatural stretch by federal incentives and the impact that commoditizing technology can have on the evolution of our technical product, you would think that EHRs would be in trouble. The reality is of course no one has the appetite to make a change right now. But with the data operating system and open APIs we can change the trajectory so it looks more like the red dotted line now, right? Now, you don’t want to wait too late in extending and reinventing products. You need to start when you’re actually pretty comfortable at that maturity phase. That’s where Health Catalyst® is right now and that’s why we’re undergoing what we’re undergoing right now because we don’t want to wait to be disrupted by someone else. We’re going to disrupt ourselves. But we can improve this curve for the EHR vendors into the betterment of the industry.

Dale: So role model vendors like Facebook, all of those folks, their collaboration is role model for us in healthcare. Do we stack up in healthcare? Absolutely not. We’re terrible. The evidence is very clear and there are even some of the vendor app stores that appear to support open APIs like FHIR. But if you look at the details of the contracts and the way they’re worded, if you submit your applications into that app store that EHR vendor has the opportunity to take your IP and profit from it if you contribute. And my mantra here is we need to collaborate on standardization and then compete on innovation. We have to do exactly what the vendors in Silicon Valley are doing. Unbelievable what those bloodthirsty
competitors will do in the name of collaboration because they know that at the end of the day it’s about innovation, not standards, not mundane and boring standards. We need to follow the same kind of role model.

Dale: These are all the different open-source technology products that are available thanks to Silicon Valley. It’s moving so fast that this slide changes literally every week. It’s probably already outdated. This is just an example of what can be done when you focus on collaboration and then innovation around that collaboration and standardization.

Dale: Same thing applies at the software development tools layer on top of all that technology. We are at the beginning of software technology Renaissance. There’s no doubt about that. This is the most amazing time of my 33-year career in IT. We have to take advantage of this Renaissance in healthcare. We’ve got to put pressure on all of us to do better.

*“EHRs would become commodity components in a larger platform that would include other transactional systems and data warehouses running myriad apps, and apps could have access to diverse sources of shared data beyond a single health system’s records.”* 

Dale: Another quote, this came out just last Friday in the New England Journal of Medicine in an article called the 21st Century Health IT System. In that article the authors would suggest that with open standard software APIs, EHRs would become commodity components in a larger platform that would include other transactional systems and data warehouses running myriad apps and apps that could have access to diverse resources of shared data beyond a single health systems records. So that sounds threatening to an EHR but it’s only threatening if they don’t participate in what’s happening. That’s the only reason to feel threatened. They can leverage these concepts in these opening APIs to be more competitive and extend their product lifestyle and value.

Dale: So why can we do this technically more than we ever have before? Let’s dive into that.

Dale: So I have a pretty deep history of what a master open system standards evolution is and so the risk of jinxing myself I think I know the major patterns of successful failure is started way back in 1983 with my first exposure to ASN.1 and now it can be characterized by things like FHIR and JSON. So I think I have a pretty good understanding of what works and what doesn’t work in the progression of open systems and APIs. I’m convinced that we’re in this Renaissance period like we’ve never had before. I want to emphasize an experience at Northwestern. So when my team and I built the data warehouse there, Eric Just and Mike Doyle and Andrew Winter and everybody else there, we didn’t call it the data operating system but we had what amounted to an early version of it ten years ago. And so what I’m trying to communicate here is that the data operating system as a concept isn’t that far away. We just now have tools and techniques and more data content than we’ve ever had before so we can do it right and better this time like never before. In the data warehouse at Northwestern we supported analytics and near real-time exchanges, single records, before there were HIEs. We were pushing data point-to-point before there were HIEs around. We were using the data warehouse to do that. We had text data
and discrete data in a single platform. We ran analytics and batch processing computations on that data. We could also pull up just single records and display that in an application, single lab result served up through the data warehouse. This was all running on a pretty early version of a SQL server in Microsoft. Microsoft SQL server is way better than it was then. It has the ability to handle these mixed environments. And then you add the big data technology coming out of Silicon Valley and this concept is easily achievable friends. This data operating system is not just a pipe dream. We’re going to do this and it’s not that far away.

Dale: - calls it the hybrid Big Data SQL – tap or the hybrid transactional analytic processing. And I’ll just state this quote here, “Because traditional data warehouse practices will be outdated by the end of 2018, data warehouse solution architects must evolve toward a broader data management solution for analytics.” Totally agree with this. I think 2018 is about the right time. Again, which is why that we’re disrupting ourselves right now.

Dale: So thanks to this diagram from Ben Stopford at Confluent, but what we have now is the ability to pull data in from the left, take advantage of streaming pipelines, for example through tools like Kafka, we’re going to run analytics on it, run an elastic search on it, populate a SQL data warehouse or a no SQL data warehouse and then push this out through APIs to the source systems and transaction systems. So we have data and I’d rather we have technology around data that we’ve never had before.
Dale: I won’t go into too much detail here but land architecture is a conceptual architecture supported by the big data world. Basically says you can split your incoming data into two branches. One that’s a bad batch layer, one that’s a real-time layer, and then you can serve that up to end-users in the serving layer underlying all of this historical, as well as results, storage.

Dale: The Kappa architecture has some appeal. Again, it comes out of Big Data in that it uses one incoming data source. So just one code set for both real-time as well as what we’ve traditionally called batch oriented analytics. Both of these architectures can be implemented with a combination of open source tools like Apache and those listed there. Kafka, HBase, Hadoop. And the cool thing too that Microsoft and other vendors are doing is they’re blending these two environments together so that SQL and no SQL work seamlessly together. The products like Hive and that sort of thing, Impala.
Dale: So here’s another view. Just another reminder of the data operating system. Granular data, core services, the fabric, apps, all around open API is designed from the beginning to support open APIs and 3rd-party application development.

Dale: Another view of the fabric. And I might mention here too that, you know, there are few people that kind of go “Oh geez, this is too hard. This is going to be impossible.” Well my response to that is “No it’s not.” Because we basically did this at Northwestern ten years ago, and we didn’t have the tools that we have today to do that. And I would also say that just because it’s hard does that mean we shouldn’t do it? Right? Climbing Denali is hard, hiking the Appalachian Trail is time-consuming. They’re both difficult for different reasons. But it doesn’t mean they can’t or shouldn’t be done. So to anybody that wrings their hangs over this, or “Oh this is going to be too hard”, or “Oh my gosh how are we going to do this?”, I say “What’s the alternative? Not doing it?” What about all those reasons for doing that I mentioned earlier. What’s the answer to all those reasons I mentioned previously if we don’t do this? What’s the alternative? That we stay in the lower left quadrant of McKenzie’s digital index? I can’t imagine any of us as patients and citizens allowing that to happen. So this is as much about determination as it is to Health Catalyst® for me. I don’t want to be a part of the digital trajectory that we’re on anymore.

Dale: There are the seven attributes again, I won’t go over those. Those are just there for reference and refresher.
Dale: I will mention a little more detail here about the initial fabric services that we’re building to give you some sense of how we’re approaching this. We got Fabric dot Identity and dot authorization. Their names imply exactly what they would do. That machine learning is the microservice that allows data science- machine learning data science to be plugged into any application. Dot EHR is the micro-services that allow us to interact with EHRs with reading and writing and interacting at the functioning level. Dot PHR is the micro-service to support the personal health record. Dot terminology is that common terminology layer that will benefit application developers. They don’t have to recreate that. First the dot FHIR micro-service is going to be used across a lot of these areas. And then dot telemetry which is the ability to actually measure and see what’s going on in the data operating system from command centers and things like that. So those are some of the initial fabric services that are in development and available right now.

Dale: Here’s an example to give you some tangible- this is really happening kind of evidence. We’re converting our existing relational models into FHIR information models and so this is an example of one of the scripts that Imran and his team have developed for that conversion.
Dale: And then you’ll see the output into FHIR in this screen. So converting what we have in the relational world into FHIR is like I said, it’s difficult but it’s not impossible. You could argue it’s going to be time-consuming but we can accelerate that.

Dale: One other thing to think about is- here’s a sampling of the 220+ Health Catalyst® reusable value sets that we have. So we’re porting these into what we call a content management system, measures builder library, or MBL. So that we can reuse these in Health Catalyst® and 3rd-party applications. There are over 2,000 value sets and quality measures in the NQF library now and that’s just a portion of what we have to measure and keep track of in healthcare. So we’re building a content management system called MBL and from MBL then we’ll be able to express those value sets and reuse those value sets in the micro-services of the fabric. Okay?
Dale: Same kind of concept here with machine learning models in the data operating system. They’ll be in that fabric dot machine learning service that we talked about earlier so that any application can invoke these.

Dale: There’s a diagram of managing reusing the measures builder library and this explosion of value sets in the industry. So measures builder is the center of that diagram. The greater than 2,000 NQF and CMS values is coming from the left. Putting that in a form that app developers can reference both programmatically as well as manually so that they don’t have to hunt these things down all over the North 40. We’ll have clients and Health Catalyst® contributing where we can’t automate the value sets as we have on the left. By the way, we’re automatically pulling in those value sets that are produced by CMS and the National Library of Medicine. Thanks to our clinical analytics team we can now suck those into the measure builder library without human touch. Where we can’t bring those in automatically we’ll have content management systems that help us do that. Precise registry builder that we have will feed measures builder, then we’ll push that out to the Health Catalyst® Fabric and make it available to our applications, 3rd-party applications, and client Aps as well.

Dale: So one of the things we’re trying to do here in addition to build this data operating system and turn traditional data warehousing and HIEs on their head. We want to be role models in software development too. So, I would argue why can’t healthcare be the role model instead of Silicon Valley and should we aspire to something less than that? Is that acceptable to any of us? I would say no. Why can’t we be the role model for Silicon Valley instead of us always looking to them. Given the importance of our mission we should be the role model. So, these are the attributes of role model software development that we’re following to implement and achieve the concepts in the data operating system. I won’t go into the details of this one, you can review those at an offline time. And then how will we
know if we are a role model? I’ve listed below the software development vital signs we have. Successfully implementing the dots is critical. We want fast and simple releases every two weeks. Constant improvement of our apps, not these giant painful upgrades. We want analytics to drive the user interface in the application. We don’t want the application to drive the data, we want the data to drive the user interface. Amazon, you’ll never have two user interfaces the same when you have an Amazon experience because analytics is driving changes to those user interfaces. We want to constantly consume and expand the ecosystem of data that we offer up. Machine learning and pattern recognition will be amazingly valuable to all of us, will be a mark of our role model behavior. Economic scalability that we’re so efficient with our products and we can work across multiple operating systems, the data topologies. But it’s economically efficient for us to constantly deploy. We want what we call auto-fill analytics, so it’s a play on words. But how do we, through pattern recognition machine learning anticipate next steps in our clients decision-making. And then finally we’ll know that we’re a role model when Google, Facebook, Amazon, and Microsoft come to us for advice about software success and value. We’re going to do our best to become that role model.

Dale: Just trying to mention that for our Health Catalyst® clients we have what’s known as the Health Catalyst® community where you can learn more and watch the development of the data operating system. That’s out of community.healthcatalyst.com.

Dale: We’ll have questions about the DOS, request for features, reviews on roadmaps and things like that all available to you.
Dale: So, summary thoughts here. There will be people who hope we fail, guarantee that. Those who are trying to protect the status quo and are afraid of being disrupted will hope that we fail. There will be people who expect us to fail because it’s hard and healthcare IT doesn’t have a great reputation for breakthrough achievement. But I believe there are many more people who hope we don’t and that’s what we’re working for and as a patients and as members of the greater community of the United States and the world, this is something we need to do, it’s something that we can do, and we will do. And I think that’s it friends.

Dale: There’s a Healthcare Analytics Summit™ that we sponsor every year. It’s not a Health Catalyst® user group by any means. We bring people from all across industries and we bring people in from what would amount to our competitor client sites. What we’re trying to facilitate here is the exchange of good ideas and role model behavior. And increase the rate of innovation, diffusion, in the industry whether it comes from Health Catalyst® or not. We’re a strong believer in karma and this will come back to us. And now I think Tyler has passes to give away...