The Myth of health data complexity, why heath IT systems integrate poorly today, and what we can do about it

Why patient data isn't as accessible as it could be
in current EHRs

By Shahid N. Shah,
@healthcareguy (www.HealthcareITGuy.com)

Netspective | Extending the Enterprise
EXTENDING THE ENTERPRISE

The government's push for patient access to their data won't work unless the approach to health data integration is modernized.

Physicians, patients, healthcare providers, and other health industry participants have been clamoring for modernization of health IT systems for years. Recently, the HITECH Act, Meaningful Use, and other major government initiatives led by the Office of the National Coordinator (ONC) have been accelerating the demand. Unfortunately, as stated eloquently in the recent New England Journal of Medicine (NEJM) article "Escaping the EHR Trap - The Future of Health IT", health IT systems are trapped in legacy infrastructures:

It is a widely accepted myth that medicine requires complex, highly specialized information-technology (IT) systems. This myth continues to justify soaring IT costs, burdensome physician workloads, and stagnation in innovation — while doctors become increasingly bound to documentation and communication products that are functionally decades behind those they use in their "civilian" life.

The problem is not that engineers don't know how to create the right technology solutions or that we're facing a big governance problem. Rather, the real cross-industry issue is much bigger: our approach and the methods we have chosen for integration are opaque, decades old, and they reward closed systems. Drs. Mandl and Kohane summarize it well in their NEJM article by saying "a few companies controlling much of the market remain entrenched in "legacy" approaches, threatening other vendors' viability." They elaborated further on what they feel is the reason:

We believe that EHR vendors propagate the myth that health IT is qualitatively different from industrial and consumer products in order to protect their prices and market share and block new entrants. In reality, diverse functionality needn't reside within single EHR systems, and there's a clear path toward better, safer, cheaper, and nimbler tools for managing health care's complex tasks.

Netspective
EXTENDING THE ENTERPRISE

From the 1950s through the mid-1990s, systems integration required every system to know about each other in advance, agree on what data they would share, engage in governance meetings, put memoranda of understanding or contracts in place, and so on. In the age of the web, the approach has changed to one where the owner of the data provides whatever they decide (e.g., through a web server) and whoever wants it can come get it through a secure access method (e.g., through a browser or HTTP client). This kind of revolutionary approach in systems integration is what the health IT and medical device sectors are sorely lacking, and something that ONC, the U.S. Department of Health and Human Services (HHS) and the National Institute of Standards and Technology (NIST) can help promote. No amount of government money will solve health IT integration issues so long as our approach is incorrect.

As users of health IT systems, Drs. Mandle and Kohane have identified the problem of legacy approaches doing a lot of damage. What can we in the technology industry do to help? Let's take a look at the major issues holding back modernization of IT and integration of systems in healthcare, and what the government and systems owners — such as electronic health record (EHR) vendors — can do about it:

We don't support shared identities, single sign on (SSO), and industry-neutral authentication and authorization

Most health IT systems create their own custom logins and identities for users, storing metadata about roles, permissions, access controls, etc. in an opaque part of a proprietary database. Without identity sharing and exchange, there can be no easy and secure application integration capabilities, no matter how good the formats are. ONC should mandate that all future EHRs use industry-neutral and well supported identity management technologies so that each system has at least the ability to share identities. ONC does not need to do anything new -- they can can simply piggyback on the The

Netspective
EXTENDING THE ENTERPRISE

White House's National Strategy for Trusted Identities in Cyberspace (NSTIC) that is already defined and being managed by the National Institutes for Standards and Technology (NIST).

I'm continually surprised how little attention is paid to this cornerstone of application integration. There are very nice open identity exchange protocols, such as SAML, OpenID, and OAuth, as well as open roles and permissions-management protocols, such as XACML, that allow identity and permission sharing. Free open source tools such as OpenAM, Apache Directory, OpenLDAP, Shibboleth, and many commercial vendors have drop-in tools to make it almost trivial to do identity sharing, SSO, attribute-based access control (ABAC), and role-based access control (RBAC). It's quite hard to believe, but most current enterprise health IT systems don't even support Active Directory or LDAP.

We're too focused on "structured data integration" instead of "practical app integration" in our early project phases

In the early days of data collection and dissemination (it's sad to say that after 50 years of computing, health IT is still in those early days, but it's true) it's not important to share structured data at detailed machine-computable levels. Instead, different applications need immediate access to portions of data they don't already manage. When industries take on structured data integration too early, they often waste time because they don't understand the use cases well enough to specify best-case solutions. Poor implementations result.

For example, instead of asking for HL7 (the health IT vendors' evolved standard) or other structured data about patients, we can use simple techniques like HTML widgets to share "snippets" of our apps. Widgets are portions of apps that can be embedded or "mashed up" in other apps without tight coupling. The Department of Veterans Affairs' successful Blue Button approach has demonstrated the power of app integration versus structured data integration. It

provides immediate benefit to users while the data geeks figure out what they need for analytics, computations, etc.

Once app integration, SSO, identity sharing, and simple formats like JSON are in good shape, we can shift our focus to structured data integration, with all the governance and analytics associated with it. Future EHRs must master the production and consumption of secure authenticated application widgets using industry-standard approaches such as JavaScript and JSON.:

We focus more on "pushing" versus "pulling" data than is warranted early in projects

A question we commonly ask at the beginning of every integration project is "what data can you send me?" This is called the "push" model, where the system that contains the data is responsible for sending the data to all those that are interested (or to some central provider, such as a health information exchange). What future EHRs should do is implement syndicated Atom-like feeds (which could contain HL7 or other formats) for all the data they can share and allow anyone who wants it to subscribe to the data. This is called the "pull" model. Data holders allow secure authenticated subscriptions to their data and don't worry about direct coupling with other apps. If our future EHRs became completely decoupled, many of our integration problems would go away. Using ATOM and JSON as formats, the Open Data Protocol (oData), which has free open source implementations, should be used to actually open patient data in days rather than months.

To make sure security and privacy are maintained in the decoupled systems, automated auditing of protected health information can be enabled by logging data transfers through use of syslog and other reliable methods with proper access control rules expressed in standards like XACML.

Netspective
EXTENDING THE ENTERPRISE

We're too focused on heavyweight industry-specific formats instead of lightweight or micro formats

Appointment scheduling in the health IT ecosystem is a major source of health IT integration pain (in fact, much worse than most other areas). If EHRs just used industry-standard iCalendar/ICS publishing and subscribing we could solve, based on my experience, a large majority of appointment schedule integration problems. Think about how your iPad can sync with your Outlook/Exchange server at work — it's not magic, it's an industry-neutral standard widely used and supported.

Another example of outmoded industry practice is the use of HL7 ADTs for patient profile exchanges instead of more common and better supported SAML (which emerged to meet the need for industry-neutral user identities and profile exchange). If you've ever used your Google account/profile to log into another app on another website, you're using SAML. Again, no magic. It works millions of times a day with "good enough" security and user-controlled privacy.

Data emitted is not tagged using semantic markup, so it's not securable or shareable by default

In many existing contracts CIOs have signed, the vendors of systems that house the data also 'own' the data and it can't be easily liberated because the vendors of the systems actively prevent it from being shared. The healthcare industry sets up large data governance structures where vendors are cajoled and are often begged for access to patient data but vendors claim that it's not easy or not possible because health data is special. However, Drs. Mandl and Kohane, like me, think otherwise by clearly stating "some types of data used in health care are stored and used in ways that are unique to the medical field, but the field is not unusual in its need to share data across diverse electronic systems." Even when systems

are opened up after data governance establishes the sources and sinks of data along with specifications of data ownership rules, vendors only do the minimal tagging possible to make sure data can't easily be propagated securely. They do structured data integration and then present information on the screen (usually as HTML), failing to tag data with proper semantic markup when it's basically free to do (no extra development is required). One easy way to create semantically meaningful and easier to share and secure patient data is to have all HTML tags be generated with companion [RDFa](#) or [HTML5 Data Attributes](#) using industry-neutral schemas and microformats similar to the ones defined at [Schema.org](#). Using microformats and RDFa as a start, EHRs can then start tagging (in backward-compatible HTML) so that it's easier to discover metadata and allow simple securing and sharing of data. None of this is technically challenging if we really care about integration and are not just giving it lip service. Google's recent implementation of its [Knowledge Graph](#) is a great example of the utility of this semantic mapping approach. Once even basic microformats are in place with RDFa for authenticated or unauthenticated semantic tagging, we can then create SPARQL endpoints to make data easier to understand.

When health IT systems produce HTML, CSS, JavaScript, JSON, and other common outputs, it's not done in a security- and integration-friendly manner

Future EHRs should start to use industry-neutral CSS frameworks like Twitter's Bootstrap (which is free and open source). When using JavaScript, EHRs should use common lightweight and integration-friendly libraries like jQuery, instead of JavaScript frameworks that take over the app and prevent easy discovery and integration. Lastly, instead of emitting just complex XML or non-semantically aware HTML, emit JSON from your APIs so that client side applications can be easily written to take advantage of data.

Netspective
EXTENDING THE ENTERPRISE

Also, be sure to offer both JSON and JSONP so that integration can occur more easily without getting into security problems like cross-site scripting. Modern engineers that care about integration should always assume that their UIs might be "scraped" or connected to other systems and make it easy to allow others to securely take UI-focused data and create secure secondary uses.

Conclusion

All of these techniques I've mentioned are widely accepted secure web practices that need to make their way into our EHRs. Drs. Mandl and Kohane summed up the benefits of these approaches perfectly in their NEJM article:

Health IT vendors should adapt modern technologies wherever possible. Clinicians choosing products in order to participate in the Medicare and Medicaid EHR Incentive Programs should not be held hostage to EHRs that reduce their efficiency and strangle innovation. New companies will offer bundled, best-of-breed, interoperable, substitutable technologies — several of which are being developed with ONC funding — that can be optimized for use in health care improvement. Properly nurtured, these products will rapidly reach the market, effectively addressing the goals of "meaningful use," signaling the post-EHR era, and returning to the innovative spirit of EHR pioneers.

I'll go one step further and say that the government's multi-billion dollar incentives push won't do much if the technical methods and approaches being promoted don't match the commonly accepted, lightweight, and modern approaches mentioned above.