Why Centralized Data Banks Won’t Work for Health Information Exchange:  
Abstract

Rex E. Gantenbein, PhD  
Center for Rural Health Research and Education, University of Wyoming  
1000 E. University Avenue, Dept. 3432  
Laramie WY 82071  
Voice: 307.766.6545  Fax: 307.766.4356  
rex@uwyo.edu

There are a number of issues affecting the development of technology for the reliable, secure, and efficient exchange of health information among diverse electronic health records (EHRs). One of the most significant is the lack of interoperability among different EHRs that is needed to support the movement of information from one record format to another. While the certification efforts being carried out by the Certification Commission for Health Information Technology (CCHIT) are addressing this problem, there are still a number of EHR systems being (or already) deployed that are not compatible with others. Moreover, even CCHIT-certified systems require an interface that allows the information in one system to be translated into the format used by another. This translation is further complicated by different, non-standardized usages of terms for various health-related conditions.

The problems with centralized data banks

As a result, the use of a single, centralized data bank to store records for multiple facilities or organizations has been proposed. On the surface, this seems like a reasonable solution to the problems of not only health information exchange but also the costs and complexity involved in maintaining an EHR system. With a centralized data management scheme, all the records of the participating organizations are kept in one location, in the same format; there is no need for “exchange” as such, since all records would be accessible by each participant. Furthermore, small facilities would be freed of the need to manage their own systems, since the central bank would manage the servers, backups, etc. (presumably for a fee).

Unfortunately, there are a number of flaws in this scheme. A centralized data bank is expensive to maintain, particularly if there are a large number of users, and the data management mechanisms used to assure timely and accurate updates are complex. It also requires all participating organizations to use the same EHR, which may not be the most appropriate one for a given facility due to size, features (or lack thereof), or interface. All patients would have to have a unique identification code, which is technically feasible but difficult to manage.

Security and reliability are also concerns. A large data bank is a much more tempting target for hackers wanting to access personal information than a small, local set of electronic records. Moreover, because of the risk of private information being disclosed to unauthorized parties, many people are resistant to the idea of their records being maintained by a centralized system. A survey by a British newspaper showed that three-quarters of the respondents believed that a national system of EHRs would be less secure than the existing systems.

Even if these problems can be solved, there is still the issue of interoperability among centralized systems. Suppose that Wyoming was able to implement a single, statewide health information management system for its roughly 563,000 residents, many of whom travel to other states (mainly Colorado, Montana, and Utah) for health care that is not available locally. In order for their records to be shared with these states, the Wyoming system would need to be interfaced with all the systems in those other states (which might, or might not, be centralized or even certified as interoperable).

Federated storage: a better approach

In the long run, the idea of centralizing the storage for health information is fundamentally flawed. A better solution is the development of federated systems, in which local information is stored locally and accessed only when there is a need to exchange that information with another organization. The information is more likely to be up to date, since local updates will be done more quickly, and should one system fail, there will be no effect on the availability of information in other systems (unlike a centralized system, in which failure would make all information inaccessible until the system is restored). Moreover, a smaller system is likely to be less attractive to hackers.

A federated model would allow existing systems to continue to be used, as long as their information could be translated into a common, standardized format readable by other systems and convertible into their own format. In fact, such an interface would permit centralized systems to
communicate with each other, or even federated systems, in a hybrid “network of networks” that works similarly to current internetworks.

The federated model is not a particularly new idea. Distributed computing systems have been in operation for many years, and the federated model, which is based on such systems, has been the underlying architecture for many health information exchange designs, including the National Health Information Network (NHIN). The major issue with the model, however, is that national standards for digitizing and connecting patient data among individual systems are still in development. The Health Information Technology Standards Panel (HITSP) has identified a substantial set of standards for a nationwide record system, but their contract with the Department of Health and Human Services ended in 2010.

An lightweight infrastructure for health information exchange

Even with these highly specific standards in place, there is still a need for a robust infrastructure through which standardized information can be exchanged. One potential solution can be based on the idea of interface engines, which are often used to automatically convert information between differing formats; however, these systems are largely designed to convert information from legacy databases within an organization, not for exchanging such information across multiple organizational boundaries.

A more easily implemented (and significantly less costly) approach is the use of Extendible Markup Language (XML) constructs to provide a Web-based interface among different information systems. XML uses tags to identify particular objects within documents; such tagged documents can be transmitted across a network using the Simple Object Access Protocol (SOAP) in conjunction with the common HTTP protocol. SOAP encapsulates its transmitted messages, which supports maintaining the privacy of the messages when sent over public networks.

Web services provide the actual transportation among different information platforms using these technologies. The “producer” of a Web service provides and XML document that describes the interface to the service. This file would reside on a server that can be accessed by other network entities. Messages to the service – which can reside inside or outside a network firewall – can be constructed to request data or processes from the producer. These messages would come from another network entity (a “consumer”) that would request particular elements of patient data from the “producer” using the tags in the XML document.

Any entity participating in an exchange-enabled network using this scheme would need to provide two interfaces to the network. When a Web service receives a data request from another entity, it would process the request within its own system and construct an XML document that would be returned to the requester. The requesting application would consume the responses and transmit the included data to its own system for processing.

Achieving interoperability

One difficulty with this approach is that every entity in the network needs to be aware of every other entity, meaning that the interfaces would need to be reconfigured whenever a new entity joined the network, and new applications would need to be added to consume services from the new participant. This difficulty can be addressed, however, by a “hybrid” federated model that uses a centralized information management system (CIMS) that provides a standard interface for all network entities and is aware of all entities in the network and thus can route a request to any (or all) of them;

With this approach, no patient data is stored centrally, except during the transmission of messages. Furthermore, the network entities can individually determine what data will (or will not) be released in response to a request. The use of standardized interface (such as HL7) supports the interoperability among systems using different data formats.

Moreover, only the CIMS would need to be updated when new entities join or make requests of the network, and the participating entities may not even need to be aware of each other. Some system intelligence could also be built into the CIMS to make processing of requests more efficient. For example, a centralized patient registry containing basic patient identifying information (but no medical data) could be used to eliminate the need for a unique patient identifier known to all the network entities.

We have created a pilot information exchange system for biomedical researchers based on this approach at the University of Wyoming, and preliminary testing has shown it to be efficient in terms of locating and requesting information among different entities in a small (admittedly local) network. We are planning a broader test of the system over the next few months.