Guide to Immunization-related Electronic Data Exchange

June 13, 2008 (rev)

by
Noam H. Arzt, Ph.D.
HLN Consulting, LLC®
8449 Christopher Ridge Terrace
San Diego, CA 92127
858/538-2220 (Voice)
858/538-2209 (FAX)
info@hln.com
http://www.hln.com/

Table of Content

1 Start at the Beginning

1.1 Introduction

Immunization Information Systems (IIS) are somewhat unique public health systems in that by their very definition they are concerned with the effective collection, storage, and retrieval of consolidated immunization records, initially for children but increasingly across the lifespan. Most public health systems are inwardly-focused, and primarily support internal agency operations. IIS equally (if not predominantly) supports the clinician out in the field – public and private – while also supporting public health programs, and traditional public health monitoring and assurance. These external organizations and individuals are users of IIS applications and data, and represent the source of much of the data that gets stored within the IIS. Yet these users are under severe pressure to work as efficiently as possible as the "business" of primary care medicine is more marginally reimbursed than it once was. Any additional administrative activities not perceived of as essential to good medical practice *or* reduction in cost are appropriately resisted by physicians, practice managers, and their staffs.

There is a growing phenomenon in the United States towards deployment of electronic data systems of many kinds to improve effectiveness and efficiency. The near ubiquity of the Internet has fueled even stronger desires to promote electronic information over paper-based records and has enabled system-to-system sharing of data more easily than ever before. This document will focus on both the "why" and, at least at a high-level, the "how" of enabling electronic data exchange between IIS and other systems. The target audiences include public health and private sector functional and information technology managers who want to gain an understanding as to the issues, opportunities, and obstacles to electronic data exchange. Individuals charged with actual implementation of electronic data exchange interfaces will be able to refer to more detailed documents in the bibliography at the end of the document.

1.2 Why Exchange Data Electronically with an IIS?

There are a number of good reasons why enabling an IIS to exchange data electronically is desirable:

1. **Reduce the burden of on-line data entry:** In many jurisdictions, reporting of immunization events to an authorized public health agency is required by law; in other jurisdictions it may simply be accepted practice for the benefit of the entire community. As more and more reporting organizations (practices, hospitals, and clinics) deploy electronic health record systems (EHR-S), more and more immunization data will already exist in electronic form even before it is ready to be sent to an IIS. Yet in many jurisdictions, users are forced to key-enter immunization data twice (or more): once in a local system, and then again in the IIS. This duplicative activity can be eliminated through proper implementation of electronic data exchange.

2. **Pressure from some users to *only* use institutional applications:** Clinical practice is complicated, and supporting clinical applications within an organization is equally complicated. As the abundance of computer applications can proliferate over time, many organizations (especially larger organizations) attempt to reign in the cost of supporting a multitude of applications by streamlining the number and type of applications that are available to users. Often, external applications – like an IIS – are perceived of as being more difficult to support

and as being distracting for users. IIS typically look different, behave differently, and require different usernames and passwords than institutional systems. For this reason, an increasing number of institutions are *discouraging* their users from logging in to IIS while at the same time looking for alternative ways to provide the data and special features that IIS offers within their local systems. Electronic data exchange represents one key strategy in fulfilling the data needs of this growing set of organizations.

3. **Leverage existing and emerging data exchange capabilities:** Whether IIS exists or not, healthcare is moving towards greater exchange of data between systems. In many cases, not only does IIS *not* have to "reinvent the wheel" but IIS can often ride along on existing data exchange relationships and installations. In public health, disease surveillance and syndromic surveillance activities often provide the foundation for basic data exchange between clinical organizations and the agency. Electronic laboratory results, ePrescribing, and telemedicine provide additional opportunities in the private sector. Over time, more opportunities will come about for public health to leverage these existing and emerging activities.

4. **It's one of the Minimum Functional Standards for IIS:** Exchange of immunization records via the HL7 protocol is Minimum Functional Standard #7 for IIS as defined by the Centers for Disease Control and Prevention (CDC).[1] Though there is no certification process as of yet for IIS, many projects are funded through by the CDC under Section 317 of the Public Health Service Act. Compliance with CDC minimum functional standards is certainly strongly advised, if not required, to receive these public funds.

1.3 With Whom is Immunization Data Exchanged?

While a full model of IIS data exchange will be offered in the next section, here is a quick summary of the major IIS data exchange partner types:

- **Providers and provider organizations:** Providers want easy and quick ways to report patient and immunization data to IIS, whether required to by law or not. In addition, as we have begun discussing, providers are beginning to want IIS data *back* so that their own local systems can benefit from what IIS knows about their patients.

- **Health plans:** Immunizations represents a core set of data used to measure the performance of clinicians. In fact, the National Committee for Quality Assurance (NCQA) has long included immunization rates in their Healthcare Effectiveness Data and Information Set (HEDIS) which is used by most health plans to measure their performance.[2] Rather than performing expensive and time-consuming manual records audits, health plans would usually rather receive data on their enrollees from IIS electronically in a HIPAA compliant manner.

- **Vital Records:** Many IIS initialize their newborns with records received from State or local vital records systems. This is an important source of data for IIS and almost always originates in electronic form. Prompt initialization of newborn records in the IIS helps to reduce missed

[1] See http://www.cdc.gov/vaccines/programs/iis/stds/min-funct-std-2001.htm
[2] See http://www.ncqa.org/tabid/59/Default.aspx

opportunities for immunization (and often other newborn screening) early on and establishes a needed base for population health monitoring.

- **Public Health:** Numerous public health programs, and the systems that support them, rely on immunization data from IIS across a wide variety of activities. Emergency preparedness and response programs, including Countermeasure and Response Administration (CRA) and Strategic National Stockpile (SNS) activities as well as Outbreak Control and Management and Communicable Disease Surveillance, use and create immunization data.[3] Immunization program activities, including administration of the Vaccines for Children (VFC) Program[4] and the Assessment, Feedback, Incentives, and Exchange (AFIX) quality improvement strategy[5] as well as influenza and pneumo campaigns for adults also rely heavily on IIS data for their implementation. State-sponsored chronic disease management programs have begun deploying systems to help clinicians with disease management that invariably have immunizations as a key component. Finally, epidemiologists and other public health professionals have long used immunization data – increasingly from IIS – for studies and analyses.

- **Other agencies and programs:** Many jurisdictions are beginning to integrate systems from different agencies or programs that offer services to overlapping groups of citizens. IIS often possesses a fairly complete, population-based set of person data which can be useful for many programs and systems. For example, a maternal and child health program might be able to identify from the IIS children who are missing immunizations and target those children for outreach not only for immunization but other likely health deficiencies (for example, exposure to lead). School and WIC programs may be able to identify children in need of immunizations and target outreach to those families.

- **Regional Health Information Organizations (RHIO):** RHIOs are broad-based organizations that promote an enable the exchange of health information between stakeholders in a particular location. RHIOs are building, from the bottom up, a Nationwide Health Information Network (NHIN) that hopes one day to be the roadway upon which health information moves swiftly and securely between all parties with a "need to know." IIS is emerging as a key player in some early RHIO projects, and can benefit from the data exchange initiatives that RHIOs initiate.

1.4 IIS Interoperability Model[6]

We have already reviewed the key players in IIS electronic data exchange. The following diagram, explained section by section below, shows these interactions by placing the IIS at the "center of the world" and displaying related systems around it:

[3] See http://www.cdc.gov/phin/activities/applications-services/cra/index.html
[4] See http://www.cdc.gov/vaccines/programs/vfc/default.htm
[5] See http://www.cdc.gov/vaccines/programs/afix/default.htm
[6] Developed in part with funding from the Vermont Department of Health.

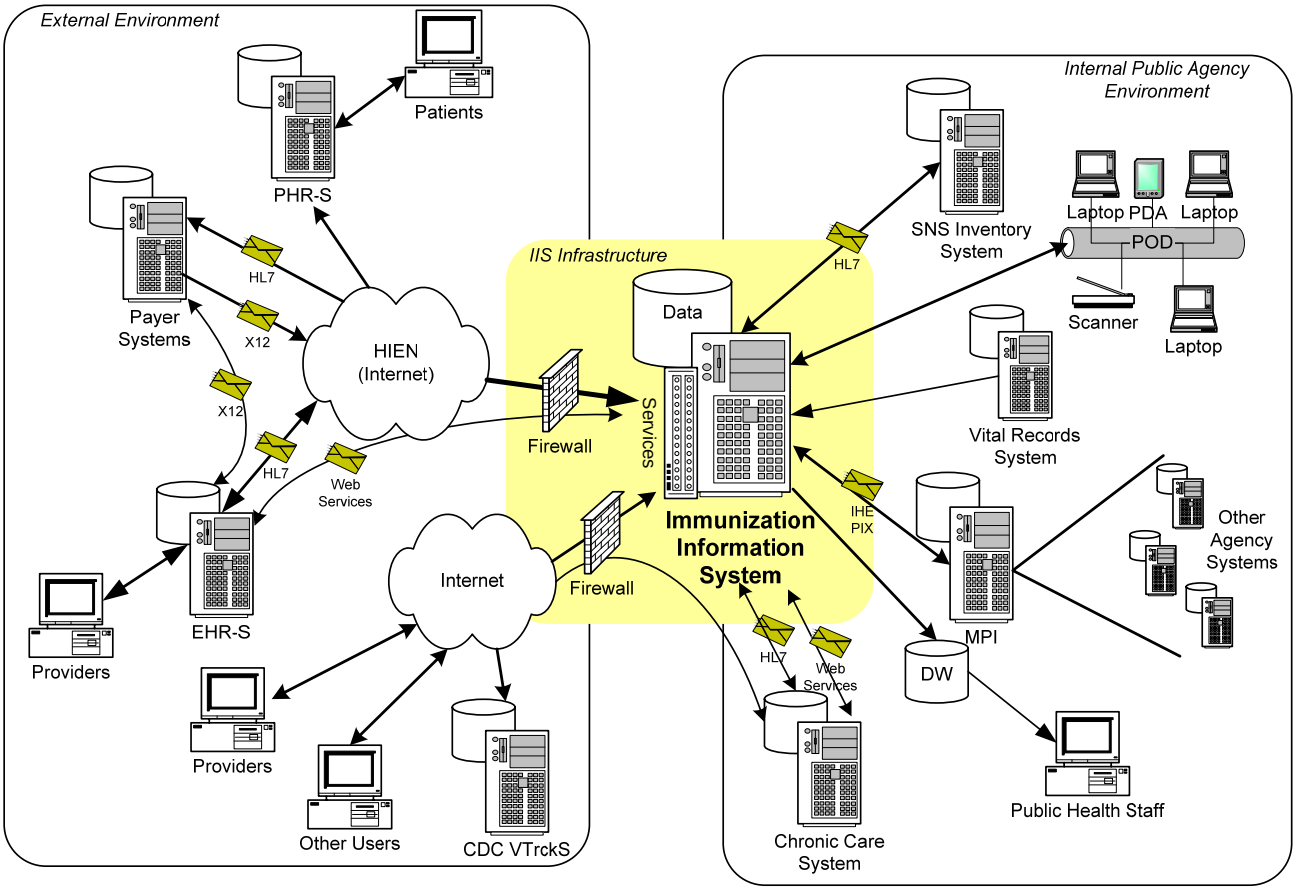
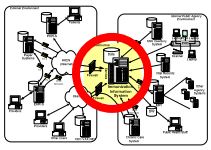
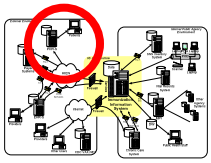


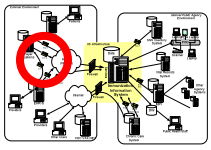
Figure 1 – IIS Interoperability Model



IIS Infrastructure: The IIS provides data and services to various healthcare activities, both in the public and private sectors. It is usually a centralized system that contains a database and one or more applications for online access. A set of services may provide more transparent access to IIS data or features by outside systems. All access to IIS data is governed by strict privacy and security rules dictated by both Federal and state/local legislation.

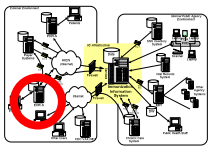


Personal Health Record System: Patients and their guardians may access immunization data through a Personal Health Record System (PHR-S). Those systems provided by an organization with which the patient has an affiliation (health plan or insurer, employer) are referred to as "tethered" systems; those provided by an independent entity are referred to as "un-tethered." A PHR-S may acquire its data directly from the patient, from provider EHR systems (see below) or even from an IIS. It is important for IIS managers to recognize the potentially growing important these systems may have in the system landscape.

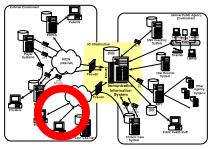


Payer Systems: Payers have always played a big part in IIS activities. They are anxious to collect data from IIS for HEDIS measures or to support other quality improvement activities. They can be a source of primarily administrative data for an IIS when their records are considered accurate and complete enough. Often this is not the case as some claims data received from providers may not have all the required data for an immunization event (e.g., sometimes the type of vaccine administered is not included in the claims data).

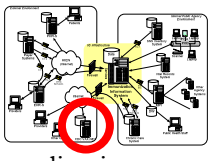
Payer systems receive claims in X12 format but need to be prepared to also receive immunization data eventually in HL7 format from IIS and other sources. For the near term, many of these interactions will take place through the use of flat files whose format and contents are defined for specific purposes.



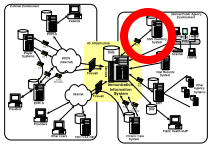
Electronic Health Record System: Providers use an Electronic Health Record System (EHR-S) to automate their clinical practice and serve as an electronic version of their patients' records. An EHR-S exchanges data with the IIS via HL7 and, if so equipped, can access its services through a published Web Services[7] interface. These services might include an immunization forecast scheduler whose algorithm can be applied to EHR-S data independent of the IIS, or tools such as practice assessment or reminder/recall capabilities. A typical EHR-S system does not contain these special features and may welcome the ability to access them from the IIS via Web Services. The EHR-S may also provide insurance claims directly to a payer system via X12. Providers access an EHR-S in their offices through many types of interfaces, including web-based application, client/server applications, and even Personal Digital Assistants (PDA).



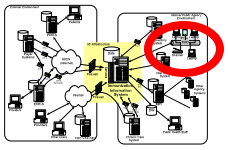
Direct Online IIS Access: Providers who do not have an EHR-S, as well as other users from schools, pharmacies, visiting nurse associations, and others access an IIS directly using an online application (usually web-based or web-presented) provided by the IIS project. For commercially- or public health-acquired IIS software these applications are provided with the system; for others these products are developed in-house. Over time, it might be expected that use of these stand-alone applications will diminish as more and more stakeholders acquire primary systems of their own that can interface with the IIS to provide access to its data.



CDC VTrckS: CDC is currently developing a new system to replace VACMAN and automate the ordering and distribution of Federally-provided vaccines. Many projects will choose to use the IIS as the primary interface for its providers to this new ordering system. In the short run, this may increase the use of direct online IIS applications; over time it is expected that service-based interfaces to these ordering functions will let primary systems interoperate with the IIS to support the necessary processes.



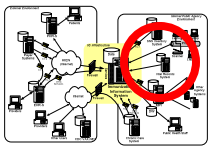
SNS Inventory System: Many jurisdictions have deployed an inventory system to support their potential use of the Strategic National Stockpile in the event of a large-scale emergency. Vaccine is just one type of materiel that might be involved in a public health response. The IIS may be called upon either during or after an event to store vaccine administration data. It may require an interface to the SNS Inventory System to support this activity.



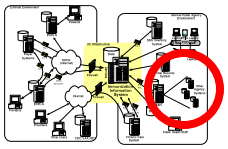
Points of Dispensing Support: During an emergency response, jurisdictions use Points of Dispensing (POD) as the locations where vaccines and medications may be distributed to the population. These PODs are often temporary locations –schools, recreation centers, churches, or mobile locations – pressed into service during an emergency with little or no permanent public health infrastructure. PODs may support data management products of their own or may access the IIS directly through standard or special

[7] Web services refers to a set of technologies that implement a Services-oriented architecture (SOA) which features software components that are assembled in a modular way to facilitate reuse and standard interfaces. See http://www.webservices.org/

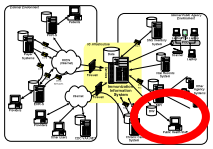
applications. In some cases, stand-alone applications may collect data that later needs to be integrated into the IIS database after-the-fact. Data may also be collected through alternate means such as scanning of paper forms or via a PDA interface to account for high volume or difficult conditions.



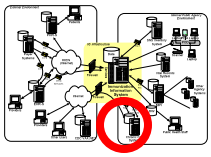
Vital Records: Children's records in the IIS are often initialized from Vital Records data within the jurisdiction which may or may not be housed within the same agency. This data transfer is usually conducted under a Memorandum of Agreement or similar instrument that dictates how and when this data can be used. CDC Minimum Functional Standards stipulate that this transfer happen within six weeks of birth, but most IIS strive to have the data transfer occur more expeditiously. IIS projects need to be careful to abide by state law related to disclosure of birth records under conditions of adoption or foster care, and also need to be diligent in updated official name changes and in recording deaths to prevent accidental transmission of reminder/recall information for a deceased child.



Other Agency Systems: Most agencies have additional internal systems that manage other aspects of their programs to which IIS either interoperates or is integrated. At minimum, there is usual some exchange of data (in one direction or both) with Vital Records systems (for electronic birth and death records), WIC systems (for providing immunization history for assessment by WIC personnel), and service-encounter systems (that support automation of services offered in public health clinics) should they exist. More sophisticated implementations may use a Master Patient Index (MPI) as a central registration point and clearinghouse for patient-related data in the agency.



Public Health Staff: IIS data supports Immunization Program functions including surveillance, coverage assessment, quality assurance at provider sites (AFIX), vaccine management and storage, and communication with stakeholders. The more widely available data is within the agency the more important this data management function. Some agencies deploy a data warehouse parallel to the production system to facilitate data query and reporting with less impact on operations.



Chronic Care System: Many jurisdictions are deploying specific systems targeted at management of chronic conditions within the population. Typically focused on clinical support, these systems provide features that an EHR-S does not yet have. Chronic Care System exchanges immunization information and accesses vaccine forecast from IIS. Some users access through their EHR-S, some directly through a "lite" client. Chronic Care Systems should also be able to access IIS functions via Web Services for a more transparent integration of these capabilities into the system.

1.5 Health Information Exchange Networks

A regional health information organization (RHIO) is a collaborative organization focused on health information exchange. RHIOs operate health information exchange networks (HIENs) which are technical implementations supporting health information exchange between RHIO participants including physicians, laboratories, hospitals, pharmacies, patients, public health, and payers. A RHIO is primarily driven by the private sector, but often has public health and state government involvement. Usually RHIOs are focused on clinical data exchange, but also may focus on health services data and providing clinical applications. They can span a metropolitan area, a county, a state or a multi-state

region. Interoperating RHIOs and HIENs across the country will form the nationwide health information network (NHIN).

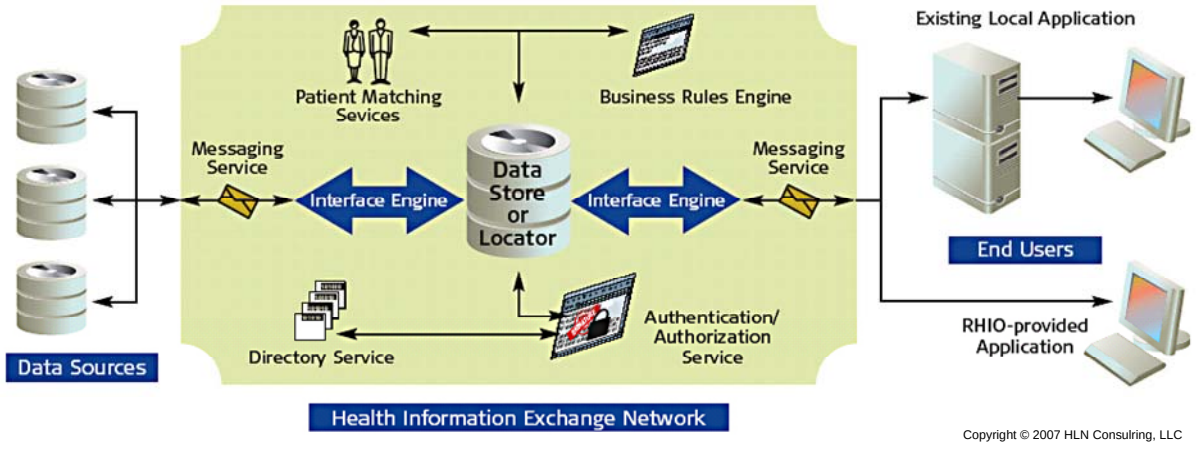


Figure 2 – Health Information Exchange Component Architecture

Figure 2 displays a typical HIEN architecture composed of the following major components:

Data Sources: Systems that provide and/or receive data from the HIEN. Examples include provider-based electronic medical records, practice management systems, immunization and disease registries, surveillance systems, laboratory information management systems, imaging systems and pharmacy information systems.

Data Store: Central or distributed databases containing patient identification and health information. Data and document locator services are examples.

Interface Engine: A utility that provides translation of standard messages containing patient identification and health information being sent to or from the data store.

Messaging Service: A system that provides for the secure transport of information in and out of the HIEN.

Matching Service: A utility that provides reliable matching and linking of patient information received by the HIEN from disparate sources. A master patient index, record locator service, and patient locator service might be major components of this service.

Business Rules Engine: A repository for the business rules that control how the HIEN treats data and the business processes it supports. These rules are available consistently to all HIEN applications.

Authentication/Authorization Service: A utility that provides reliable identity verification of system users (authentication) and ensures that users access information and services that are appropriate (authorization).

Directory Service: A central repository for information about system users. In Vermont, this might incorporate the emerging master provider index.

End User Applications: Existing or new applications used to access HIEN patient and health information. Examples include provider and patient portals and electronic health records systems.

As described in the IIS Interoperability Model, IIS will likely need to interact with HIENs sooner or later to either supply or receive data.

2 It's All About Interoperability

2.1 Introduction

Some agencies have recognized the limitations of deploying systems purely within individual programs when the information related and their limited funds for technology could be better spent if leveraged across multiple projects. These agencies developed a broader vision and some of their systems evolved into *integrated systems* supporting a wider variety of patient-centered or case-centered functions. By bringing together information from many programs, agencies hope to be able to better serve their citizens by reducing the oppressive burden of forms and paperwork faced by many people that became barriers to receiving service in the first place, and increase the coordination across agency programs to ensure that only eligible people were receiving the appropriate services. These efforts paralleled the activities in hospitals to implement unified patient intake and scheduling systems as a precursor to computerized patient order entry (CPOE) systems for inpatient facilities or comprehensive electronic health record systems for both inpatient and ambulatory settings. While many IIS projects have integrated their systems, the challenges they have faced have caused some to refocus the issue: in addition to concentrating on healthcare systems integration *within* the agency, these organizations are focusing on healthcare systems *interoperability* with organization *outside* of the agency.

But just what is interoperability? In early 2007 the Health Level 7 (HL7) Electronic Health Record (EHR) Interoperability Work Group published a white paper[8] whose purpose was to consider the meaning of interoperability, develop a consensus definition, and discuss implications for future standards work. After reviewing and analyzing more than 100 definitions, the work group agreed upon a three-part definition:

Interoperability as Defined by HL7:

Technical Interoperability: Structure, syntax, reliable communication

Semantic Interoperability: Full meaning preserved

Process Interoperability: Integration of systems into work flow

1. **Technical interoperability** focuses on the physical transmission and receipt of health data, its transport between participating systems. Much of the work here is on message formats and reliable, secure message transport.

2. **Semantic interoperability** focuses on ensuring shared meaning between sending and receiving partners – ensuring that the meaning of what was sent is consistent with the understanding of what was received. Much of the work in this area is focused on medical terminology which can be referenced consistently by all parties.

3. **Process interoperability** focuses on higher-order workflow concepts that make data sharing a richer and more valuable experience. Work in this area tries to understand how shared health data supports the specific activities and workflow of the organizations that use it and the integration of health data into the work setting. Issues of data usability and timeliness are examples of process interoperability concerns.

[8] Health Level 7 EHR Interoperability Work Group, *Coming to Terms: Scoping Interoperability for Health Care,* February 2007. < http://www.hln.com/assets/pdf/Coming-to-Terms-February-2007.pdf>

The HL7 EHR Interoperability Work Group went on to define an Interoperability Model Draft Standard for Trial Use[9] (DSTU) which defines the characteristics that records need to have to meet these three levels of interoperability.

Let's now review these three components of interoperability one at a time with some examples, since they form the foundation of proper electronic data exchange.

2.2 Technical Interoperability

Technical Interoperability between systems is achieved primarily through *messaging*. Several message standards are available and relevant, including Health Level 7 (HL7). How this messaging typically works is illustrated in Figure 3.

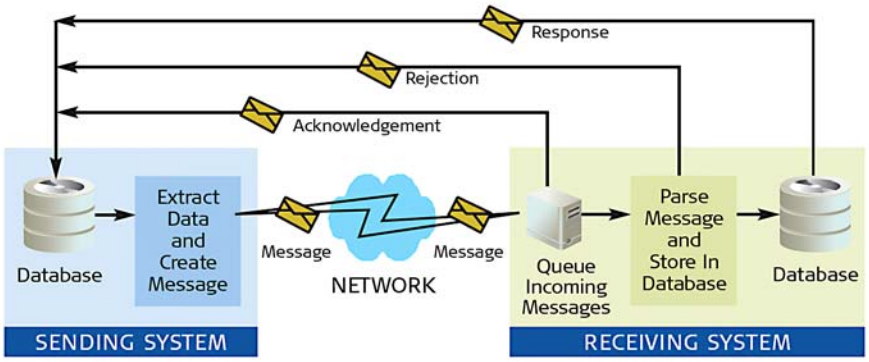


Figure 3 – System-to-system Messaging

The sending system extracts data from its database and creates a standard message. This message includes a message header, or envelope (information about who is sending it and where it is intended to go), and a message body (some kind of data payload to be carried). The message body might contain data bound for the receiving system, or it might contain a structured request for data from the receiving system. The message is then transported across a network (often using a secure and encrypted channel over the Internet) to a message queue in the receiving system to wait its turn for processing (this is in the integration engine described previously).

When the message is processed, it is read and interpreted by a message parser in the receiving system which checks the message header and evaluates the contents in the body. Based on the business rules in the receiving system, the contents are stored in the receiving system's database, held for further processing, or rejected. An acknowledgement message might be sent back to the sending system, or a rejection message might be sent instead. If data was requested and found, a response message would be sent back.

Public health agencies in general, and IIS projects in particular, have been sending and receiving data via messaging for years. In many cases the formats, transport protocols, and message contents have been proprietary, and the interactions have not been fully automated. But the basic concepts are the same and much of an IIS prior experience in exchanging "flat files" is still relevant.

[9] See http://www.hl7.org/ehr/downloads/index_2007.asp

2.3 Semantic Interoperability

Data must be understood correctly: what the sender intended should be recognized unambiguously by the receiver. Just as two people who are not speaking the same language cannot be understood, even if their two languages use the same letters to make up their words and some of the words have meaning in each language, data cannot be exchanged with another system unless you can ensure that the "language" of the data is consistent. In practical terms, ensuring semantic interoperability often comes down to agreement on the naming and use of standard data elements, and agreement on the use of a standard set of codes and terminologies to instantiate data within those elements.

The American Immunization Registry Association (AIRA) has done a lot of work in the area of semantic interoperability for IIS. AIRA's *Data Codebook* provides a useful resource to IIS projects by identifying commonly-used code set relevant to IIS deployment.[10] Figure 4 shows an excerpt from that codebook which is delivered as a Microsoft Excel workbook:

Category	Table No.	Subject	Source - who is responsible for maintaining code set	Codes - values	Codes - Description	Suggestions on coding?	Description	Rationale	Related To	Required in which code sets (also see Tab 2)	Active - Is this field currently used
Client	0001	Sex	User	F	Female	N	This is the Patient's Sex.	In an immunization event, this is used to identify the sex of patient.			
	0001	Sex	User	M	Male						
	0001	Sex	User	O	Other						
	0001	Sex	User	U	Unknown						
HL7	0003	Event Type	HL7	A28	ADT/ACK-Add person information	N	Provides a finer level of definition in the HL7 message type. For ADT messages, it indicates whether the message adds, deletes, updates a person's record. For the immunization messages, its value is always determined by the HL7 Message type.	For some HL7 messages it is necessary to indicate what action is to be taken as the result of receiving the message (e.g. "delete the patient record"). For other messages, while required, it provides no additional information.		HL7 0003	
	0003	Event Type	HL7	A29	ADT/ACK-Delete person information						
	0003	Event Type	HL7	A30	ADT/ACK-Merge person information						

Figure 4 – Selection from AIRA Data Codebook

This codebook was developed through a consensus process within the IIS community and represents an invaluable resource for addressing semantic interoperability issues. It continues to be maintained by AIRA's Data Definitions Workgroup.

2.4 Process Interoperability

It's not enough for systems to be able to exchange data reliably, and for the meaning to be intact. For systems to be truly interoperable they need to be able to support the business processes important to the users with that data as naturally and effectively as possible. One way to ensure that IIS and the data they possess supports users properly is through the development of Use Cases which detail, in a structured way, the functions to be supported by a system or set of systems from the users' point of view. Use case documents often become the point of intersection between the user's world and the technology world. In the fall of 2007, the Office of the National Coordinator for Health Information Technology (ONC)

[10] See http://www.immregistries.org/docs/IIS_Data_Codebook_101607.xls

of the Federal Department of Health and Human Service (HHS) began developing a use case to describe how systems and data should support the process of administering vaccines and medicines to patients during the course of both routine care and emergency response. A final use case was published in March 2008 and handed over to the Health Information Technology Standards Panel (HITSP) Population Health Technical Committee whose responsibility it is to develop interoperability specifications to support the use case.[11] The following diagram from this use case illustrates the flow of information to support vaccine and drug administration and reporting:

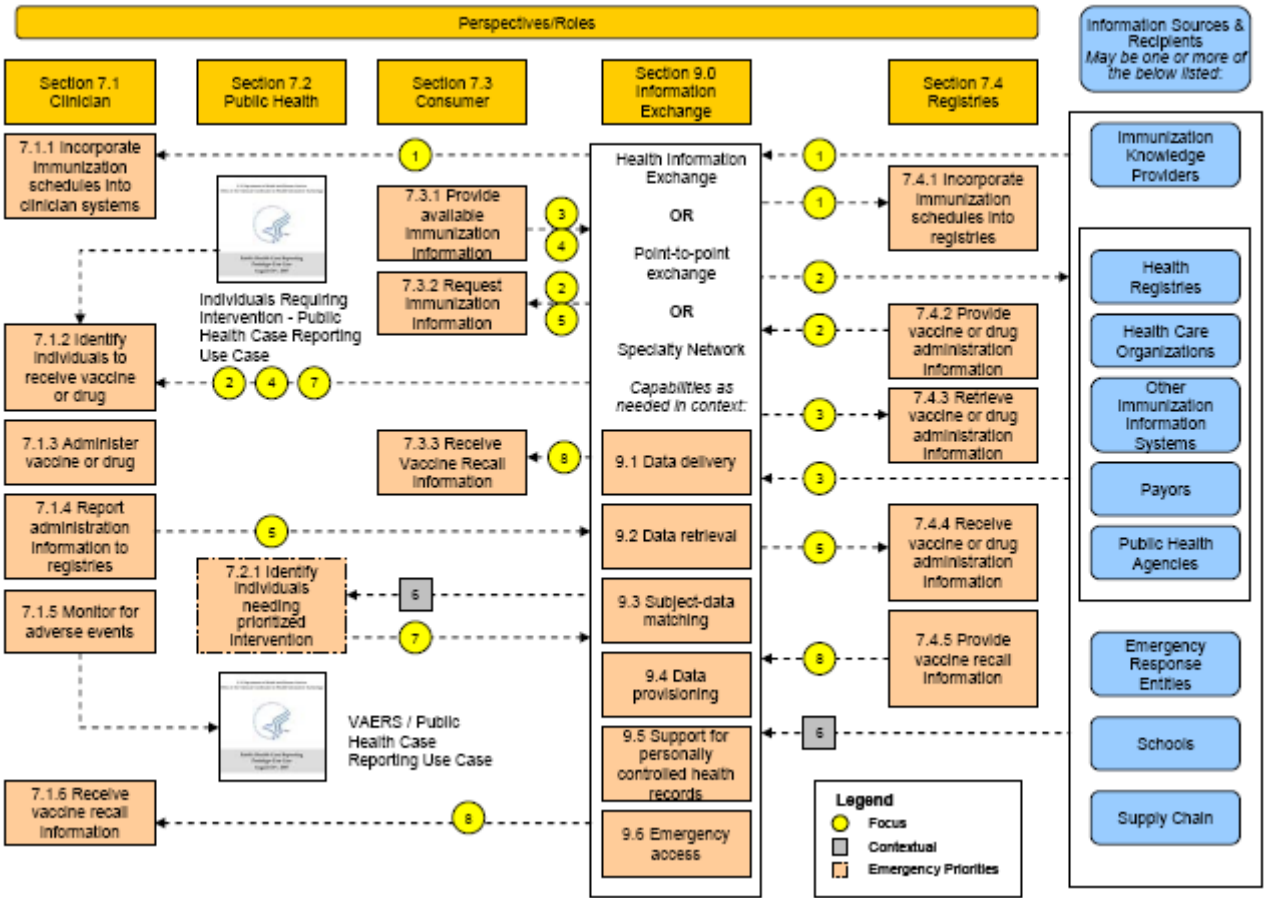


Figure 5 – Vaccine and Drug Administration and Reporting

[11] See http://www.hhs.gov/healthit/usecases/respmgmt.html

3 Practical Approaches and Best Practices

3.1 Introduction

In the following section we will offer a set of practical approaches and best practices – tried and true techniques for ensuring high-quality interoperability. No single solution is a panacea, and not all suggestions will be a perfect fit in all circumstances. In many cases, further details and examples are referenced in additional documents.

Topics to be covered include:

Technical/Semantic Interoperability:

- Overall architecture and information flow
- Data-centered or Document-centered
- Message/file format
- Record contents
- Direction of data transfer
- Person matching
- Data quality assurance
- Data transport

Process Interoperability

- Record modification and deletion
- Batch mode versus real-time mode
- Process monitoring and error correction

3.2 Overall Architecture and Information Flow

Though we begin with a discussion of overall architecture, it is likely that this cannot be determined with certainty until all the issues in this guide are examined and considered. Yet it does help to begin with a clear picture of the whole before considering its specific parts:

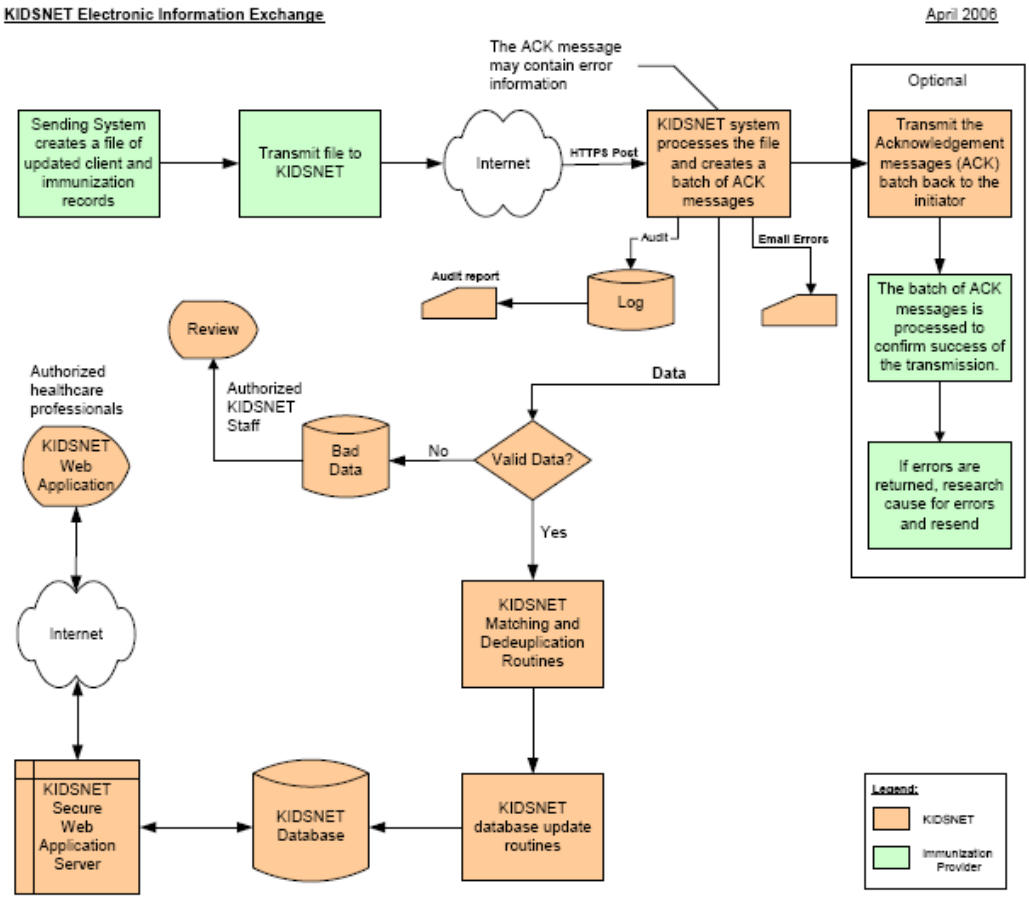


Figure 6 – KIDSNET Data Flow Diagram

The diagram above from the State of Rhode Island's KIDSNET integrated child health information system captures pictorially how data moves between key systems that need to interoperate in that project's environment, namely provider EHR systems and the KIDSNET IIS. The diagram highlights not only the basic processing steps and exceptions, but also includes the points of quality review. A more detailed description of these steps is found in the document from which this was borrowed, the *RI KIDSNET HL7 Implementation Guide for Immunization Transactions*[12] which is provided to data partners so that they can understand how the data exchange is supposed to work and what they need to do to comply with the standards that ensure that it works smoothly and reliably.

If you can, think in pictures! But recognize that some people understand better by reading text. Both are necessary to convey the overall intent of your project and the details.

[12] See http://www.health.ri.gov/family/kidsnet/dataexchange.php

3.3 Data-centered or Document-centered?

There are two predominant styles of data *storage* for public health data repositories. In a *data-centered* approach, systems store data in a conventional relational database (RDBMS) with tables for different entities and rows for instances of those entities. Structured Query Language (SQL) is used to send queries to the database and extract the rows from the appropriate tables that meet specified criteria (Figure 7).

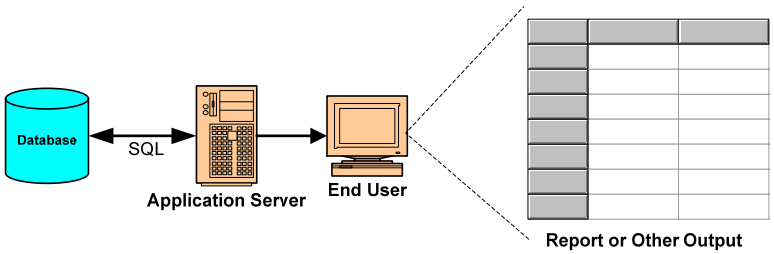


Figure 7 – Data-centered Storage Approach

In a *document-centered* approach, data is stored in a formatted document within the repository for retrieval as a unit. The contents of the document cannot usually be searched directly; a smaller amount of metadata is saved in a database at the time the document is stored or updated which contains key information useful in selecting the document for retrieval (Figure 8). Documents can be stored in many formats, including simple image files, Adobe Acrobat format, or more sophisticated XML documents.

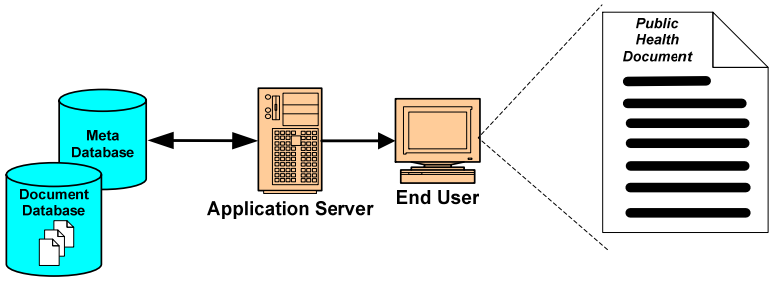


Figure 8 – Document-centered Storage Approach

Similarly, there are two predominant styles of data interoperability *between* systems. A *data-centered* approach uses traditional structures to represent the data set being transported. In simple cases, this means a row in a file for a record, and either delimited or fixed length fields within the record. Metadata describing the structure of the file may be a simple header row before the data or a separate file containing more detailed field descriptions, code sets, or semantic explanations. More sophisticated examples include HL7 or X12 messages which follow a well-developed, standards-based syntax detailed in implementation guides or profiles (see Figure 9).

```
MSH|^~\&||||||VXU^V04|19970522MA53|P|2.3.1|
PID|||221345671^^^^SS||KENNEDY^JOHN^FITZGERALD^JR|BOUVIER^^^^^^M|19900607|M|||~^^^^MA^^^BDL|
NK1|1|KENNEDY^JACQUELINE^LEE|MTH^MOTHER^HL70063|
RXA|0|1|19900607|19900607|08^HEPB-PEDIATRIC/ADOLESCENT^CVX|.5|ML^^ISO+||||||||
MRK12345||MSD^MERCK^MVX|
```

Figure 9 – Sample HL7 V2.n Message

The second style of data interoperability is *document-centered*. In this case, the data is pre-arranged in a document format which is usually quite structured. Simply opening up and examining the document itself conveys its contents in an organized, labeled fashion. The best example of this approach in the clinical world is ASTM's Continuity of Care Record[13] (CCR), which contains a pre-determined set of data in a pre-determined format (Figure 10). HL7 has created a more generic architecture for creating data in this style (Clinical Document Architecture, or CDA), and the two organizations have combined the two by developing an implementation of the CCR using CDA technology called the Continuity of Care Document (CCD). Both CCR and CCD represent *summaries* of clinical information about a specific patient.

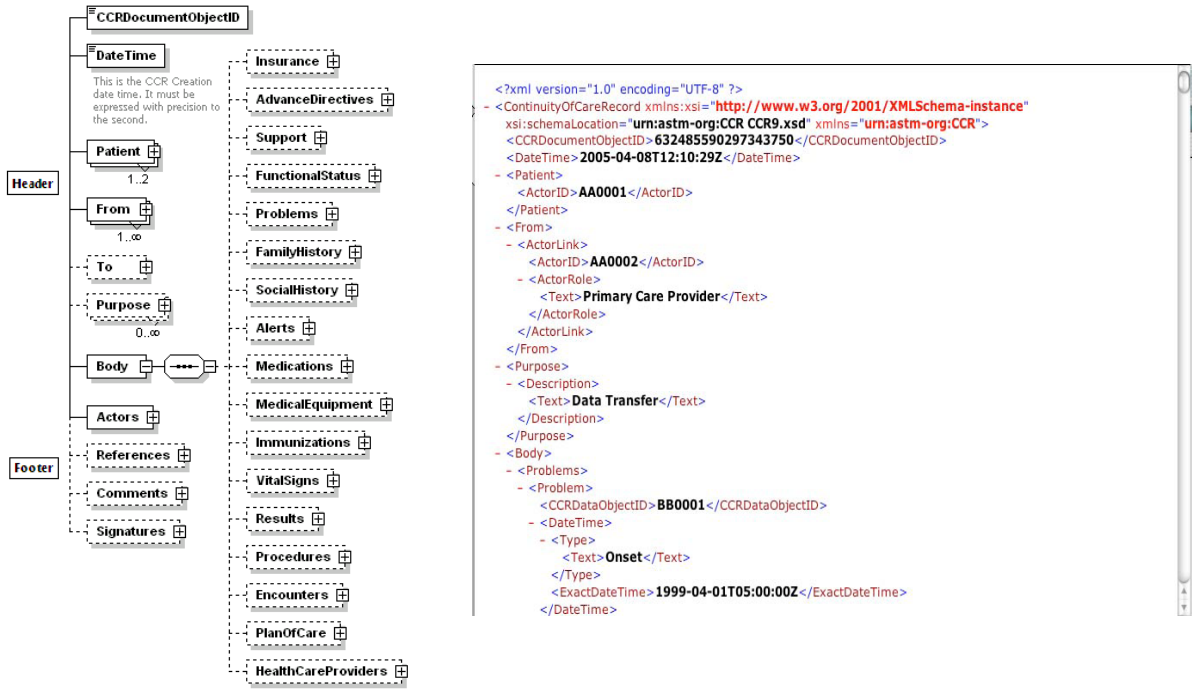


Figure 10 – Make-up of CCR Record[14] and Sample XML Representation[15]

Since data exchange partners do not always control the attributes of the other's system, the style of data storage and the style of data interoperability may not be the same for a particular interaction. However, certain combinations present particular benefits or drawbacks and shown in Figure 11. When data is made to interoperate in its native storage style (data-centered or document-centered) there are fewer challenges (lower left quadrant; upper right quadrant). When these strategies do not match, it may be more challenging on one end of the transaction or the other (upper left quadrant; lower right quadrant).

What does this mean for an IIS? Typically, IIS store data in a data-centered way (we are not aware of any IIS with a document-centered storage approach). And typically, IIS exchange data in a data-centered way as well, as that is the dominant style for existing flat-file and HL-7-based messaging solutions. Other public health data exchange initiatives, like the Public Health Information Network (PHIN) Messaging System (PHINMS) deployed in many jurisdictions to support disease or syndromic surveillance, are also

[13] http://www.astm.org/cgi-bin/SoftCart.exe/DATABASE.CART/REDLINE_PAGES/E2369.htm?L+mystore+yirv0526+1174970569
[14] Modified from diagram found at http://www.astm.org/COMMIT/E31_CCR0305.ppt
[15] See http://www.centerforhit.org/PreBuilt/chit_ccrnyc.pdf

data-centered. Data-centered approaches also continue to be dominant *within* organizations or agencies. However, as the healthcare industry continues to develop interoperability standards and specifications, there appears to be a movement towards document-centered approaches to interoperability.

Data Storage	Data-centered	Document-centered
Document-centered	May be difficult to extract discreet data from documents and assemble into the desired message or file format. Receiving data-centered messages and storing them in the databases as documents is less challenging.	Relatively easy to extract documents, transport them as such, and store them as documents in the destination system.
Data-centered	Relatively easy to extract data and assemble in the desired message or file format. Interface engines exist which facilitate parsing data from databases into messages and vice versa.	Relatively easy to extract data and assemble in the desired document format. May prove more challenging to parse documents back into discreet data elements for storage in the destination system, depending on the form of document used.

Data Interoperability

Figure 11 – Data-centered and Document-centered

This typically puts IIS squarely in the lower right corner of the matrix in Figure 11. When developing their overall architecture for data exchange (see section above), IIS projects will need to consider their data storage strategy as well as the appropriate data interoperability strategy and plan solutions accordingly. The data interoperability strategy will likely need to be developed in consultation with data partners in the project's environment.

Consider the appropriate strategy based on system capabilities and the needs of data trading partners you are targeting. It may seem obvious that data-centered strategies dominate today, but the HIEN world is shifting to a document-centered direction.

3.4 Message/File Format

A number of message/file formats are possible when considering how to structure the data that needs to be sent and received. Here are a few important considerations:

- Consider the capabilities of both the IIS as well as the systems with which the IIS wants to interoperate. The "data-centered versus document-centered" discussion above is but one consideration.

- Fewer formats are usually better than many formats. This allows you to focus your efforts on developing correct, robust interoperability without spreading your resources (particularly people resources) too thin. Projects are tempted to accept data however the sender wants to send it as a way to encourage data exchange, but in the end poor quality or unreliable data exchange is worse than no data exchange at all.

- Leverage existing formats over creating new ones. Reuse existing formats from your agency (or from your data partner organizations) if they are a good fit.

- Use formats that are provided by national and international standards development organizations (SDO) over proprietary formats. This will increase the likelihood that the specifications of the format will be well documented, well thought out, and adopted by your data partners.

- Start simple: often simple text files are easier to implement than other, more complex, formats.

Various file formats are possible, both proprietary and standards-based. They fall along a continuum of simpler to more complex solutions:

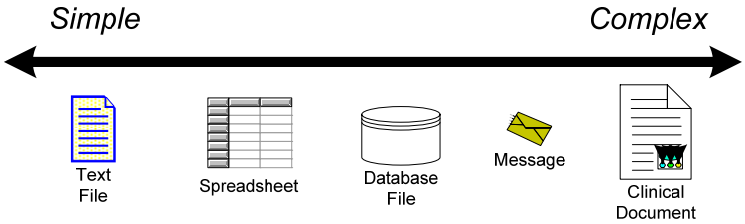


Figure 12 – File Format Continuum

Format	Attributes and Comments
Text File	Usually the easiest to deal with. Many IIS projects begin here, and many began here in the past as no other standards-based format existed years ago. Can be delimited (comma or symbol separating data fields) or fixed-length (fields begin and end based on their position in a row in the file). Can contain different types of structures (see below).
Spreadsheet	Since electronic spreadsheets are fairly ubiquitous, some IIS accept data in this format. It is advised to proscribe a standard order and format for spreadsheet columns, but given their flexible nature it is easy for data partners to make mistakes so all files need to be examined for quality and completeness.

Format	Attributes and Comments
Database File	Simple database management systems like Microsoft Access are also fairly common, so some projects accept data in this format and use local processes to extract, transform, and load data into the IIS. More skill is required on the part of both data partners to work within this type of format.
Message	Increasingly, sending systems are becoming capable of generating standards-based messages (like HL7) and this is consistent with emerging national standards. This represents a data-centered approach to data exchange. Messages can be generated individually in real-time or transmitted as a batch, and can go in one direction or be bi-directional (see discussion elsewhere in this document). Many message formats exist – standard and proprietary – though most use either an ASCII text format or XML.
Clinical Document	Use of clinical document formats such as ASTM's Continuity of Care Record (CCR) and HL7's Continuity of Care Document (CCD) are gaining in popularity. They represent a document-centered approach to data exchange and have not yet been adopted within the IIS community.

One additional consideration for text and message files is the overall file structure. Several options exist, including:

Style	Example	Comments
Single, complete row per transaction	NYC DEI (Data Exchange Format) used primarily to send records to health plans	Simpler to use, but need to have a strategy for dealing with records that may have differing quantities of immunization transaction records
Multiple rows with different record segments	HL7 v2.x messages	More complex but easily standardized. Usually employs standard message "segments" which can be repeated if necessary to accommodate multiple parts of a record if needed.
	NYC UPIF (Universal Provider Interface Format): demographic records are submitted in one record, immunization transactions in another	Also more complex, but there is no minimum number of record segments required for a file. Care needs to be taken to process files in the correct order as there may be data dependencies between record segments that may appear in different files.

Start as simply as you can by considering the current *and future* capabilities of your data exchange partners. Standards-based formats will increase in prominence and importance over the next several years.

3.5 Record Contents

Choosing exactly what information to include in a data exchange is important. In this section we will focus primarily on choices about record contents for submissions *to* the IIS since the recipient of data *from* the IIS will largely determine the record requirements in that direction. The CDC has defined a core data set for records exchange[16] (required data elements appear in bold). In 2007, the National Vaccine Advisory Council (NVAC) issued an IIS Progress Report in which some additional recommendations were made regarding data elements.[17] The following represents a consolidated recommendation:

Patient/System/State Identifiers:

- **Patient name: first, middle, last**
- Patient alias name: first, middle, last
- Patient address, phone number, birthing facility
- Patient Social Security number (SSN)
- **Patient birth date**
- **Patient sex**
- **Patient race/ethnicity**
- Patient primary language
- **Patient birth order** (multiple births)
- Patient birth registration number
- **Patient birth State/country**
- Patient Medicaid number
- **Mother's name: first, middle, last, maiden**
- Mother's SSN
- Father's name: first, middle, last
- Father's SSN
- Patient status indicator

Immunization Event Identifiers:

- **Vaccine Type**
- **Vaccine Manufacturer**
- Vaccine expiration date
- Vaccine injection site
- **Vaccination date**
- **Vaccine lot number**
- Vaccine provider
- Historical vaccine indicator
- VFC eligibility
- History of varicella indicator

Note that this data set was developed originally for *IIS to IIS* data exchange and not EHR-S to IIS data exchange, so the list should be critically examined by an IIS project in this new context. Some data elements – even required data elements – may not be available in practical terms in EHR systems making their requirement in data exchange difficult (e.g., mother's name, patient race/ethnicity).

Some additional issues need to be considered, including:

- **Demographic data**: Some minimal set of demographic data should be required in order for a record to be accepted. In many jurisdictions it is no more than first name, last name, date of birth, and gender. But the reason to require a minimal set of data is to ensure a good-enough set of data to identify a record when querying the IIS database will be present. While mother's

[16] See http://www.cdc.gov/vaccines/programs/iis/stds/coredata.htm
[17] See http://www.hhs.gov/nvpo/nvac/NVACIISReport20070911.doc

maiden name may be a required CDC data element, it is often not routinely captured in an EHR-S and should usually not be required for records received from that source.

- **Immunization history**: Most IIS would like to get the immunization history that the data partner may have, if it is available. It is important to identify the *source* of each immunization – whether it was administered by the data partner organization itself or whether it was collected and recorded by the data partner from other sources (like a parent's immunization "card" or from a payer claim.). In this way the IIS can judge the veracity of the data.

- **Currently-administered immunizations**: The primary focus of the IIS is usually on the immunization events from immunizations currently being administered by the data partner organization. The provider may be required by law to submit this information or it may be voluntary.

- **Completeness versus compliance**: There is a basic tradeoff: the more an IIS requires of a data partner the less likely the data partner will be able to comply consistently. IIS needs to consider the nature and profile of data partner systems when making decisions about required data content, while at the same time not compromising the data quality standards of the IIS.

- **Semantic interoperability**: See the section below on code sets for a discussion of the important of consistent, standards-based data coding.

There is one important issue with respect to transfer of data *from* an IIS:

- **Decision support data**: Most IIS can generate immunization recommendations based on a patient history and age. This information can optionally be provided to data exchange partners within a message or file. Many data exchange partners may be unable to store this information in their local systems and may not even want it included in the transmission, while others may want it to populate in their local system.

Consider the CDC Core Data Set as well as the capabilities of the systems with which you want to exchange data before determining the specific data exchange record contents. Be sure to test with *real* data to make sure the selected data set is feasible.

3.6 Direction of Data Transfer

We have touched upon the topic of the direction of data transfer several times in this document already, and it is captured in the IIS Interoperability Model above. Data can flow from an external system to the IIS, from the IIS to an external system, or both. IIS projects much decide which direction is required to support the primary business needs and implement solutions accordingly. Some additional points:

- It is usually best to begin with the most important function: the import of immunization records from providers *to the IIS* to satisfy basic reporting requirements mandated by law or being encouraged to populate the IIS's core data. Work directly with providers, medical or hospital societies, and community coalitions to define the data submission needs in your area.

- The first transfers of data *from the IIS* usually involve sending immunization data to health plans to satisfy their HEDIS reporting requirements. Try to develop a standard format and interface for doing this across all health plans. Health plans are usually required to submit demographic data from a panel of patients to be used by the IIS to match records for extraction.

- As IIS moves to implement HL7, this often translates into a phased implementation approach:

  - Begin with the VXU message for data submission which represents a one-way flow of information *to* the IIS. Keep in mind that this may mean the implementation of acknowledgement message (ACK) which does represent two-way communication. Many external systems simply ignore the acknowledgement messages anyway.

  - Later add the VXQ unsolicited query message and corresponding VXR response message to provide data *back* to participating external systems.

Carefully consider the direction of data movement and don't bite off more than you can chew! A phased approach is usually better than rushing to implement it all.

3.7 Person Matching

One of the most critical issues for data exchange with an IIS is the strategy for matching and de-duplicating person records between systems. Often, a Master Patient Index, or MPI, is created or acquired to match records or de-duplicate sets of records that appear to have the same patient represented more than once. The IIS can serve as that MPI in many circumstances, or just concern itself with its own matching needs. Systems can implement matching on the front-end, that is, when new records are being integrated into a database and one wants to determine whether the person represented by the new record is already represented in the database, or on the back-end, that is, when one wants to examine an existing database or data set and determine if it already has duplicated records within it.

There are many software products available to perform provide these services, but software acquisition is only part of the puzzle. It is estimated that only a third of the total cost of deploying an MPI pays for the software itself (Figure 13). Another third of the cost pays for the development of an appropriate architecture within which the MPI will

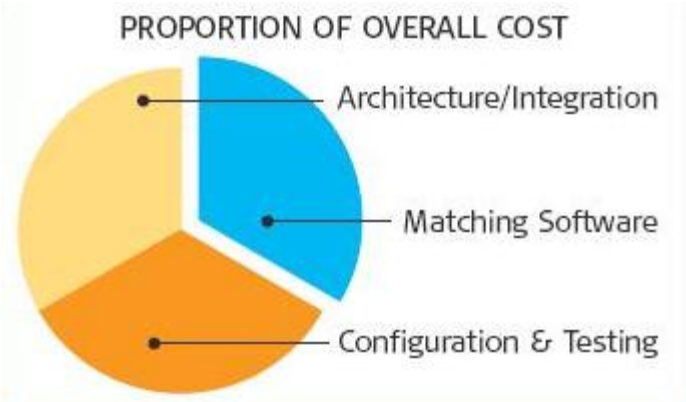


Figure 13 – Cost of Deploying an MPI

function, as well as the integration of the MPI into the other components of the agency's systems. The final third pays for configuration and testing—by agency staff or the MPI vendor—to ensure that the software works properly with the agency's data and systems.

The core of any de-duplication and matching service is the type of algorithm that is used to determine if two records are the same or different. There are two basic types of matching strategies:

1. **Deterministic Matching** uses sets of predetermined rules to guide the matching process. The rules rely on a series of exact matches between data elements to identify when records match. It is most successful when the data is of relatively high quality or is dominated by reliable unique identifiers for records. Deterministic matching is less successful when the data is incomplete or inaccurate, when there are many spelling or transcription errors, or lots of inconsistencies (e.g., frequent name changes).

2. **Probabilistic Matching** is a process whereby an estimate is made of the probability that two records are for the same person based on the degree to which certain fields in the two records match. Two thresholds are then set:

   - All record pairs whose probability is above the higher threshold are considered to be matches.
   - All record pairs whose probability is below the lower threshold are considered not to be matches.

   The disposition of record pairs whose probability falls in between the two thresholds is considered to be uncertain and they require additional review, likely by a trained staff member.

In 2006, the Public Health Informatics Institute released its Unique Records Portfolio, a detailed guide for public health agencies to use to develop a strategy for person record matching and de-duplication.[18] It contains not only a theoretical background in the key concepts, but practical checklists for action as well as case studies describing successful efforts. This section will not attempt to reproduce the entire contents of this seminal document.

IIS projects need to decide on an appropriate set of de-duplication strategies both when sending and receiving records. The issues differ slightly depending on the direction of the records movement. When an IIS is receiving data,

- IIS needs to determine if the incoming record is for a new or existing patient. Hopefully, services exist within the IIS software to determine this with some high level of confidence.

- It is important to prevent both record fragmentation (i.e., the same person is represented by more than one record in the database) and false merges (i.e., records from two different people incorrectly linked, joined or merged together).

- If possible, rely on the same algorithms to determine whether an incoming record is a new record or not as used by the IIS' interactive (web) software. This will ensure a consistent

[18] See http://www.phii.org/resources/UniqueRecordsPortfolio.asp

understanding by users and staff of the sometimes complex decision as to whether a record is new or not.

- IIS needs to provide tools for its own trained staff to review ambiguous incoming records easily and quickly and decide if the new record should be matched with an existing record in the database.

- IIS should encourage data exchange partners to retrieve and store the internal IIS ID of each person in local systems, and to provide that IIS ID in subsequent data submissions to facilitate matching accurately. Matching should never be done *solely* with this IIS ID but should be corroborated with additional data elements found in the incoming record (e.g., name, date of birth).

When sending records *from* an IIS to another system, a slightly different set of issue present themselves:

- IIS should encourage data exchange partners to retrieve and store the internal IIS ID of each person in local systems to facilitate matching records that are sent from the IIS.

- Systems requesting data from the IIS need to prove that they have a legitimate reason to ask for the specific records. Most IIS projects have policies surrounding how (and how many) records can be requested.

- It is possible that a request for data from an outside data exchange partner will result in *multiple* records matching the request criteria, either because the IIS has potentially duplicative or fragmented records, or because the criteria used to request records from the IIS were not specific or narrow enough. The IIS project needs to determine a policy for these circumstances – most systems do not return *any* data to the requestor if a unique record did not result from the query.

Use PHII's *Unique Records Portfolio* as a guide to developing policies and strategies surrounding person record matching and de-duplication.

3.8 Data Quality Assurance

Data quality assurance is critically important to any successful electronic data file transfer. It is important in all aspects of interoperability: technical, semantic, and process. What good is data properly received and understood if it is inherently incorrect or misleading? While this section does not provide a comprehensive guide for establishing a formal data quality process, it does hit upon some of the important aspects of such a program. More details can be found in the AIRA MIROW recommendations on this topic in *Data Quality Assurance in Immunization Information Systems: Incoming Data*:[19]

[19] See http://www.immregistries.org/pdf/AIRA_MIROW_Chap3_DQA_02112008.pdf

- Data quality assurance requires a formal process, applied consistently to all incoming and outgoing data exchange actions. Many IIS projects dedicate specific staff members to this task who develop specialized skills and techniques for ensuring quality.

- Part of this formal process is the qualification of a data exchange partner's data as being of sufficient quality as to permit the data to be accepted routinely by the IIS (or by the other systems). While the focus here is on the IIS' acceptance of external data, other data exchange partners have just as much right as the IIS to demand that IIS data bound for their systems be qualified through a rigorous process.

- Data quality assurance requires vigilance. Data received from data exchange partners may be of an acceptable quality one day, but not acceptable another. Because of these potential inconsistencies over time, periodic re-testing of data is recommended.

- Create and maintain a formal guide document to explain all the requirements for data exchange. This document should also guide the requirements for the testing process.[20]

- A formal data quality assurance program should have a documented set of steps for qualifying a new data submission. This qualification should ensure that technical, semantic, and process interoperability goals of the data exchange program are upheld consistently. Here are some tips for such a program:

    - If possible, compare data in the source system with its final representation in the destination system to ensure consistency (in other words, does the data seem to look the same where it finally ended up as it did when it started?).

    - Document all code sets used in the data exchange process. Be sure to keep data exchange partners up-to-date as to any changes in accepted or required codes. Use AIRA's IIS Codebook whenever possible. Translation tables may be necessary to convert codes being received from (or sent to) data exchange partners into (or from) IIS project code values.

    - For vaccine data, consider using the AIRA MIROW recommendations in *Vaccine Level Deduplication in Immunization Information Systems*.[21]

- When it comes to testing incoming data, here are some useful tips:

    - Use real data, rather than just sample or artificial data, to test transfers. This ensures that anomalies are not introduced into the testing process that may be a factor of the creation of a test data set only.

[20] For examples, see the New York City *HL7 Web Service Integration Guide for Immunization Transactions* and the Rhode Island *KIDSNET HL7 Implementation Guide for Immunization Transactions* <http://www.health.ri.gov/family/kidsnet/dataexchange.php>

[21] See http://www.immregistries.org/pdf/AIRA_BP_guide_Vaccine_DeDup_120706.pdf

- Stage your testing in a non-production testing database, but one which mirrors the production database as closely as possible. This will ensure that no mistakes occur with test data being commingled with real data.

- Test a sufficient quantity of data (usually a week's worth is good for all but the smallest sized practice or site) to ensure that a representative set of diversity of data is represented.

- Check first for the presence of required fields per your documentation. Data exchange files or messages should be rejected if required fields are not present.

- Check for acceptable code values per your documentation (semantic interoperability). Data exchange files or messages should be rejected if acceptable code values are not used.

- Check for a logical combination of codes and the presence of combinations of fields you might logically expect to find together (process interoperability). For instance, you would not expect to see an immunization for Varicella if a disease indicator shows that the child already had chicken pox.

- Test the data communications that you have defined for the exchange process, including authentication and authorization steps (technical interoperability). Be sure to simulate an *inappropriate* access attempt to be sure that security rules are in force.

♦ Your IIS project needs to consider what to do about files or messages that are received which contain either erroneous data (e.g., missing required data) or inappropriate data (e.g., adult demographics and/or immunizations when only childhood data is accepted). While clearly such data should not be processed and added to the IIS database, it is less clear whether a submission file or message should be rejected in its *entirety* should only part of it be erroneous or inappropriate. It is simpler for the IIS to reject an entire submission if it is problematic. However, IIS projects run the risk of "throwing out the baby with the bath water" by rejecting good as well as bad records. If the IIS chooses to keep just the good records and reject the problematic ones, additional processing and reporting capabilities will need to be provided (for example, see Section 3.12 on Process Monitoring and Error Correction below).

Finally, the following issues deserve special attention when examining files for quality:

Issue	Strategy
Date Fields	Date fields are critical to IIS data and require some particular attention and checking. Specifically, be sure that immunization transactions dates do not occur before the patient's date of birth. Some errors concerning dates are more subtle and may not be noticed unless searched for specifically. A data file/message may be suspect if many or all dates in the file are identical, or if components of many or all dates are identical (for instance, all dates on the first, middle, or last day of the month; date component missing (usually the day); use of impossible date components suggesting a default date of some kind (usually '99'). You may want to alter your testing or actual production processing to scan an *entire* submission for some of these anomalies.
Combination Vaccines	Combination vaccines may not be submitted consistently by all data exchange partners. Sometimes their components are submitted as individual antigens. You may want to check for this condition and work with data exchange partners to ensure that no misinterpretation of immunization data takes place.
Missing Manufacturer Information	In some cases, the actual manufacturer of a vaccine that is administered affects the immunization schedule for the patient who receives it. Missing manufacturer information (even though it *is* part of the minimum core data set identified in Section 3.5 above) may occur in data exchange. You need to develop a clear policy on whether to enforce this rule or not, or whether to selectively enforce the rule when it is likely to have the strongest impact.
VFC Eligibility	Increasingly, IIS is being asked to help track compliance with the use of Vaccines for Children (VFC) vaccines within your jurisdiction. Since certain eligibility for the VFC program changes as a child's circumstances change, it must be verified as each VFC dose is administered. Two problems present themselves: 1. Many EHR systems do not have the capacity to record VFC eligibility and therefore cannot pass it along to an IIS along with the immunization data. 2. If it is recorded, often it is recorded as a function of the *patient* globally and not a feature of the individual immunization event. This makes it difficult to accurately keep track of eligibility or to verify eligibility at some point in the past. There is no "quick fix" for this problem, but awareness of the issue is important to ensure that data is not misinterpreted or mis-used.

Issue	Strategy
Inadequate Vaccine Type Coding	Some data sources do not use coded schemes for vaccine type, let along standard coded schemes. This may result in some data exchange files/messages containing text fields with vaccine type. Projects need to decide whether to outright reject such submissions or whether to apply some type of pre-processing to change the text field values to standard vaccine type codes. In addition, various coding schemes exist for vaccines, including CPT codes from clinical systems and CVX codes from HL7. IIS may need to store one or both of these codes, with CVX codes representing the national standard.
Middle Name/First Name Issues	Some data sources may include a patient's middle name embedded in the patient's first name field, consistently or inconsistently. Projects need to decide whether to outright reject such submissions or whether to apply some type of pre-processing to identify and parse the middle name from the first name and store it in a unique field.
Race Code	Many different coding schemes exist for patient race, and some EHR systems do not even bother to record it at all. Use the standard coding found in the AIRA IIS Codebook or translate externally-received codes to the Codebook values before storing the data in the IIS.
Immunization Source	It is important for submission of immunization data *to* an IIS for the sender to differentiate between immunizations actually administered at the provider location (to which they can attest) and immunizations provided by other clinical locations whose data is being provided as a part of a child's recorded history. Be sure you are receiving, and can capture and save, this distinction.

Many strategies have been included here, and the discussion is far from comprehensive. Develop a set of strategies that work for you, and rely on the AIRA MIROW chapters referred to in this section for guidance.

3.9 Data Transport

Data transport is an important component of technical interoperability. If data cannot be reliably moved from where it is to where it needs to be interoperability cannot take place successfully. Data transport strategies must be acceptable to all parties involved, which is one reason why IIS projects often support *multiple* techniques. While this may be necessary to accommodate a variety of data exchange partners, projects should strive to support *as few techniques as possible* for data transport. Finally, whatever strategies are selected must be compliant with relevant security/privacy laws and regulations in your jurisdiction, including national laws (e.g., HIPAA), state/local laws or ordinances, and institutional/organizational policies.

Though there are many possible strategies, the following are the most common ones in use by IIS projects. They fall along a technical continuum from simpler (though not necessarily more reliable or safe) to more complex:

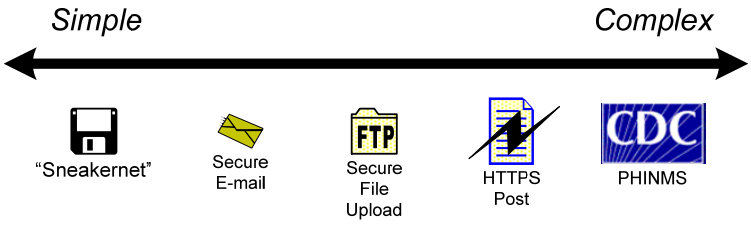


Figure 14 – Data Transport Strategy Continuum

Strategy	Attributes and Comments
"Sneakernet"	By far the simplest method of transport, this usually involves exchanging of removable media such as CD-ROM/DVD, tape, or perhaps USB drive. Data is secure to the degree that the media itself is secure. The biggest barrier to this kind of transport is the added time it takes for the media to physically move from sender to recipient. Once it arrives, and the data is extracted from it, the data is handled just as any other data receipt.
Secure E-mail	Some IIS projects allow data exchange partners to send files or messages via e-mail. If this is permitted, it *must* be done using a secure e-mail product like PGP.[22] If at all possible, files received in this manner should be removed from the email environment (and purged) as soon as possible. Once the files or messages are removed the data is handled just as any other data receipt.
Secure File Upload	There are a variety of secure file upload products available, some better than others. Typically, these are web-based environments or file systems that assign accounts (i.e., usernames and passwords) to authorized users and provide some disk space on server to which they upload (or from which they download) files. Various Internet protocols can be used, including Secure FTP, SSH, and SSL. Some of these products support more automated (as opposed to keyboard sessions) for file operations. Some IIS projects also implement various file submission status reports/screens as well as online submission error correction.[23]

[22] See http://www.pgp.com/products/desktop_email/index.html
[23] For an example of a product such as this see HLN's Web File Repository (WFR), <http://www.hln.com/collaboration/tools/wfr.php>

Strategy	Attributes and Comments
HTTPS Post	This method of data transport is used for both files and messages, though it has become most popular for message-based transport. It is advantageous because it uses standard Internet HTTP protocols and SSL encryption during data transport that allow messages to easily and securely transverse organizational firewalls and other physical security devices. HTTPS Post should always be implemented with some type of authentication means (usually username and password) so that transmissions can be validated as authentic.
PHINMS	The CDC's Public Health Information Network Messaging Service (PHINMS) was developed as a peer-to-peer messaging protocol for bio-surveillance and other PHIN-related communications.[24] CDC offers a set of messaging guides to assist in setting up the PHINMS environment. It is built on a transport protocol called ebXML.[25] PHINMS is a good choice if the IIS project needs to rely on standards already in place within the health agency. It does, however, require implementation by all communicating parties and may be considered more difficult to implement in smaller sites with less information technology support services. Hospitals are more likely to be running PHINMS as part of a separate disease surveillance or syndromic surveillance project, so some leverage of that effort may be possible.

Authentication of data exchange transactions is very important. Once again, several options are available:

Strategy	Attributes and Comments
Single-factor Authentication	The simplest and most prevalent form of authentication is single-factor authentication, usually a username and password. This is the minimum level of authentication that is acceptable. Passwords should be "strong" (several schemes are available that require mixtures of upper case and lower case letters, as well as numbers, symbols, and a minimum password length) and always encrypted. For server-to-server communications passwords are also used to authenticate data communications partners (sometimes referred to as "shared secrets" in some implementations).

[24] See http://www.cdc.gov/phin/activities/standards/messaging-transport.html
[25] See http://www.ebxml.org/

Strategy	Attributes and Comments
Two-factor authentication	If practical, traditional passwords can be augmented by a second factor of authentication provided by additional software installed on the server receiving the file or message communication. This second level of authentication is usually in the form of an additional password challenge that prompts for a one-time password good only for the specific session. These one-time passwords are generated a number of ways, including a hardware token device[26] or even a unique, one-time communication to the user.[27] For some file/message transport schemes – particularly those that take place in an automated or semi-automated fashion with little direct user interaction – two-factor authentication may not be practical or possible.
X.509 Digital Certificates	Digital certificates are installed on a computer and used to authenticate the identity of the user *on that computer* (it is possible to have certificates authenticate an individual wherever on the Internet that individual happens to be located, but that use is far less common). Digital certificates are a good way to provide a second factor of authentication for server-to-server data transport especially since the identity of the participating servers can be validated by all participants.

Data transport must be reliable and secure. Use the least complex alternative available, and as few alternatives as necessary to meet your needs. Encryption of data during transport and strong authentication of participants is essential.

3.10 Record Modification and Deletion

Special attention needs to be taken with respect to determining a strategy for accepting modifications or deletions of IIS data via electronic data transfer. Many EHR systems do not in fact modify data stored for patients, but rather they delete the particular record or portion of a record and then add a new record as it needs to be (that is, with the old data to be retained and the new data to be modified). This is done primarily for record auditing purposes, to ensure that the state of a patient's record can be reconstituted at any moment in time to establish what the clinician had available for consideration clinically at that time. While an IIS may or may not share this auditing concern, it may be common for projects to receive data "modification" requests that come in the form of requests for deletion and addition of data.

Deletion requests should only be considered valid if coming from the facility or site that originally submitted the data. In addition, some IIS may consider only deleting the data if it represents an immunization actually administered at the sender's facility or practice (see Section 3.8 above). Extreme

[26] See http://www.rsa.com/node.aspx?id=1156
[27] See http://www.anakam.com/web/products/authentication.jsp

care should be taken when deciding how to handle these requests, and whether to automate deletion requests at all or process them manually. IIS projects should also consider the potential that records may have been modified since their initial submission, so requests for deletion may no longer correspond to the records as they exist in the IIS. This condition needs to be considered before deleting (or further) modifying records found in the IIS.

One more note: it is important to ensure that records received via electronic interface that contain empty or missing values do not cause actual data in the IIS to be overwritten. Be sure that appropriate business rules are in place to deal with this condition.

Be mindful of the implications of record deletion and modification to data integrity and choose strategies widely to audit and protect your data.

3.11 Batch Mode versus Real-time Mode

There are two predominant modes of receiving data into an IIS. *Batch Mode* involves receiving data without presuming that it will be processed into the IIS database immediately. *Real-time Mode* involved receiving data and presuming a fairly immediate processing of that data with the intention of storing it in the database (assuming it passes the various business rules and requirements) as soon as possible. IIS projects usually pick one, or both, of these modes for their operations. The following attributes may help you select which mode is best for your project:

Issue	Batch Mode	Real-time Mode
Typical number of records	Any number of records, but the more the merrier	Usually a single record or request per file/message
Predominant File/message Format	Almost any type	Usually message-based
Predominant Data Transport	Almost any type	HTTPS Post, PHINMS
Direction of Data Transfer	One way, as there is no provision for response to the sender	Usually two-way; sometimes response is no more than an acknowledgement.
Best Use	Transmission of data on a routine schedule or when data transport is less sophisticated. Often simpler for data exchange partners to implement.	When immediate response to a transmission is desirable, or when there is no routine schedule for transmissions in the business process

If possible, start simpler with a batch mode interface to data submission and gradually work towards a real-time mode if data exchange partners require the immediacy. As EHR systems become more prevalent real-time interfaces will become more necessary.

3.12 Process Monitoring and Error Correction

The more routine interoperability becomes, the greater the volume of data exchanged, the more it may become necessary to provide additional features and tools for data exchange partners interacting with the IIS. *Process monitoring* tools allow data exchange partners to view the status of data submissions or data requests through an online computer application, usually on the Web. *Error correction* tools allow data submitters to review data submissions that the IIS determines were incomplete or suspects were erroneous in some way and submit corrections to these records on-line. Together, these tools provide a greater degree of "self service" to data exchange partners and reduce the amount of support time required by IIS staff to answer questions and make corrections.
In New York City, for example, the Citywide Immunization Registry allows data submitters to upload files anytime to its secure Web File Repository over the Internet. Once files are uploaded, they appear in a listing along with their current status (see Figure 15):

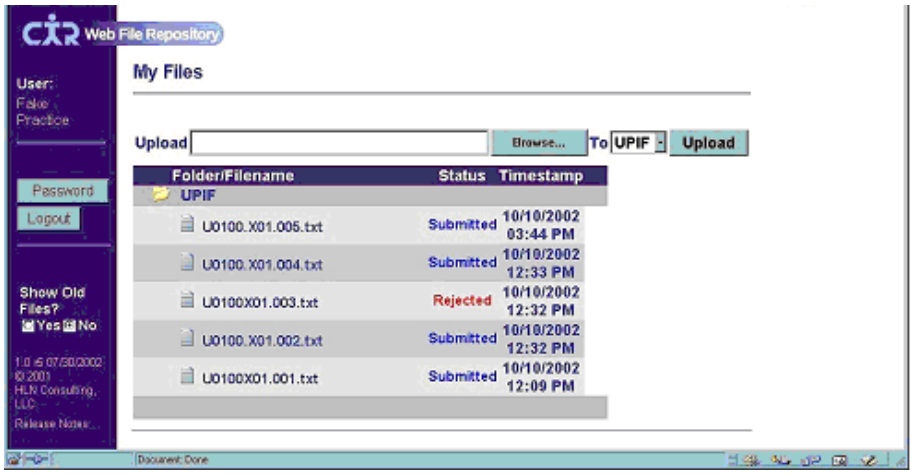


Figure 15 – Web File Repository File Listing

If the user wants to see the details surrounding the status of any particular file, the user merely needs to click on the Status link. Figure 16 shows the details for the rejected file from above:

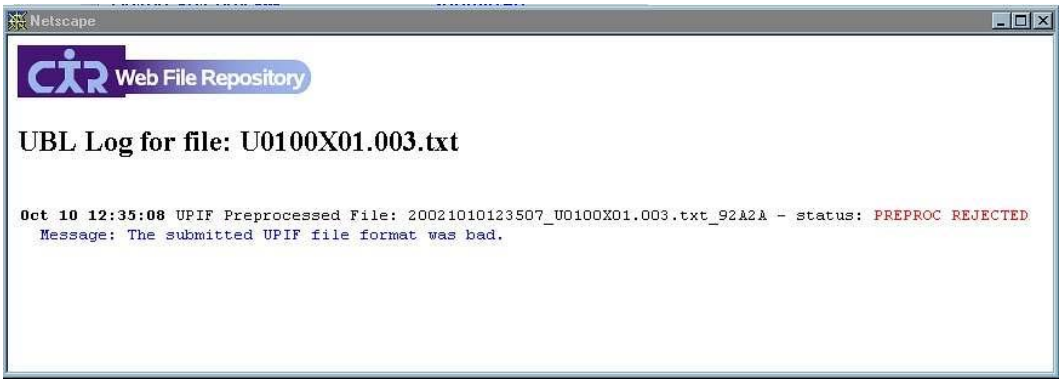


Figure 16 – Web File Repository Rejected File Explanation

At this time, users are unable to modify file submissions to correct errors, but that capability may be added in the future. Users can receive email messages, however, as file status changes through routine processing.[28]

Consider advanced features such as online process monitoring and error correction as a way to improve customer service to data exchange partners and reduce IIS staff time requirements.

[28] See Arzt, Noam H. and Amy Metroka, *Implementing a HIPAA Compliant Electronic File Transfer Application for Reporting to the NY Citywide Immunization Registry,* 2002 Immunization Registry Conference, Philadelphia, PA, October 29, 2002. <http://www.hln.com/noam/WFR-IRC-1002.pdf>

4 Looking to the Future

Trends in the HIT marketplace may bring some unintended side effects. Deployment of EHR systems in provider practices may interfere with the providers' ability to submit data to the IIS the way they used to, as underlying systems change and the interfaces once carefully worked out with the IIS staff no longer function. Similarly, the IIS may move to new, more technically sound methods of data exchange and interfere with existing data exchange processes and capabilities unless care is taken to make sure those older techniques are examined and either replaced or maintained until no longer needed. Finally, the deployment of EHR systems, public health-sponsored chronic care systems, or HIEN portal applications may reduce IIS users' reliance on traditional web-based IIS applications as users strive to simplify their application environment and reduce the number of competing web-based applications all with overlapping sets of patient data. These factors should *increase* the importance of electronic data exchange both to and from IIS and increase the healthcare community's reliance on these capabilities.

In the standards world, there has been much activity when it comes to immunization data interoperability, though it will take years for these standards to fully evolve and become implemented. Initial efforts have been uneven, and somewhat complicated to develop and deploy. The Health Information Technology Standards Panel is finally taking up the issue of immunization data interoperability in 2008 with the publishing of the Immunizations and Response Management Use Case for which interoperability specifications are being developed.[29] HL7 continues to work on a domain model and message constructs for its Version 3.0 standard, while the Integrating the Healthcare Enterprise (IHE) collaboration begins to consider IIS needs in its profiles and documents. Technology planners also are realizing that all parts of interoperability – technical, semantic, and process – need to be addressed in order for data to flow smoothly and intelligibly. Increasingly, they will focus on such concepts as service-oriented architecture (SOA) to help move data transparently, reliably and securely between locations. Over time, we hope that these many worthy efforts converge to form a fully-interoperable health information exchange environment for the country and the world.

[29] See http://www.hhs.gov/healthit/usecases/respmgmt.html

Appendix A: Terms and Acronyms

AFIX	Assessment, Feedback, Incentives, and Exchange http://www.cdc.gov/vaccines/programs/afix/default.htm
AIRA	American Immunization Registry Association http://www.immregistries.org/
CDC	Centers for Disease Control and Prevention http://www.cdc.gov/
CPOE	Computerized Physician Order Entry
CPT®	Current Procedural Terminology http://www.ama-assn.org/ama/pub/category/3113.html
CRA	Countermeasure and Response Administration (part of PHIN) http://www.cdc.gov/phin/activities/applications-services/cra/index.html
CVX	Codes for Vaccines Administered http://www.cdc.gov/vaccines/programs/iis/stds/downloads/hl7-cvx.pdf
DEI	Data Exchange Interface (NYC) http://www.nyc.gov/html/doh/downloads/pdf/cir/DEIspecifications_113006.pdf
EHR-S	Electronic Health Record System
ebXML	Electronic Business using eXtensible Markup Language http://www.ebxml.org/
ePrescribing	Software function that allows clinicians to order/renew patient prescriptions online directly to participating pharmacies.
FTP	File Transfer Protocol
HEDIS	Healthcare Effectiveness Data and Information Set http://www.ncqa.org/tabid/59/Default.aspx
HHS	US Department of Health and Human Services http://www.hhs.gov/
HIE/HIEN	Health Information Exchange/Health Information Exchange Network
HITSP	Health Information Technology Standards Panel http://www.hitsp.org/

HL7	Health Level Seven http://www.hl7.org/
IHE	Integrating the Healthcare Enterprise http://www.ihe.net/
IIS	Immunization Information System
MIROW	Modeling of Immunization Registry Operations Workgroup http://www.immregistries.org/pubs/mirow.phtml
MPI	Master Patient Index/Master Person Index
NCQA	National Committee for Quality Assurance http://www.ncqa.org/
NHIN	Nationwide Health Information Network http://www.hhs.gov/healthit/healthnetwork/background/
NVAC	National Vaccine Advisory Council
ONC	Office of the National Coordinator for Health Information Technology http://www.hhs.gov/healthit/onc/mission/
PDA	Personal Digital Assistant
PHIN	CDC's Public Health Information Network http://www.cdc.gov/phin/
PHR-S	Personal Health Record System
POD	Point(s) of Dispensing
RHIO	Regional Health Information Organization
SDO	Standards Development Organization
SNS	Strategic National Stockpile http://emergency.cdc.gov/stockpile/index.asp
SOA	Service-oriented Architecture http://www.webservices.org/
SSH	Secure SHell
SSL	Secure Sockets Layer

UPIF	Universal Provider Interface Format (NYC) http://www.nyc.gov/html/doh/downloads/pdf/cir/upif110_102507.pdf
VACMAN	Vaccine Management System http://www.cdc.gov/vaccines/programs/vacman/default.htm
VFC	Vaccines for Children http://www.cdc.gov/vaccines/programs/vfc/default.htm
WIC	US Department of Agriculture's Women, Infants and Children Program http://www.fns.usda.gov/wic/
WWW	World Wide Web http://www.w3.org/2002/ws/
X12	Accredited Standards Committee X12 http://www.x12.org/
XML	eXtensible Markup Language

Appendix B: Selected Sources

Organization	Document URL
AIRA	Immunization Information Systems Codebook http://www.immregistries.org/docs/IIS_Data_Codebook_101607.xls
	MIROW Chapter #2: Vaccine Level Deduplication in Immunization Information Systems http://www.immregistries.org/pdf/AIRA_BP_guide_Vaccine _DeDup_120706.pdf
	MIROW Chapter #3: Data Quality Assurance in Immunization Information Systems: Incoming Data http://www.immregistries.org/pdf/AIRA_MIROW_Chap3_DQA_02112008.pdf
CDC	Implementation Guide for Immunization Data Transaction using Version 2.3.1 of the HL7 Standard Protocol http://www.cdc.gov/vaccines/programs/iis/stds/downloads/hl7guide.pdf
	Core Data Set http://www.cdc.gov/vaccines/programs/iis/stds/coredata.htm
HHS ONC	Immunizations & Response Management Use Case See http://www.hhs.gov/healthit/usecases/respmgmt.html
HL7	Coming to Terms: Scoping Interoperability for Health Care http://www.hln.com/assets/pdf/Coming-to-Terms-February-2007.pdf
IHS	Working Together on Data Exchange: A Guide to IHS and SIIS Interfaces http://www.immregistries.org/pdf/IHS_Interface_Guide.pdf
NYC	CIR Data Exchange Interface (DEI) Specification http://www.nyc.gov/html/doh/downloads/pdf/cir/DEIspecifications_113006.pdf
	CIR UPIF Provider's Guide http://www.nyc.gov/html/doh/downloads/pdf/cir/upif110_102507.pdf
	CIR HL7 Web Service Integration Guide for Immunization Transactions
	CIR HL7 Web Service Technical Design
PHII	The HOWs and WHYs of Developing a Health Level 7 (HL7) Implementation Guide http://www.phii.org/HL7ImplementationGuide.htm
	Unique Records Portfolio http://www.phii.org/resources/UniqueRecordsPortfolio.asp
RI	KIDSNET Immunization Export File Specification http://www.health.ri.gov/family/kidsnet/KIDSNETImmunizationOutputFile.pdf
	KIDSNET HL7 Implementation Guide for Immunization Transactions http://www.health.ri.gov/family/kidsnet/dataexchange.php