Solicitation:

ONCHIT-3

Developing a prototype for a nationwide health information network architecture

Award Type:

Submittal for Small Business Set-Aside

Open source implementation

Offeror:

Records For Living, Inc.
34 Church Street
Sudbury, MA 01776
(978) 460-0539
info@RecordsForLiving.com

Persons authorized to negotiate on the offeror's behalf and to sign the proposal:

Simone L. Pringle
President and Founder
(978) 460-0539
simone@RecordsForLiving.com

This page intentionally left blank.

Executive Summary

RFP ONCHIT-3 requests proposals to "develop and evaluate prototypes for a NHIN architecture maximizing the use of existing resources such as the Internet to achieve widespread interoperability among health care software applications, particularly EHRs. Another key objective is to spur technical innovation for nationwide sharing of health information in patient care and public health settings."

Competitive market opportunities foster innovation. Our Proposal specifies a NHIN architecture that encourages innovation, creativity, and technical experimentation by delegating policy-making responsibility to the lowest levels possible. Our architecture supports national standards, with localized control. Information will be securely and privately accessible across the entire NHIN, and yet the architecture will put as much responsibility and control as is possible within the RHIOs.

The proposed architecture also supports further delegation within the RHIOs to sub-RHIOs, certified provider systems, and in the future, to consumers themselves (through patient portals or similar mechanisms).

The architecture's ability to delegate to lower layers in the policy hierarchy will support localized, market-driven solutions. Providers and vendors will have the freedom to explore solutions that best meet the needs of their patients, taking into account, factors such as geography and cost considerations. Individualized, competitive, innovative "local" solutions will roll-up to the national level, in a manner controlled by the national standards.

This proposal is based on widely used networking tools and standards, such as SOA, XML Messaging, LDAP directory services; as well as security and caching techniques and solutions. All software produced as part of this effort is to be made available as an open source implementation.

This page intentionally left blank.

TABLE OF CONTENTS

This page intentionally left blank.

Architecture

I. Overview

The June 2005 Summary of Nationwide Health Information Network (NHIN) Request for Information (RFI) Responses report (http://www.hhs.gov/healthit/rfisummaryreport.pdf), states that,

> Among the many opinions expressed, the following concepts emerged from the majority of RFI respondents:
>
> - A NHIN should be a decentralized architecture built using the Internet linked by uniform communications and a software framework of open standards and policies.
> - A NHIN should reflect the interests of all stakeholders and be a joint public/private effort.
> - A NHIN should be patient-centric with sufficient safeguards to protect the privacy of personal health information.
>
> […]
>
> - Key challenges will be the need for additional and better-refined standards; addressing privacy concerns; paying for the development and operation of, and access to the NHIN; accurately matching patients; and […]

The concept of RHIOs has been defined in the July 21, 2004 Framework for Strategic Action, The Decade of Health Information Technology: Delivering Consumer-centric and Information-rich Health Care (http://www.hhs.gov/healthit/documents/hitframework.pdf) as,

> The development, implementation, and application of secure health information exchange across care settings requires a local leadership, oversight, fiduciary responsibility, and governance. These regional health information organizations (RHIOs) are critical to health information exchange that reflects the health care priorities of a local area as well as the legitimacy and trustworthiness of this activity to clinicians and consumers.
>
> […]
>
> To create a more permanent and accountable infrastructure to support health information exchange, there is a need for a common approach to the formation and operation of RHIOs. The government could help define a common set of practices by incorporating minimal performance requirements into its contracts with, or grants to, communities. Another approach, commonly used in health care, is private sector accreditation to ensure that these organizations meet minimal standards.

The Framework for Strategic Action further states,

A national health information network that can provide low-cost and secure data movement is needed, along with a public-private oversight or management function to ensure adherence to public policy objectives.

Such a technology should be nonproprietary, available for broad use, and shared within the public domain in a manner that is available to all. It should be integrated with public health surveillance and response in accordance with existing statutory provisions, and deployed and operated in a secure, HIPAA-compliant and decentralized manner.

The nationwide health information network (NHIN) will be based on widely accepted networking principles, and widely used internet networking technologies. The overall architecture will be multi-layered and distributed, in order to provide scalability and robustness. Security, auditability, reliability, cost, and flexibility for future growth will be crucial quality attributes of any acceptable technical architecture.

II. Problem segmentation

The problem of organizing the communications within the NHIN can be broken into largely orthogonal domains of

- person identification service,
- authentication,
- roles management,
- access control,
- patient de-identification,
- data location,
- data storage and retrieval,
- data meaning.

For each of these problem domains, we will attempt to specify only enough information about what the RHIOs will do in order to provide for a coherent overall design, and in order to provide the ability for cross-RHIO communication and function. At every step – care will be taken to place as little specific *requirement* on *how* these areas are addressed – so that each RHIO has freedom to innovate. Yet, at every step, we will specify suggested ways of addressing these areas to illustrate how the overall system might operate.

The principle role the government (or any NHIN governing body) should have is to define the high level requirements for what happens within a RHIO, and to define quality metrics, and perhaps a scheme for evaluating RHIOs on those quality metrics (for example, responsiveness, robustness, ease of use, breadth of support for various features, etc).

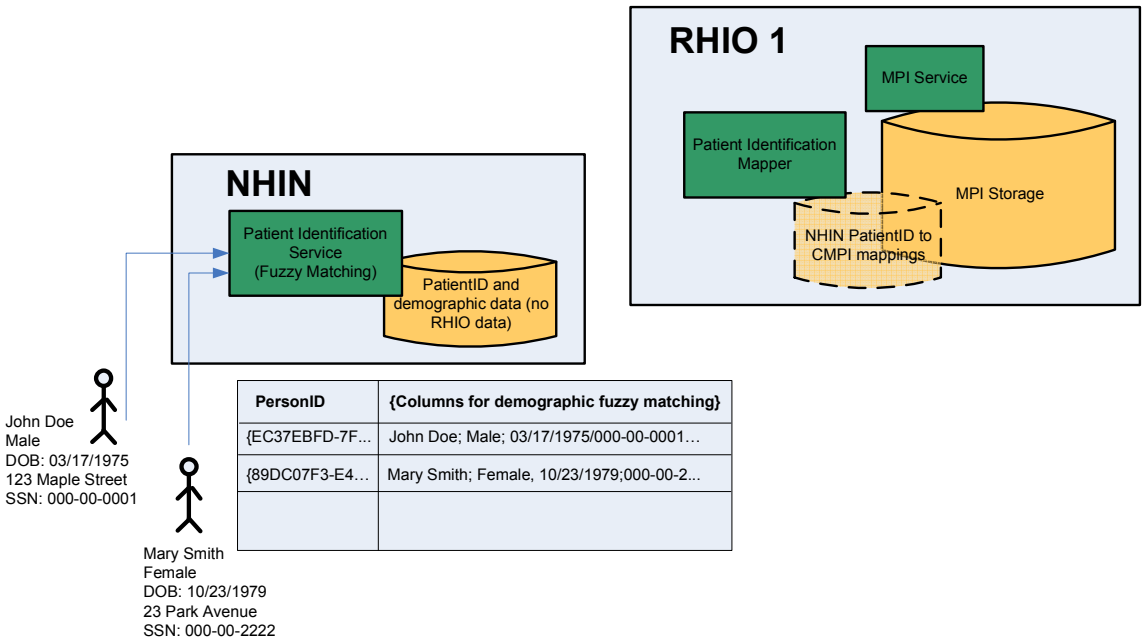
a) Person Identification Service

Some mechanism needs to be devised, to uniquely identify a person across the entire NHIN system, even if his/her medical records cross RHIO boundaries.

We propose a system whereby each person in the NHIN will be identified by a globally unique identifier – referred to as a *PersonID*. (Note that the proposed PersonID system would apply equally to patients and care providers such as doctors, nurses, technicians, etc, and would otherwise be similar in concept to the internal ID used in MPI systems, and would integrate neatly with existing MPI-type systems).

A NHIN-level service will be required to map between personally identifying information (such as name, possibly address, place of birth, birth date, or social security number) – and their unique ID. The service must be able to accept partial specification of these related attributes and provide fuzzy matching to produce a rank ordered list of matches.

An LDAP server, coupled with name and address standardization tools would be a possible implementation strategy for this service. Another possible basis for implementing this would be to leverage some existing master patient indexing (MPI) system.



The diagram above illustrates a patient identification implementation that leverages already in-place MPI systems to track NHIN global PatientIDs, by simply adding the NHIN PatientID to the list of mapped identifiers.

This database could be populated in a number of different ways. The database could be initially seeded with data from government databases, such as social security data, birth registries, state school registration data, medical data, and so on. In any case, RHIOs would need the ability to add new person records (in case they encountered people who were not in the system – for example – foreign nationals, newborns, etc…). This suggests an alternate population mechanism (which might be more costly in overhead, but also more politically acceptable) – which is a bottom-up, opt-in-based population of the mapping table just when people opt-in to having their data stored in their RHIO (and the RHIO becomes part of the NHIN).

Access to this person identification service will require appropriate access control provisions to assure privacy. (See the section on privacy concerns for discussions of the implications of this service; also see authentication).

b) Authentication

When someone (a person or a process) accesses data, for security reasons, it will be necessary to know who they are. For service-based queries (such as enterprise-wide management or monitoring tools) digital certificates and trust relationships can be used to establish identity.

For individual access, since there are a wide variety of different authentication mechanisms one might want, each with different strengths and weaknesses, and since technology over the next decade will likely provide many more (and a changing landscape of relative advantages and disadvantages) – it is important that any scheme chosen by a national health information network be quick and easy to change (i.e. it should be flexible and adaptable).

Still – we clearly want an authentication to be usable for access across the entire NHIN (single sign-on or SSO), for convenience sake.

We propose that the NHIN establish standards for authentication and the actual maintenance of the authentication services be delegated to third parties. A set of authentication providers that met specific service level agreements (SLAs) – would have those SLAs published with the NHIN. Any SLA would include at least include minimal technical and management criteria defined by the NHIN (including emergency authentication, lost password and problem management, etc), but – since the details of the SLAs are published with the NHIN – any participating RHIO could examine the agreement, and decide on the level of 'trust' they wish to confer to entities authenticated through that authentication provider.

These authentication providers would collaborate with the NHIN to produce a globally (across the NHIN) usable authentication ticket. This technique is like that used in

Kerberos, and might possibly be implemented using SAML (http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security) – or at least some very similar technique.

These tickets would allow participating systems (e.g. RHIOs) to

- identify the person logged in,
- identify the authentication technology used,
- and know when the ticket expires.

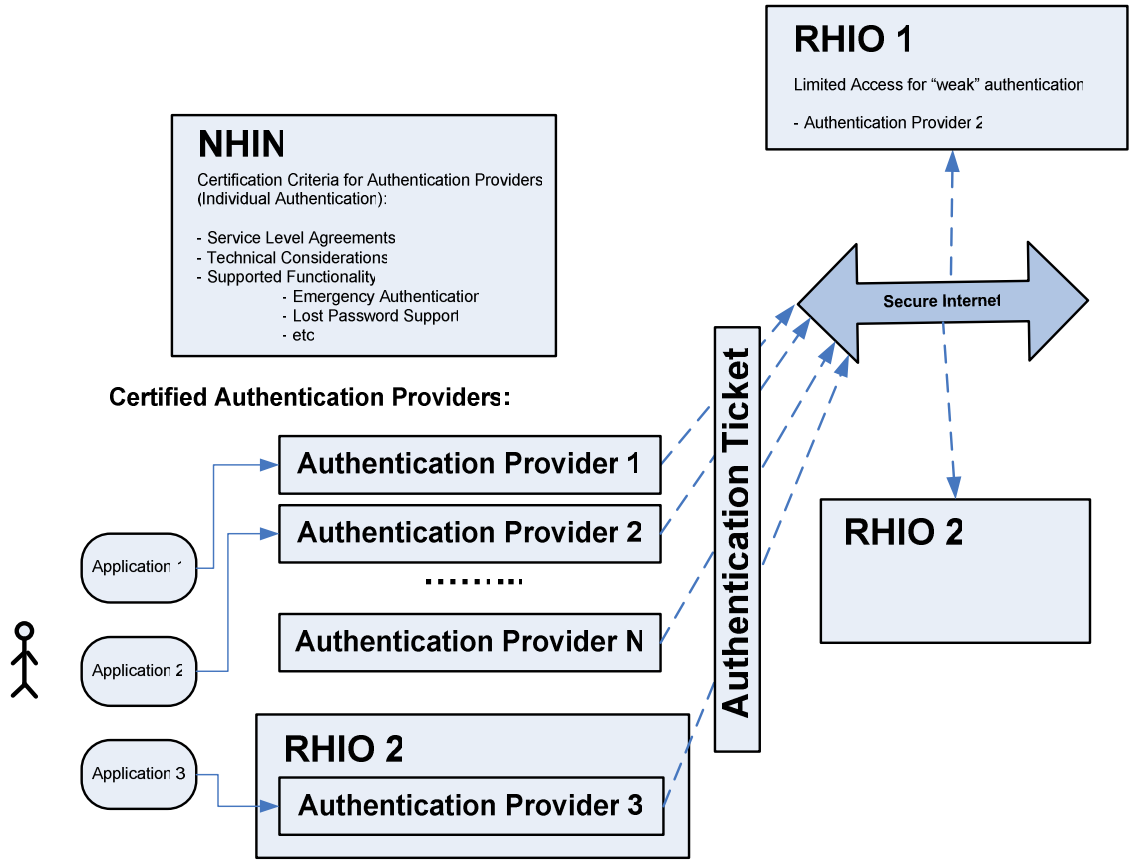
Individual RHIOs (and by delegation – individual providers or members inside RHIOs) could reject particular forms of authentication as a policy – but it is anticipated that this would be rare. A more likely scenario is that individual RHIOs might have stricter access control rules for authentication schemes that seem relatively insecure (in other words – a RHIO would be permitted to only allow *limited* data access if a requestor used a weak form of authentication).

This sort of system would allow for the important requirement of 'single sign-on' – SSO – while still allowing the delegation of most responsibilities for designing and maintaining authentication schemes to third-party businesses. It would further delegate to consumers and RHIOs the choices about which schemes they want to use and allow – making the choice about security vs. convenience vs. cost.

The NHIN would create the standards; the free market would provide the solutions; and the consumers would make the choices.

The diagram below illustrates users having access to multiple applications (web pages, EMR applications, etc). These applications use one of the NHIN-certified authentication providers.

By successfully authenticating via one of the certified authentication providers, the user is issued an authentication ticket, as described above. Note that in this diagram, RHIO 1 does not consider authentication services provided by "provider 2" to be adequate for certain types of access. Users who are authenticated utilizing provider 2 will have limited access to information originating from RHIO1.



Finally, it is worthy of mention that a RHIO might choose to be NHIN certified as supporting authentication services, though over time, RHIOs will likely choose to use existing, certified authentication providers, due to cost and maintenance considerations.

By utilizing a federated authentication system, the NHIN can be established on top of existing infrastructure, allow and encourage that infrastructure to grow and evolve, and still achieve the required harmonization and integration required for secure and private national information access.

c) Roles Management

Many users of the NHIN system will be ordinary consumers. Others will be doctors, healthcare administrators, emergency care professionals, nurses, pharmacists, medical researchers, and so on. The system is designed to support individuals having multiple roles.

A directory service mapping Person IDs to a set of roles is provided for use in the access-control system. A NHIN-based registration service (manned by humans and human-generated policies for each role) – will populate and manage this directory service. Some of the management can be automated by directly populating this directory service with data from other registrations (such as the AMA, or state medical registries).

While this roles management system *could* be done locally, within each RHIO, doing so would cause confusion in the access control systems managed at the RHIO level (since fragmentation of role names/meanings across the NHIN would make defining access rules difficult). Still, since most use of roles would be within the RHIOs access control system, it could – as an implementation detail – define its own private (to that RHIO) roles.

d) Access Control

Though authentication needs to be global – across the entire NHIN – access control does not. Each individual RHIO can design its own access policies, fostering experimentation, and competition to provide the best solutions. It *could* even define its own local roles (groups of users with common access privileges) to augment the national roles standards and definitions (at least on an experimental basis).

There is value in the simplicity of a common access control standard, and the harmonization process in ONCHIT-1 might develop such a single standard, but it might not, and need not. Even if standards for access control do evolve – within the standards – there is room for extension and variation.

Conceptually, access control will grant or deny access to data location or data storage or retrieval activities. Since exact policies will be determined at the RHIO level, they could in turn be delegated, in part, to member institutions (such as doctors' offices or hospitals). Access control systems would probably, at least initially, be simply limited to basic rights such as view/add/modify/delete to records of a given type within the system (see data meaning for more details).

Access control policies also govern what logging and audit trail system will be in place for data access, and modification. A typical policy might be to log all access and modification requests.

Note that the NHIN would need to implement some sort of access-control regimen itself – but only to manage access to the part of the data location service where a query is made about which RHIOs contain information about a given Patient-ID, and for access to the person identification service.

e) Patient De-Identification

Many services, such as biosurveillance, and medical research, will profit from access to people's health records. Still, substantial security and privacy concerns abound with regard to this kind of heath record access (see privacy).

We propose dealing with this issue by defining virtual patient records (in essence – 'limited database views') whose elements are dynamically synthesized by the RHIO's data location and data retrieval systems to create de-identified information which can then be accessed by authorized personnel and services.

A *key* to this process is that – though the information contained in these virtual records will be specified globally (at the NHIN level) – their access control will be defined at the RHIO level – assuring appropriate levels of patient privacy. Clearly, the information contained in the virtual records would be a subset of the total medical record information, specifically designed to be non-patient identifying.

The process of synthesizing de-identified health records (within each RHIO or its designee) might leverage pre-existing patient-data de-identification tools and systems, but more likely would follow trivially form carefully defined virtual summary information records.

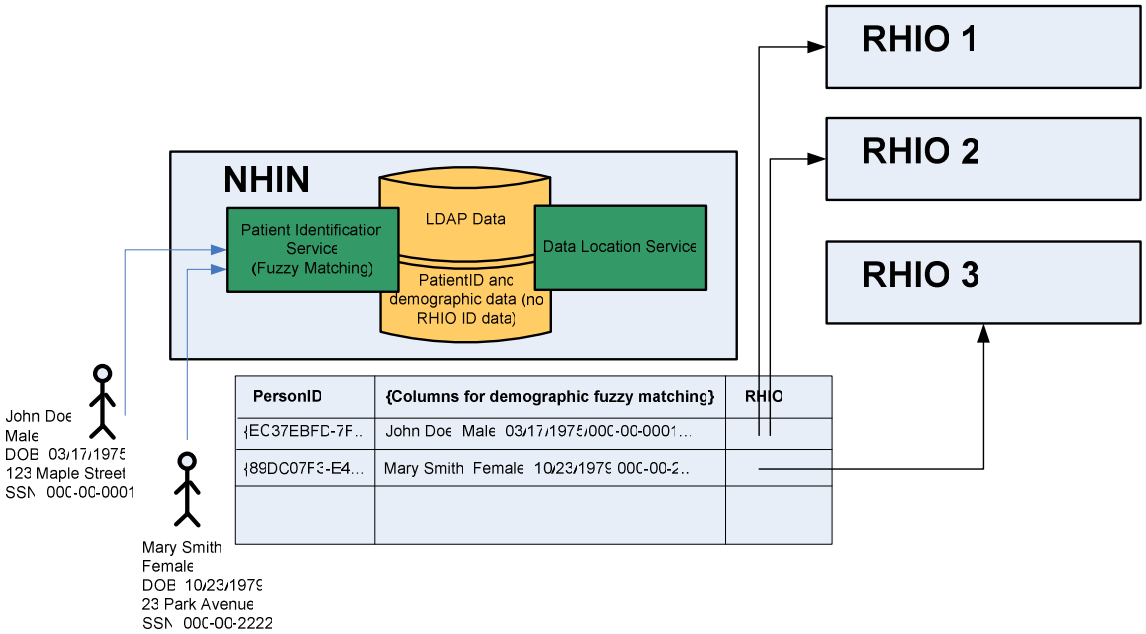
f) Data Location

Patients' data can be distributed among several RHIOs. Additionally, certain patient records can be quite large, with information gathered at different dates and of different types. There needs to be a mechanism that can locate the entirety of that information.

A patient's data is a collection of *records*- whose type information is specified in the data meaning section, and each record has a globally unique Record-ID (as described in more detail in the data storage and retrieval section). Examples of records might include visits, lab results, referrals, etc.

We propose a hierarchical system of LDAP servers, using delegation. There would be a single NHIN LDAP server (replicated for redundancy sake) – which would contain records for each PersonID – saying what RHIOs contain records for that person ID. It is the sole responsibility of each RHIO to add and remove LDAP records from the root

NHIN LDAP server if and when the RHIO transitions from a state of having no records to and from a state of having any records, for a given individual. That is to say – updates would be extremely infrequent.



A query to find data for a given patient would start by asking the global NHIN LDAP server for all RHIOs that contain records about the given patient. Then, each RHIO with that patient's data would be queried (by the requesting client).

The form of a query to a RHIO would be an XML-formatted message, with search / selection criteria (to be specified). Queries would result in either an error condition, or a list of RecordIDs. Searches would be subject to access control restrictions.

Data Location RHIO Implementation Options

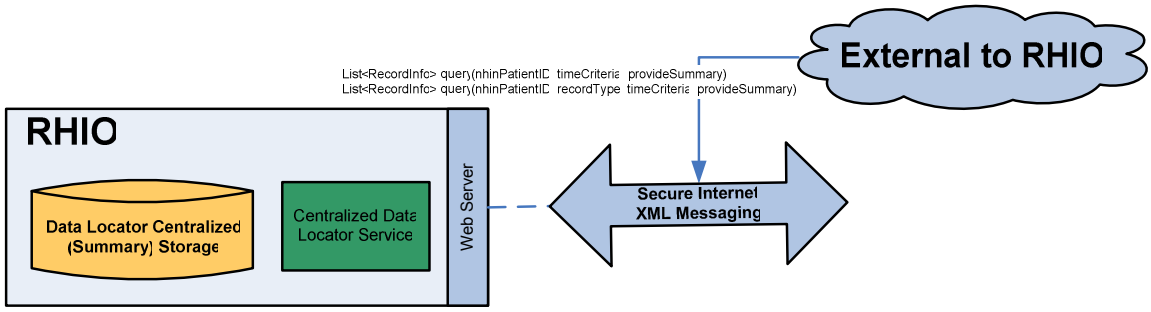
Though a small and simple piece of the data location service is implemented in the NHIN, each RHIO would be responsible for the far more substantial part of the data location service implementation. This RHIO piece can be implemented and configured to fit the constraints (financial, technical, organizational, access to connectivity) of RHIO participants.

The options listed in this section do not represent an exhaustive sampling of alternatives, but merely serve to illustrate the flexibility supported by the proposed architecture. Additionally, since other data related services are discussed later in the document, these diagrams are limited to the data location service.

All diagrams capture the flow of record location information to a NHIN entity that is external to the RHIO (represented by the 'cloud'). All communications occur with XML messaging. In all scenarios below, the NHIN entity, external to the RHIO contacts the centralized RHIO service (via RHIO-specified url). The external NHIN entity *never* directly contacts the RHIO participants' federated services. The details of the implementation within each RHIO is hidden from external entities.

Pseudo interface specification code is used to represent the system's ability to query for record information (optionally containing summary in addition to IDs) based on patient identification, specification of interested time period and/or record type.

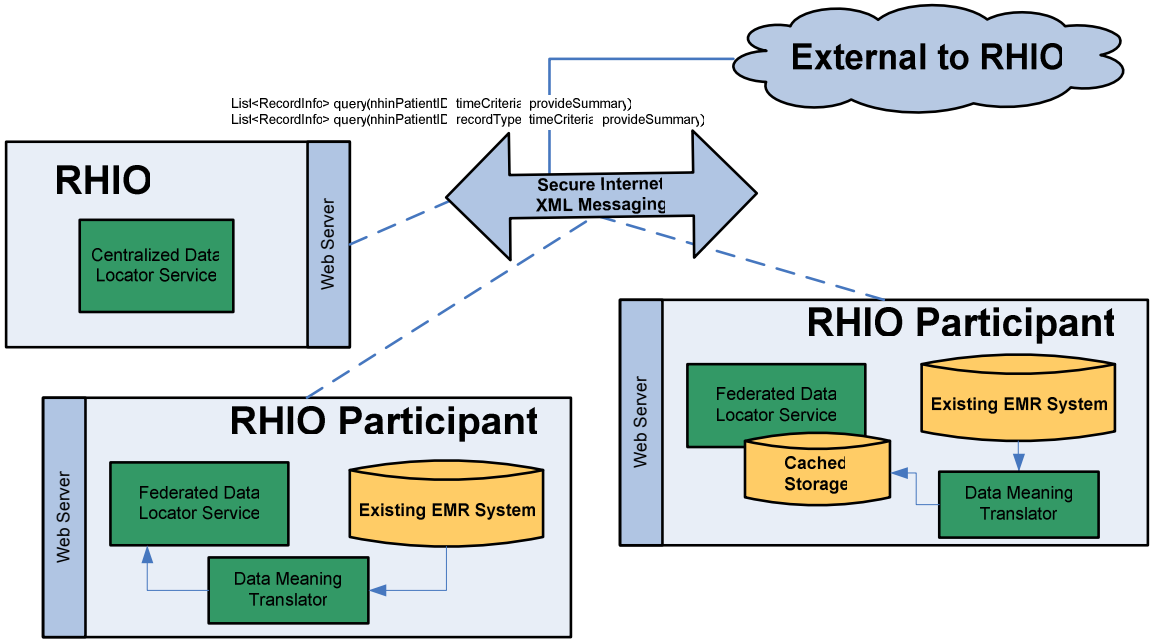
"RHIO Participants" as shown in some of the diagrams below represent organizations (providers, portals, etc) that engage in the exchange of healthcare-related information within a particular RHIO, and abide by the local RHIO standards and policies.



Option 1 RHIO centralizes data locator service implementation

One implementation option for the data locator service would be to capture centralized record summary information for the entire RHIO. Note that RHIO participants are not shown in this option, since the transmittal of information to the centralized location occurs out of band with the transmittal of information back to the requestor.

The Indiana Network for Patient Care RHIO implementation (discussed later in this proposal) supports this type of centralized implementation.



Option 2 Centralized RHIO data locator service delegates to federated services
RHIO participants can optionally cache just-in-time mapped/translated records

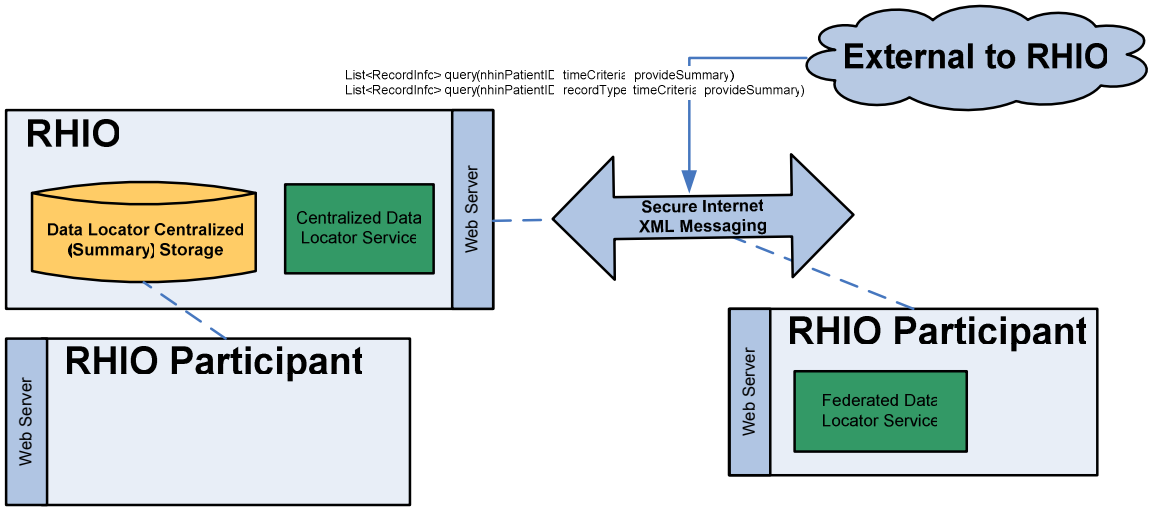
In an alternative implementation, there is a centralized data locator service, which delegates the job of retrieving recordIDs and summary information to RHIO participants. The centralized data locator is then responsible to aggregate the collected information, prior to replying to the requestor.

A couple of options for RHIO Participant implementation are shown in the diagram above. (Note again, that these options are not exhaustive either, of the possible implementations at the RHIO Participant level.)

The data meaning translator is better understood by reading the data meaning section below. It is sufficient at this point to understand that in order to provide summary information, some level of data translation (transformation, conversion) is typically required.

The RHIO Participant option on the left illustrates on the fly retrieval and translation of RecordIDs and summary information. The Santa Barbara County Care Data Exchange supports a model similar to this. Note that we believe implementations may be underway that leverage caching.

The RHIO Participant option on the right illustrates also on the fly retrieval, with the ability to cache previously retrieved information.



Option 3 Hybrid approaches designed to support technical and/or financial constraints on RHIO Participant's infrastructures

The final option shown for this service illustrates a hybrid approach where some RHIO Participants implement federated data locator services, while others simply leverage the centralized storage location.

This is a useful approach for RHIO communities that have small medical practices that do not have access to IT staff. Providers in those practices might access a web site, through which clinical activity is entered or updated, but is ultimately stored in a centralized location. Alternatively, periodic bulk data loads could transfer data from RHIO participants to the centralized location. This would imply that members from these RHIO participants would not have real-time access to their medical data. Whether this is acceptable or not, depends on service-level agreements within the RHIO.

Details of federated data locator services are not shown in this picture, but it is understood that those RHIO participants have the means to store and retrieve their own clinical data.

g) Data Storage and Retrieval

One of the most basic functions of the NHIN architecture is to actually store and retrieve medical and health related data. The storage and retrieval of data is all governed by access control policies (and auditing policies) defined earlier.

The structure of data and types of data are defined in the data meaning section.

Data retrieval, update and additions are done by record ID (*RecordID*). A record ID is a globally unique ID that may be found using the data location service. A RecordID would be a persistent moniker (name that could be re-used days, weeks, or even years later) – to

uniquely refer to a given medical record (that ID, coupled with an identifier for the given RHIO would constitute a URL as specified in http://www.w3.org/Addressing/).

It is not required that data objects be versioned, but contain a last-modified date stamp attribute (versioning maybe a good implementation strategy, and might make sense for a future extension of this specification). Data objects are defined to exist in a single RHIO. This is a naming and ownership policy – as other RHIOs can have cached copies of the record. The record may or may not physically reside in RHIO computers; it maybe stored in a member organization system, and the RHIO provides its own mapping from that RHIO record ID to the internal physical location.

Storage and retrieval of data will be done through a simple XML-messaging API, secured via SSL. Access (storage or retrieval) requests will contain a Kerberos-like ticket ID (a temporally expiring value generated from the NHIN's authentication service), the record ID, and optionally (for writes) the data to be written. It will return a success status code (and with successful reads – the actual data).

Each RHIO will provide a single web URL for data access and retrieval operations. The XML messages contain the ID information and the kind of operation (update/read/etc, and the auxiliary data if any).

For reasons of performance, robustness, and scalability, it is anticipated that RHIOs will likely use any of a number of common web techniques to load-balance and replicate data across redundant servers, but the details of how this is done will *not* be specified (an implementation detail left to each RHIO administrator).

Though not initially required, it is strongly recommended (and should eventually be required) – that all data updates be fully logged and reversible (in case of erroneous updates). A possible way to satisfy this requirement would be to store data in a transactional database, and to preserve the transaction log.

Data Storage and Retrieval Options

The data storage and retrieval service can be implemented and configured to fit the constraints (financial, technical, organizational, access to connectivity) of RHIO participants.

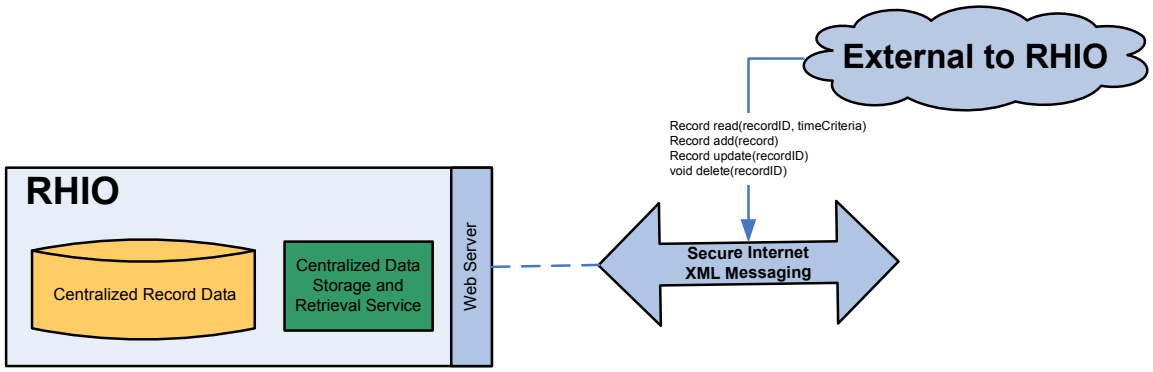
The options listed in this section do not represent an exhaustive sampling of alternatives, but merely serve to illustrate the flexibility supported by the proposed architecture.

All diagrams capture the flow of record location information to a NHIN entity that is external to the RHIO (represented by the 'cloud'). All communications occur with XML messaging.

Pseudo interface specification code is used to represent the system's ability to create, read, update, and delete records in the RHIO.

"RHIO Participants" as shown in some of the diagrams below represent organizations (providers, portals, etc) that engage in the exchange of healthcare-related information within a particular RHIO, and abide by the local RHIO standards and policies.

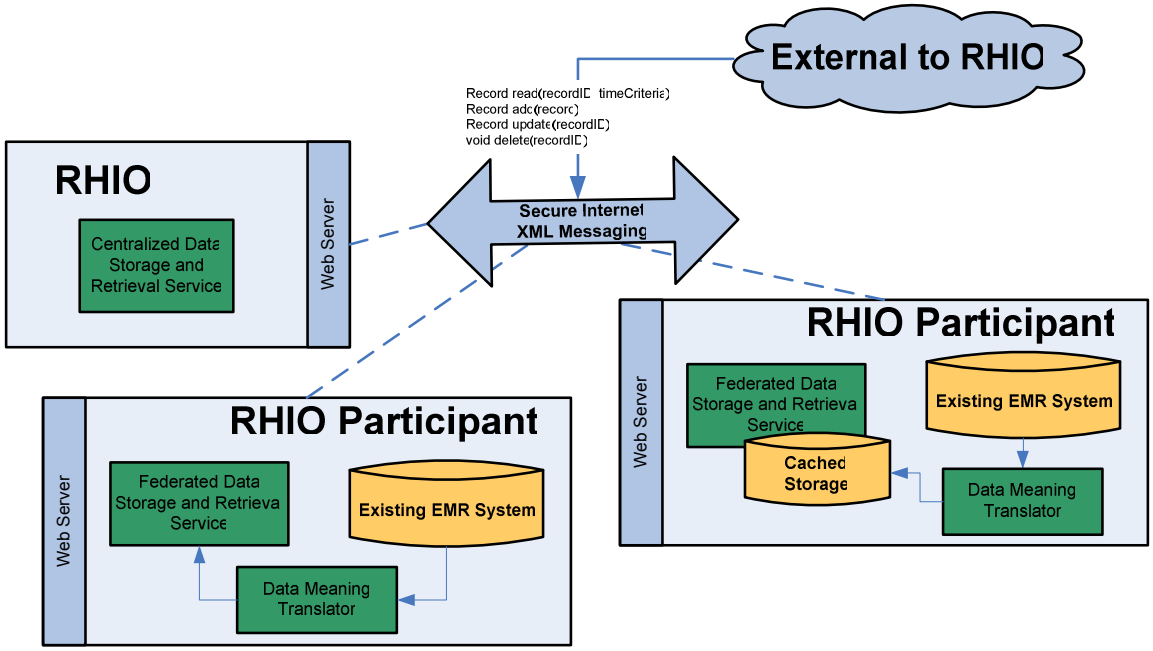
The first few options illustrated in this section will focus on the data storage and retrieval service, without including the data location service. Later in this section we will reintroduce the data location service to further illustrate flexibilities introduced by this architecture.



Option 1: RHIO centralizes data storage and retrieval service implementation

One implementation option for the data storage and retrieval service would be to capture centralized information for the entire RHIO. Note that RHIO participants are not shown in this option, since the transmittal of information to the centralized location occurs out of band with the transmittal of information back to the requestor.

This picture is drawn in such a way as to imply that the format of the centralized record data is either based on the standard data protocol for the NHIN, or it can be interpreted to imply that the centralized data storage and retrieval service is converting and transforming data on the fly.



Option 2 Centralized RHIO data storage and retrieval service delegates to federated services RHIO participants can optionally cache just-in-time mapped translated records

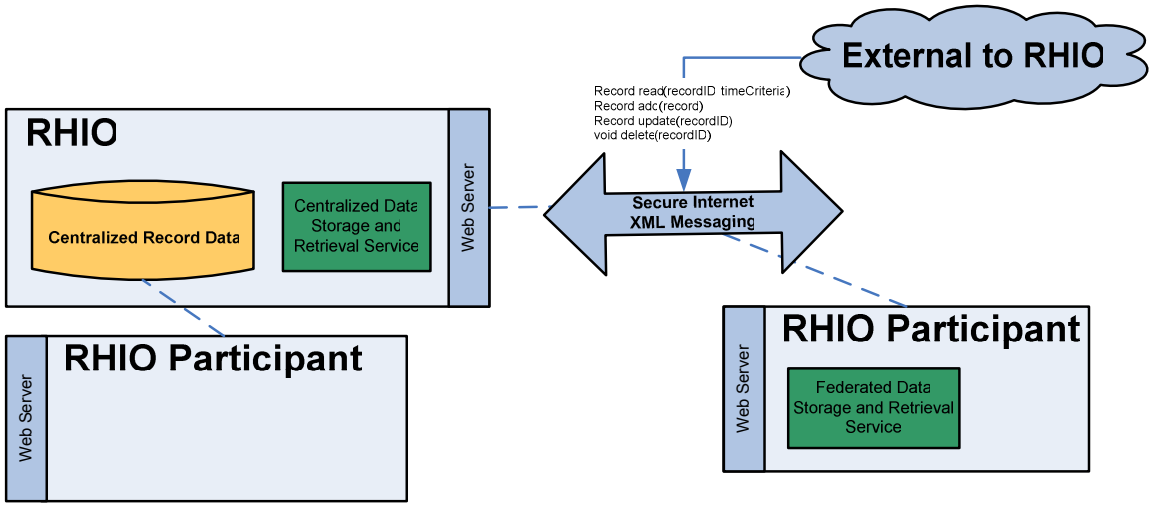
In an alternative implementation, there is a centralized data storage and retrieval service, which delegates the job of retrieving information to RHIO participants. The centralized service is then responsible for aggregating collected information, prior to replying to the requestor, and/or dispatching requests as appropriate.

A couple of options for RHIO Participant implementation are shown in the diagram above. (Note again, that these options are not exhaustive of the possible implementations at the RHIO Participant level.)

The data meaning translator is better understood by reading the data meaning section below. It is sufficient at this point to understand that in order to return compliant data to the NHIN, some level of data translation (transformation, conversion) is typically required.

The RHIO Participant option on the left illustrates on the fly retrieval and translation of RecordIDs and summary information.

The RHIO Participant option on the right illustrates also on the fly retrieval, with the ability to cache previously retrieved information.

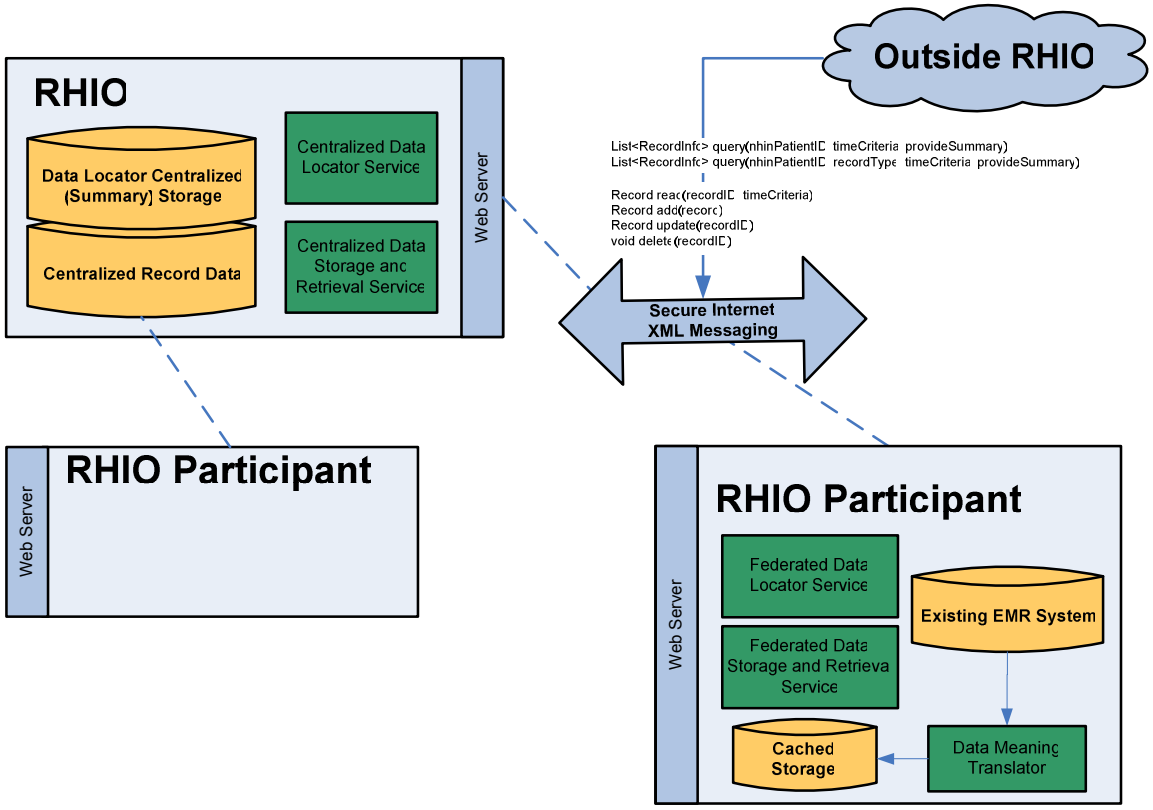


Option 3 Hybrid approaches designed to support technical and/or financial constraints on RHIO Participant's infrastructures

This option illustrates a hybrid approach where some RHIO Participants implement federated data storage and retrieval services, while others simply leverage the centralized storage location.

This is a useful approach for RHIO communities that have small medical practices that do not have access to IT staff. Providers in those practices might access a web site, through which clinical activity is entered or updated, but is ultimately stored in a centralized location. Alternatively, periodic bulk data loads could transfer data from RHIO participants to the centralized location. This would imply that members from these RHIO participants would not have real-time access to their medical data. Whether this is acceptable or not, depends on service-level agreements within the RHIO.

Details of federated data storage and retrieval services are not shown in this picture, but it is understood that those RHIO participants have the means to store and retrieve their own clinical data.



One final aspect to consider is that there are options still available to RHIO designers and implementers that support storage configuration in support of the data location as well as the data storage and retrieval services.

Storage repositories can be shared between the services or not (for instance, summary information may be tracked in a centralized location for a given RHIO participant, but the same RHIO participant may choose to retain the full record information, in a federated storage).

Similar hybrid options exist at the RHIO participant implementation, where cache information may be supported for summary but not detail-level information.

h) Data Meaning

This is one of the most long-term challenging aspects of the entire NHIN project. There are numerous efforts at creating standards, or specifications of what data might be captured in a medical record, health record, test result, and so on. These include:

- HL7 RIM (http://www.hl7.org/Library/data-model/RIM/modelpage_mem.htm)
- PING (http://ping.chip.org/)
- Records For Living's HealthFrame schema (http://www.recordsforliving.com/Schemas/)

as well as external data specifications uses by the dozens of other EMR and PHR vendors, and many more standardization efforts.

How we approach this problem of representing medical transactions will be strongly driven by the standardizations harmonization process (ONCHIT1) – harmonizing these varied and alternative approaches.

As for code-sets and vocabularies – we are strongly inclined to leverage the UMLS (http://www.nlm.nih.gov/research/umls/) system as a way to unify the various code systems like LOINC, ICD-9, CPT, etc. Again, the ONCHIT1 process will take a leading role in deciding this question.

Transforming

Since the process of harmonizing standards is not yet complete, and won't be for many years to come (at least not so that all implementations and actual data are harmonized with the standards), a process of **transforming** data between formats will be required. Briefly, the strategy we propose is as follows:

A set of data records will be defined (by an XML-schema) as the set of records which can be exchanged across the NHIN (or prototype RHIO). The smaller, and less redundant this set of record types can be made – the better, but depending on the technologies used already by the participants in the RHIO this may need to be an initially large set of record types. As it has already been covered – a major responsibility of the ONCHIT1 process will be to define a smaller, more rational set of record types, towards which the set in use will gradually migrate.

Only records of types designated as allowed will be transmitted over the network. As a governance matter, initially rules for adding types will be quite liberal and pragmatic, and be mostly limited to requiring the provision of XML-schema specification of the format. Over time, this set will diminish in size.

Then, when requests are received by a particular RHIO (or participating data storage system) for records of a particular type, the system will either already natively store records of that type, or know how to transform the records it has into records of that type (possibly using commonly available XML transformation tools such as XSL).

The net effect of this is that at any level of administrative authority – there will be a specified set of 'acceptable' record types, and the XML schema specifying the meaning of those types. Each participant will only converse with others on the system using those types – internally transforming their own data to meet those type specifications as needed.

i) Responsibility Summary

Responsibility in each of these areas must be placed somewhere. Conceptually, responsibility starts at the NHIN, and is then delegated downward to the level of the RHIO (and possibly therein even lower, to the consumer).

Sometimes – responsibility will be shared; but mostly primary responsibility can be clearly designated:

Domain	Responsibility	
	NHIN	RHIO
person identification	Primary	
authentication	Primary	
roles management	Primary	
access control	Minimal	Primary
patient de-identification		Primary
data location	Minimal	Primary
data storage & retrieval		Primary

Analysis

I. Security

It's important to understand the scope of NHIN security concerns as covered in this proposal. No system of medical record storage can withstand the obliteration of all of its underlying infrastructure, and this proposal makes no such attempts. If there is an interruption of electrical or internet access to a large swath of the country – that area will be with limited or no medical record access.

This system will work as well as – no better or no worse – than is typical for commercial banking or other types of systems.

a) Availability of Service

In keeping with the design of most of the rest of the Internet, the protocols here are simple, XML-based, and point-to-point, based on naming indirection (typically through DNS) – which allows replication of data transparently across many machines – even geographically separated machines.

This strategy means that a hacker attack which destroyed access to some machines would be less likely to prevent access to *all* machines.

b) Destruction of Data

The reason for requiring/suggesting reversible transaction logs in the Data Storage and Retrieval section was that so if and when a destructive worm or other process makes unwanted changes to health record data – these changes can be identified and reversed automatically (not magically – but using automated processes under the control of trained database professionals).

II. Performance

This proposal recommends the leveraging of conventional performance enhancing techniques used throughout the world-wide-web. These include caching, clustering, load balancing, and so on.

Because the protocols applied here are simple, XML-based, and point-to-point, based on naming indirection (typically through DNS), they can straight-forwardly leverage widely used web performance enhancing techniques. Services like Akamai's (http://www.akamai.com/) can be used to cache answers to important questions and to dynamically react to internet disruptions, and balance the load to servers (and cache information request results).

III. Audit Trail

To find and correct problems with the health infrastructure system, to correct human or machine error (whether deliberate or accidental), and to further illuminate and help understanding of the state of medical records, an audit trail of all important actions in the system is crucial. However – for *none* of these purposes is any standardization of the methods for creating, maintaining, or reviewing the audit trail strictly necessary (for some it might be nice).

Since we want to reduce costs, and increase the flexibility the implementers have in creating their RHIO systems, we impose no specific guidelines on how an audit trail is maintained. It's likely that each implementation system will have its own provided (such as a database transaction log).

Future versions of the NHIN protocol and specifications could be extended to allow for basic reporting (rolling up) of audit trail statistics and information.

IV. Privacy Concerns

a) Person Identification Service

Here – as elsewhere throughout the system – privacy concerns abound. There is significant, well-founded resistance among privacy advocates against utilizing a national ID system, or using social security numbers as a unique identification system.

One important item to note about *this* nation-wide ID system is that our proposal does not require that these IDs actually get published. They are used internally in transactions, and are *not* required to appear anywhere publicly, in reports, or ID cards.

It would be best if somehow these concerns could be assuaged – and a unique ID system created. Our proposal, however, will comfortably accept whatever consensus approach comes out the ONCHIT-1 harmonization process. Still – it should be understood that MPI systems provide no more privacy consideration (and in some ways less) – than our proposed person identification system.

b) Patient De-Identification Service

De-identification of patient data (for the purpose of biosurveillance) can be accomplished in compliance with HIPPA regulations (http://www.hipaadvisory.com/regs/finalprivacymod/deid.htm) either via commercially available HIPPA complaint tools, or via carefully designed virtual records (see patient de-identification). By delegating responsibility for access to these records to the individual RHIOs – privacy concerns can be double-checked. By defining the nature of

the virtual 'de-identified' records across the entire NHIN (data meaning) – cross-NHIN analyses maybe performed.

c) Biosurveillance

This is largely covered in Patient De-Identification Service and Privacy Concerns / Patient De-Identification Service.

The data needed to be collected for bio-surveillance would be defined as virtual records (data meaning). Then – participating RHIOs would generate these virtual records in response to any requests (following their own access control principles), and analyze the data.

d) Overall

The *key* privacy risk of the entire NHIN system is that any *effective and efficient* computer system to exchange information about people's health information is subject to weak-link security breaches. Any effective system *must* allow for rapid access to private information: you don't want people dying on the operating room table while the surgeon is typing the 90th character of his password for the 14th time.

There are lots of trade-offs to be made between privacy, and personal safety. We propose no answers as to how people should make this trade-off, because we think it's a largely personal and societal choice that should be allowed to evolve.

Instead – we propose a system whereby choices about this trade-off are clearly articulated and delegated downward from the national system to the RHIO level and – hopefully then - further to the level of the individual consumer. As much as possible, the individual consumer should be making the choices about privacy policies for their own personal information (either through their choice of providers, or eventually through slightly more fine-grained options providers might give their customers).

V. Adoption Path

An important feature of the entire NHIN / RHIO architecture which is respected, and even accentuated by this technical proposal, is that we anticipate a gradual migration path from the status quo to a world where medical records are regularly and easily shared.

Any organization which does not support these standards or mechanisms is effectively 'outside of the NHIN' for the purpose of records exchange. The actual *required* standards

for membership are kept exceedingly minimal, and so it should be easy for electronically enabled RHIOs to 'join' the NHIN (albeit perhaps with a low level of quality of service, depending on the quality of their adherence to the standards).

Each RHIO should be evaluated over time to measure how well it conforms to (as yet unspecified) quality metrics, and should be encouraged over time to improve its performance.

VI. Comparisons with Prior RHIO architectures

A number of different approaches have been and continue to be experimented with for the design and implementation of health information exchange networks.

A fair amount of differentiation exists in the approaches vis-à-vis patient identification and data policies.

This section is not designed to provide an exhaustive analysis of the various approaches, but rather, it should provide a comparative backdrop for our proposal.

Almost every state in the union has some effort underway to implement RHIO prototypes. We focus on a small subset of these efforts, including the *Indiana Network for Patient Care*, the *Santa Barbara County Care Data Exchange*, and the *Massachusetts eHealth Collaborative/MA-SHARE*.

a) Santa Barbara County Care Data Exchange (CDE)

The CDE is based on a peer-to-peer architecture that integrates federated, non-standardized *clinical data repositories* (CDRs) via a set of transactional interfaces. By design, no clinical records are stored at the CDE central repository.

Key architectural components of the CDE are the *identity correlation service* (ICS) and the *information locator service* (ILS).

The ICS leverages patient demographic information as input to matching algorithms. It correlates patient identities from different sources. The ICS is designed as a federated system and is used in patient searches (CDE-oogle.com).

The ILS links to patient clinical records in the participating systems, and is a multi-node design, with an ILS instance serving each CDR, in support of queries.

b) Indiana Network for Patient Care (INPC)

Patient identifiers from separate institutions were linked to a single global ID, using a series of progressively fuzzy matching algorithms (based on patient name, social security number, birth date, and gender).

A centrally managed global patient registry interconnects a patient's data from separate vaults, uses patient-specific virtual medical record – an aggregation of all content encoded by the participating sites for a given patient. All vaults share the same database structure and standardized terminology. Data is delivered via HL7 based messaging and is standardized prior to being stored in the institution-specific vault.

c) Massachusetts eHealth Collaborative/MA-SHARE

A key component of MA-SHARE is the *record locator service* (RLS). The RLS supports master patient indexing (MPI) with medical record pointers.

The *Clinical Data Exchange* (CDE) is a component of the architecture responsible for the maintenance and retrieval of medical history, lab information, imaging, etc. It fetches, formats and aggregates records and in effect acts as a "Google search" to multiple EHR systems.

d) Our Proposal

A key aspect of our NHIN Proposal is not so much how it bears similarities or contains differences with any of these prototype systems, but that our proposal is carefully crafted to delegate enough authority to RHIOs so that *any* of these approaches could easily integrate with a NHIN based on our architectural specification.

Each of these systems contains a slightly different way of identifying patients (different from each other, and different from ours). But in each case they know who their patients are and could use the system we propose to find corresponding PatientIDs within the NHIN, and they could gradually, at their discretion, move towards using those patient ID values internally.

Each of these systems uses a different way to locate patient information, but each has a mechanism. A translating service could be provided (within each of these RHIOs) to map between information location requests (i.e. data lookups) – and the systems they have for looking up data.

A more moderate challenge for each of these RHIO systems to integrate with our proposed NHIN would be the requirement to have globally unique 'record IDs' for each medical record (note in this context – a record is not of ALL the information for a particular patient, but rather something like a test result, or a patient encounter). Still – back-end implementation systems all of these are built with have this ability straight-

forwardly available – including the 'date last modified' requirement, and whatever infrastructure layer they have would likely be easy to adapt.

By far the most difficult requirement, and most difficult technical challenge, lies in the area of 'data meaning'. Whatever EHR / EMR systems they are using represent medical interactions and information a particular way. Though our proposal has a migration path facility to manage arbitrarily different and incompatible data representations – a system where these incompatible representations were not harmonized (through a transformation process) would be low quality and inefficient. In order to get a high-quality result, significant effort at harmonizing these data formats will be required (see ONCHIT1).

VII. Other Questions

a) Why not one RHIO per person?

Why not have a single RHIO be the one designated storage place for all data for a given person? One advantage this system would have is that it would be easier to find all the data for a given person (you wouldn't need to look across RHIOs). Another advantage is that it would provide an obvious point of consolidation if one wanted to rationalize conflicting information across medical records.

In summary, requirement of a sole RHIO per individual:

- Wouldn't support the guiding principles behind having RHIOs, which are in effect to support existing small networks of systems to share data they already have. Any time anyone added data about a patient – they would have to potentially add it to some remote RHIO system they didn't control. What if that remote system was unavailable? What if they natively stored information differently? What if conversions of data were not lossless?
- Would force the question, "How would you choose *which* RHIO to use?" If there was no choice (e.g. if based on HMO, or primary residence) – then the lack of patient choice would needlessly eliminate a positive motivator for better patient care (consumer choice fosters better quality and lower cost of offerings).
- Would force migrating data across system. Would all such transfers be lossless?
- Would really not simplify as much as it appears – since actual providers would then generally need to be located in different RHIOs than the patient's record itself.
- Would not be as big a win as some proponents might insinuate – since allowing the choice of what RHIO to reside in to be driven by the providers – still only means typical users would have data in a small handful of RHIOs (often just one).

b) Secure Email: Why not specify how secure email will be handled?

There already are internet standards for secure email (S-MIME - http://www.ietf.org/html.charters/smime-charter.html). This protocol is already supported in many commonly available email clients, and most internet email infrastructure, and there is really no connection to NHIN.

The principal requirement to support S-MIME is that people sending emails get a digital certificate which guarantees their identity and allows for encryption. We suggest people who wish to send secure electronic mail get a personal digital signing certificate from any of a number of certificate authorities (such as Verisign, or VisionShare).

Medical institutions, which wish to require all their personnel to have the ability to send secure email, can either contract with a certificate vendor to manage the process of assigning digital certificates to individuals, or they can create their own certificate authority (CA) – and manage it themselves (so long as it complies with the standards of some other higher level CA which will 'trust' it).

c) The PersonID system vs. existing MPI systems?

Why the proposed PersonID system, as opposed to using something more like existing MPI (master patient index) systems?

One key difference between the mandate of this proposal and the exigencies that motivated the existence of current MPI systems is that the NHIN proposal is for a nationwide health system where patients' medical records (and healthcare providers) can be easily identified unambiguously. This nationwide system would invite and require buy-in from participants, and would gradually be built up according to defined standards.

The existing MPI systems had no ability to prescribe what would be done with patient information outside those systems where it was used, and had no requirement that systems outside the domain of use of the MPI be able to unambiguously identify a patient or healthcare practitioner.

If existing MPI systems were extended to a national scale – they would essentially come quite close to the PersonID system we've proposed, with a negative exception: they would be filled with undetected mistakes, and have no mechanism to eliminate such future mistakes.

In our proposal and existing MPI systems, when individuals first present themselves to the system by name (and other commonly available identifying information) the system must generate a unique identifier for that person, and then use it. If that system worked *perfectly, always*, then the differences between our proposal and MPI systems would be inconsequential. MPI systems would share all the benefits and defects of our proposal, except that our proposal would be more efficient (since storage/communications would

be based on a short, structured identifier, rather than the string of information used to uniquely identify a person in an MPI system – which is essentially a longer, less structured PersonID).

But understanding the failure modes of the two systems is instructive.

When a difference between the MPI-generated PersonID and the PersonID in this proposal was detected – an attempt at correcting past mistakes in medical records can be attempted (much as in the VA's MPI based system). But in a system where the 'key' – the unique patient identifier is left unspecified (just based on whatever MPI system is being used) – then there is no way to go back and correct past medical records. It is not even clear that any of them are 'wrong'. They simply will frequently produce the 'wrong answer' when later queries are done to identify information about a particular patient.

Analyses such as the MA-SHARE http://www.mahealthdata.org/ma-share/projects/communitympi/20040416_UPIpaper.pdf erroneously conclude that, in general, national unique identifier systems are a bad idea. Their conclusions are based on reasoning that (mostly) does not apply to our proposal. In particular, they assert that a national patient identifier system "would require the establishment of an enumeration facility that could quickly address changes in the population base in Massachusetts". As was pointed out earlier – though our proposal *could leverage* existing databases of person information (such as social security numbers etc – which could be out of date) - it would not have to – and would certainly contain a facility to automatically add new national PatientIDs as-needed (instantaneously) via request from the NHIN person locator service. They further argue that France has had some difficulties along the way to implementing their system. It should be noted that the French system is in fact in place, and we should simply learn to avoid reproducing the same mistakes along the way. In particular, our proposal should have no issues of technological issues relating to ID enumeration. Technologically trivial support for GUID generation can be used in our proposal, since there is no required expectation of external use or human readability of these identifiers, in our proposal.

Counter analysis argues that having a national patient identification system would be a privacy risk. This argument is absolutely *specious*. First – any *effective* MPI system would be effectively a national ID system (just with longer, more human readable 'keys'). Besides – as everyone knows – pragmatically – we have already had a national ID system for quite some time (social security numbers). Yes – this does reduce our privacy – but not in any new ways. See the section above on privacy for more details.

d) What about consumers and personal health records?

How does this proposal affect consumers directly? What about personal health records?

A crucial aspect of the design is how it interacts with consumers. Though this wasn't explicitly stated throughout most of the Architecture proposal – nothing in the proposal prohibits that the actors reading, updating and modifying medical records might be consumers themselves! Of course – this must be done in a way which doesn't compromise the existing standards and practices of the medical community (which can be assured locally, via each RHIOs access control policies). However – it is extremely important that the design allows technically for consumers to interact with the NHIN – either directly through NHIN-enabled personal health record software, or through NHIN-enabled patient gateway systems – to download, view, review, and in some ways – amend (perhaps through requests to responsible parties) errors in the medical record.

Also – consumers could further embellish the medical record (again – with careful consideration of their privacy concerns – these embellishments would be voluntary) with information they themselves collected about their lifestyle (e.g. exercise), health-relevant statistics (e.g. weight tracked at home, blood pressure, breath peak flow, glucose levels etc). Many healthcare consumers already track this sort of information. Making it available (in a traceable, consumer-controlled fashion) to their healthcare providers can significantly help in improving healthcare quality and costs, as well as provide interesting data analysis possibilities for biosurveillance analysis.