

*Prepared for the:*  
Administration for Children and Families (ACF)

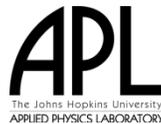
**National Human Services Interoperability Architecture**

**Master Person Index Services  
White Paper**

**DRAFT Version D0.2**

**June 2012**

*Prepared by:*  
The Johns Hopkins University  
Applied Physics Laboratory (JHU/APL)



## Draft Issue

It is important to note that this is a draft document. The document is incomplete and may contain sections that have not been completely reviewed internally. The material presented herein will undergo several iterations of review and comment before a baseline version is published.

This document is disseminated in the interest of information exchange. The Johns Hopkins University Applied Physics Laboratory (JHU/APL) assumes no liability for its contents or use thereof. This report does not constitute a standard, specification, or regulation. JHU/APL does not endorse products or manufacturers. Trade and manufacturer's names appear in this report only because they are considered essential to the object of this document.

Note: This document and other NHSIA-related documentation are available for review from the NHSIA SharePoint site. Updates and any additional documents will be published on that site. The URL for the site is <https://partners.jhuapl.edu/sites/HSNIA>. The version D0.1 and D0.2 documents may be viewed or downloaded from the document library named [NHSIA Drafts](#).

Review and comments to this document are welcome. To comment, either post your feedback in the [NHSIA Drafts Comments](#) library or send comments to [NHSIAArchitectureTeam@jhuapl.edu](mailto:NHSIAArchitectureTeam@jhuapl.edu).

Valerie B. Barnes  
The Johns Hopkins University Applied Physics Laboratory  
11100 Johns Hopkins Road  
Laurel, MD 20723  
Phone: 717-352-0131  
E-Mail: [Valerie.Barnes@jhuapl.edu](mailto:Valerie.Barnes@jhuapl.edu)

## Table of Contents

1	Introduction .....	1
2	Terminology .....	1
3	What do we mean by “master person index services”? .....	3
4	How would MPI services be used? .....	6
5	The MPI concept in NHSIA viewpoints.....	9
5.1	Business Viewpoint .....	9
5.2	Systems Viewpoint .....	9
5.3	Infrastructure Viewpoint .....	9
5.4	Information Viewpoint .....	9
5.5	Capability Viewpoint.....	9
5.6	Project Viewpoint .....	10
6	Summary.....	10
7	References .....	11
	Appendix A – Patient Cross-Referencing Profile .....	13
	Appendix B – Accessibility Appendix .....	17

## List of Figures

Figure 2–1. Master Person Index Concept .....	3
Figure 4–1. Different "Regions" May Set Up their Own MPIs .....	7
Figure 4–2. Scenario for "Match Person" Service.....	8
Figure A–1. Process Flow .....	13
Figure A–2. Actors and Transactions .....	13

## List of Tables

Table 5–1. Capabilities Supported by MPI Concept .....	9
Table A–1. Patient ID (PID) Segment Elements in ITI-8.....	14
Table A–2. OPT Codes.....	15

# 1 Introduction

This document discusses a particular topic relevant to the National Human Services Interoperable Architecture (NHSIA) project. This is one in a series of white papers on specific topics. The focus of this paper is the master person index (MPI) services. This paper summarizes our current thoughts about:

- What do we mean by “master person index services”?
- How would the services be used?
- How the architecture addresses the topic

A major challenge in health and human services is matching information about a person from different sources. Is my Jane Doe the same as your Jane Doe? Today in many regions, each service agency or service provider holds records about the same person but those records are not “discoverable” or usable by anyone outside the agency or service. Caseworkers and service providers have no way to get a clear and complete picture about the client without collecting the information from the client. Information is replicated in many distributed systems. When some basic information changes (e.g., a person’s address), it must be changed in every system. From the client’s perspective, when basic information changes, it would be preferable to update the information once and have all information systems be aware of the change. These are primary motivators for master person index services.

In health care, the [National Committee on Vital and Health Statistics](#) (NCVHS) has worked on concepts for enhanced information capacities for health, health data stewardship, and privacy and security. As part of the Connecting for Health public-private collaboration, the [Markle Common Framework](#) suggests foundations for sharing personal health information in a way that preserves privacy and security. NHSIA builds on those concepts. Some jurisdictions (e.g., Virginia and Maryland) have implemented or are planning to implement solutions to improve interoperability using shared information technology services to match records about people. NHSIA calls those services MPI services.

One of the drivers for NHSIA is to improve the delivery of human services by sharing information using information technology (IT) services. This paper describes a set of IT services to support sharing information about a person. The services should be built to be shared (re-used) by many jurisdictions.

## 2 Terminology

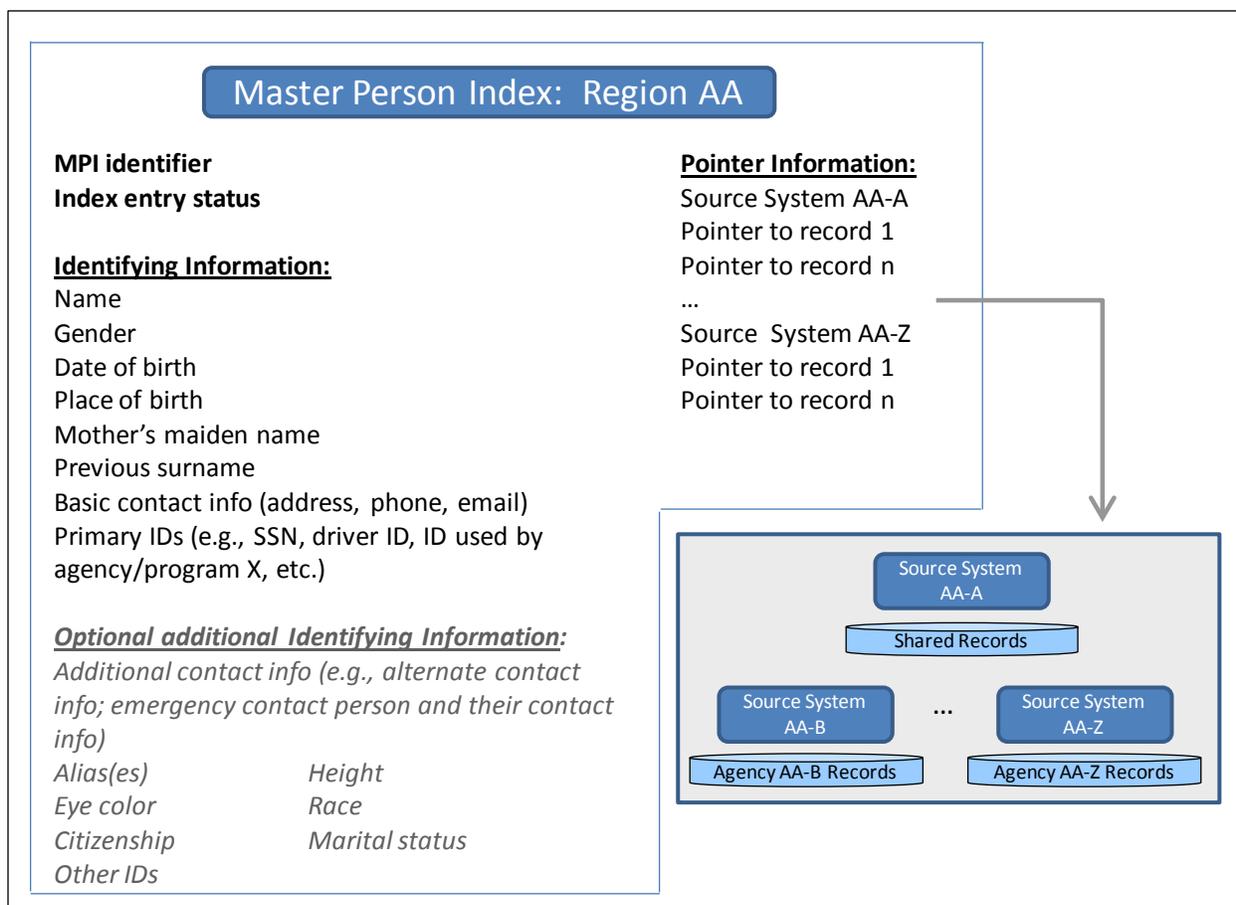
This section defines some of the terms used in this paper.

**Master person index (MPI):** A structure of identifying information about persons and pointers to information about those persons. It may be a physical or virtual structure. The MPI is used to link to information about the person in difference source systems. NHSIA expects each jurisdiction or coalition of jurisdictions (called a “region” in this paper) to maintain its own MPI and make their MPI information discoverable to all authorized users through a set of shared services.

In the first draft of this paper, we suggested that a person should be active in only one MPI at a time. Based on feedback from stakeholders, in this version we revise that idea. Instead, in limited circumstances, a person may be active in more than one MPI at a time. For example, when a court order sets the terms for child support, the related case in the human services child support agency in the court’s jurisdiction would be active as long as the court order is in effect. So the MPI in that jurisdiction would contain an active entry for the person receiving child support. If that person moves to another jurisdiction and seeks other human services, the new jurisdiction would establish an MPI entry accordingly.

In general, though, when a person moves from one region to another, the two regions may want to use a shared MPI service to transfer the person as an active member from the old to the new MPI. The old region would not transfer responsibility for existing human services agency records to the new agency. Instead, the old region would change the status of the person in the MPI to be “inactive”. The new region would establish an “active” entry in its MPI for the person.

Figure 2–1 illustrates the MPI concept.



**Figure 2–1. Master Person Index Concept**

**Metadata:** Data that describes other data.

**Person instance data:** An individual set of data about a person.

**Shared information technology (IT) services:** Software components that are made accessible over a network so that they can be combined and reused in the development of applications to support business functions. The services communicate with each other by passing data or by coordinating an activity between two or more services.

### 3 What do we mean by “master person index services”?

Master Person Index services provide the capability to

- Enable a worker to look for information about a person
- Automatically choose the best match for a person using probabilistic techniques
- Maintain the master person index

- Link diverse records from distributed sources

In this context, a person is an individual who is a client for human services. The MPI services enable users to identify, match, access, merge, de-duplicate, and cleanse records to establish a complete and single view of the person.

NHSIA proposes that every jurisdiction implement these shared MPI services:

- **Find Person.** Locate candidate (potential) information in one or more indices or identity sources about the person based on search criteria input. Return one (or more) set(s) of basic demographic and contact information about candidate matching person(s).
- **Identify Available Person Records.** Identify human services information available about the person. Based on the MPI entry parameter input, return a list of pointers to available records about the person.
- **Match Person.** Using matching algorithm(s), identify the best match in the MPI about the person based on search criteria input. Return best possible match that exceeds the specified minimum matching level. (This service would likely be used by an automated process.) Based on the best match MPI entry, return basic demographic data, contact data, and a list of pointers to available records about the person.
- **Create Person Index Entry.** Create an entry in the MPI for a person.
- **Register Person Data.** Update an existing MPI entry to add one or more pointers to additional information. Specify metadata fields necessary to access the information.
- **Update Person Identifying Information.** Update an existing MPI entry to modify or add to the information that identifies the person (e.g., name or address).

This service is nominated as a candidate for implementation in each jurisdiction:

- **Transfer Person Index Entry.** Shift responsibility for maintaining a person's active index entry from one MPI to another. This would occur when the person applies for services in a region with a different MPI or when the person legally attains residency in a different region. This service also registers a pointer to the previous MPI so that the inactive records are easy to locate.

Additional common MPI services may be proposed in future phases of the project.

For each jurisdiction to implement these MPI services the jurisdictions will need to agree on standard semantics (e.g., meaning of terms used to describe business processes and data elements), how the services will be invoked, what information will be returned, how quality will be maintained, how access to information will be

controlled, and other aspects of information sharing. The MPI services provide a foundational capability for additional information sharing about people receiving human services.

Each MPI service should provide responses that adhere to the person's authorization about sharing information, as well as interagency agreements, rules, and policies that affect sharing information. The person's Confidentiality/Privacy Authorization should be one of the items registered in the MPI entry. This will enable the service to look up who should have access to different kinds of information about the person. The service should not return information to the requester unless the requester is authorized to receive it.

All models for MPI services rely on a set of metadata and person instance data. The MPI services may be implemented using different models, including but not limited to:

- **Centralized MPI.** Store a limited set of person instance information in a physical master file. The limited set of information would include enough to uniquely identify the person and give pointers to other systems that hold additional information about the person. The MPI is the system of record (authoritative source) for the identifying information. Some system element actually creates and maintains the Centralized MPI as a physical repository. Member systems register pointers to related information. This approach is sometimes called "consolidated". In some versions of this approach, instead of storing pointers to records in member systems as part of the MPI itself, the centralized MPI provides a unique identifier (primary key) for the index entry for a person and member systems use that key to identify records about the person.
- **Virtual MPI.** Create a virtual registry, not a physical one. The source systems provide information to the registry; thus the MPI is only a view of the identifying information elements (and pointers to records) from person instance records maintained by the source systems. This approach is sometimes called "federated" or "reference" MPI. MPI services would access the information from the source systems and provide the same kind of responses as in the Centralized MPI approach. This approach allows different source systems to have their own versions of the same information (e.g., address), which must be resolved by the user of the information.
- **Hybrid.** Combines the Centralized MPI and Virtual MPI approaches. Some systems retain and use identifying information from their own person instance data and others use the identifying information data stored in the Centralized MPI as the authoritative source. All use the pointer information, whether stored in the Centralized MPI or generated on the fly from the Virtual MPI.

- Replicated MPI. Replicate the limited set of person identifying information from various sources in a central index and give pointers to other systems that hold additional information about the person.

Other models are also possible. The Centralized MPI model may be chosen across a region or jurisdiction where it is relatively straightforward for agencies to agree on control of and to share the identifying information for a person. This approach promotes consistent use and management of identifying information. The Virtual MPI allows existing systems to keep their own instances of index entries; the MPI services provide a way to share the index information. The Hybrid MPI provides a reasonable transition model for those who wish to move towards a Centralized MPI. The Replicated MPI reduces the “hits” against legacy systems until the information they hold needs to be accessed from an external system. However, replicating index entry information in this model poses maintenance and synchronization challenges.

NHSIA does not promote a particular information storage model to support the MPI services. By focusing on the functionality of shared MPI services, the architecture gives flexibility to NHSIA adapters to implement any of a variety of approaches. Conceptually, NHSIA discusses the notion of a Master Person Index as a set of information that supports the MPI services. The index would provide an entry for each “person”.

Fundamental principles associated with these services are:

- Person privacy protection
- Vendor and platform neutral
- Best practices; agile and extensible
- Promotion of widespread adoption
- Flexible implementation models

## **4 How would MPI services be used?**

Objectives:

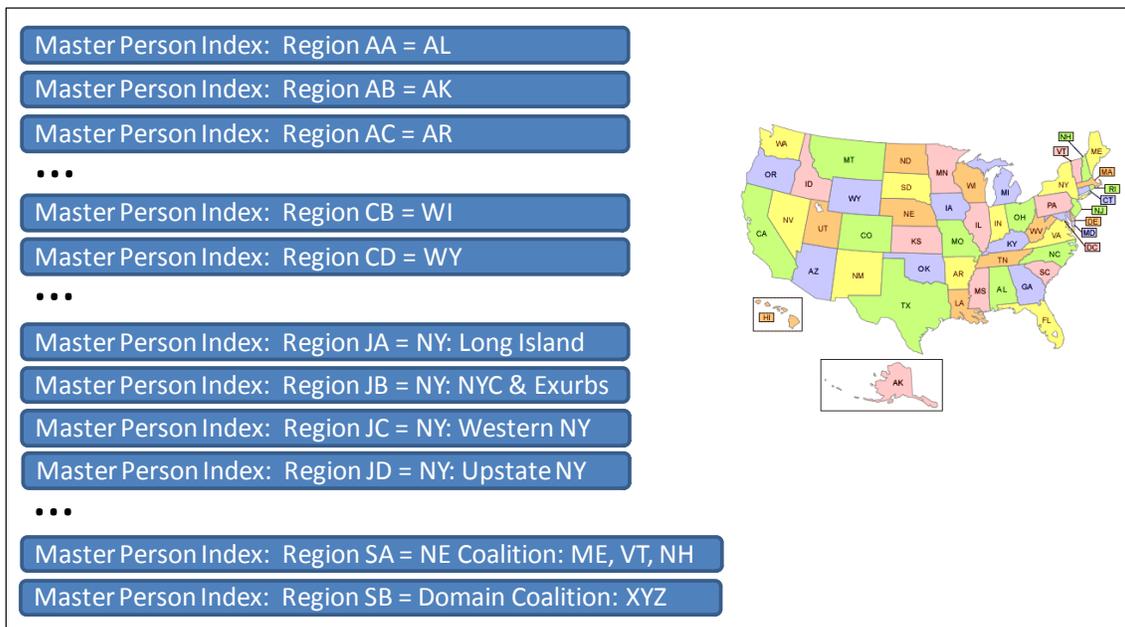
- Identify existing record(s) about a person with a high degree of accuracy (few false positives or false negatives)
- Avoid duplicate records about the same person
- Retain the pedigree of the information sources
- Keep accurate current information
- Register data about a person
- Link to data about a person
- Support “Unique Person Identifier” if adopted
- Maintain historical look-up data

In today’s environment, likely search criteria (and, hence, identifying information about a person in the MPI) include

- Name
- Gender
- Date of birth
- Place of birth
- Mother’s maiden name
- Previous surname
- Address
- Phone
- ID (e.g., SSN, case number, person ID)

As technology and security measures improve and costs go down, additional or substitute criteria may become practical. For instance, biometric data such as fingerprints or retinal images may become routine identifiers. Some other unique person identifier (perhaps a long “serial number” that has no particular meaning) may be assigned and accepted if fraud, spoofing, and identity theft can be avoided.

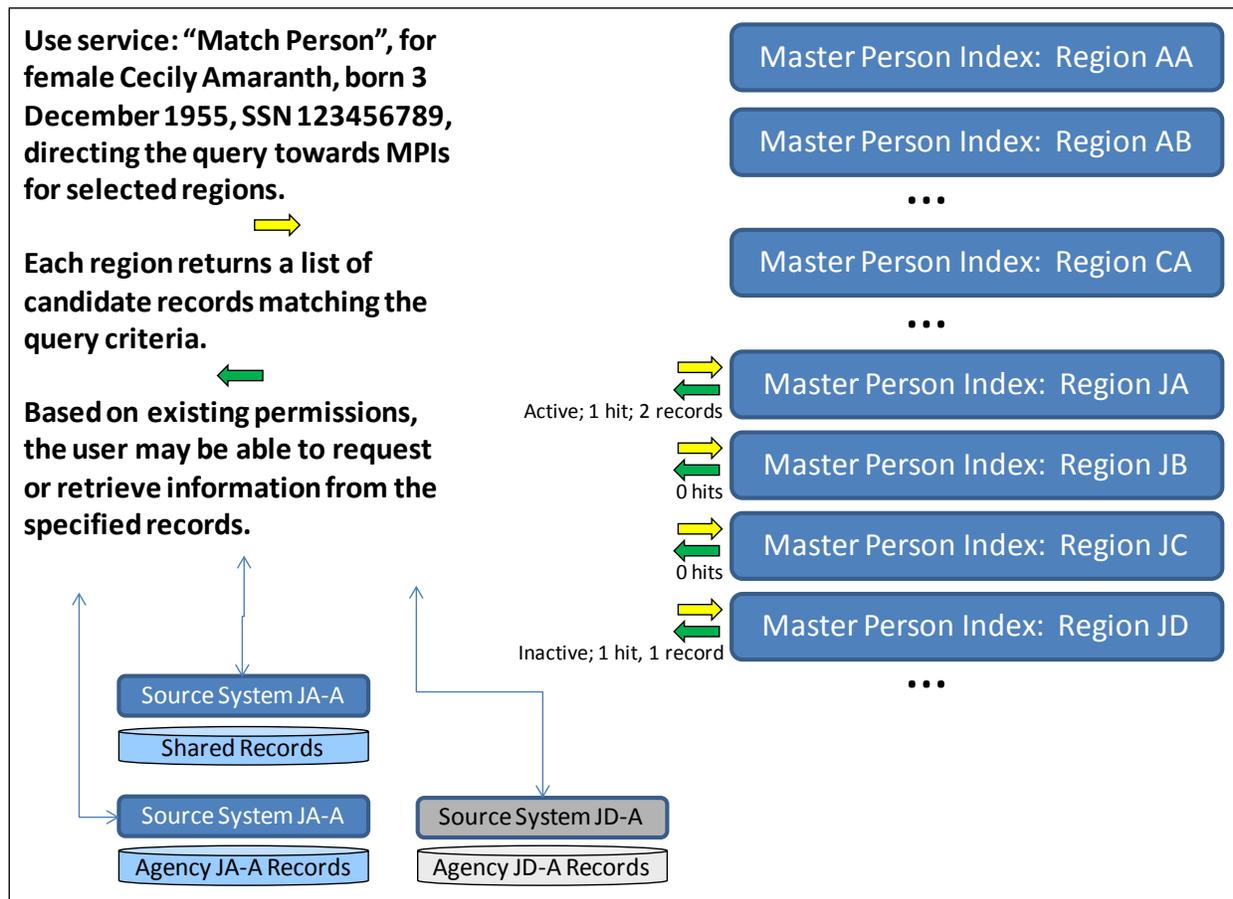
Regional deployment of MPI services would be achievable in the near term. Nationwide deployment may be considered as a longer-term option. Regional shared MPI services could be adopted within regions of different sizes (statewide, countywide, citywide, multi-state), depending on the organizational complexity, service complexity, service population, and other factors. Figure 4–1 shows how different “regions” might organize their MPIs.



**Figure 4–1. Different "Regions" May Set Up their Own MPIs**

MPI services that span several jurisdictions could enhance information sharing across those jurisdictions. If people are displaced as a result of a natural disaster, a regional approach could provide quicker responses to meeting their needs. Jurisdictions that join a regional MPI coalition would need to agree on the implementation strategy.

Figure 4–2 illustrates a scenario for the MPI service “Match Person”.



**Figure 4–2. Scenario for "Match Person" Service**

Some application would invoke the Match Person service. In this example, the query is for a female named Cecily Amaranth, born 3 December 1955, SSN 123456789. The user directs the query to four MPIs, in regions JA, JB, JC, and JD (notionally as shown in Figure 4–1, these might be the MPIs that cover New York state). Each region returns a list of the candidate records that match the query criteria. The user got lucky in this case: Only Region JA has a single active MPI entry that matches the criteria; there is a shared record managed by system JA-A and an agency record managed by system JA-B. Region JD has a single inactive MPI entry that matches the criteria. The responses from Regions JA and JD include pointers to the records and if the user has appropriate permission, he/she may be able to access the records immediately.

In an alternative approach, an application may direct the Match Person service towards a single MPI and set controls to instruct the request be directed towards additional MPIs.

## 5 The MPI concept in NHSIA viewpoints

Each NHSIA viewpoint deals with different aspects of the MPI concepts.

### 5.1 Business Viewpoint

The business model identifies business processes and activities that involve the collection of information about persons and manage the MPI information. Those processes and activities will use MPI services to avoid duplication of information and link records together.

### 5.2 Systems Viewpoint

The systems viewpoint describes the MPI services listed above and suggests which applications will use them.

### 5.3 Infrastructure Viewpoint

The infrastructure viewpoint describes Information Aggregation architecture patterns that support the MPI services options identified earlier.

### 5.4 Information Viewpoint

The information viewpoint describes the metadata necessary to support the MPI concepts and a standard information exchange to support the MPI services. IHE (Integrating the Healthcare Enterprise®) defines a Patient Identifier Cross-Referencing Profile in their group of IHE IT Infrastructure Profiles that may provide a model (Appendix A – Patient Cross-Referencing Profile).

### 5.5 Capability Viewpoint

The MPI concept supports the capabilities listed in Table 5–1. The concept provides primary support to the Locate Client Data capability. For the other capabilities, the MPI services enable the capability but other services are also necessary to accomplish the intended functionality.

**Table 5–1. Capabilities Supported by MPI Concept**

ID	Support	Capability Name	Capability Description
55	Primary	Locate Client Data	An assistor/ caseworker can locate client data based via a Master Person Index (MPI) using an

ID	Support	Capability Name	Capability Description
			index or statistical matching algorithm.
17	Secondary	Access Client Data	An assistor/caseworker can access client-related data from authoritative sources anywhere in the country from anyplace with Internet access.
3	Secondary	Change Jurisdictions	An applicant or client can move from one jurisdiction to another and conveniently transfer information and maintain benefits or move to the equivalent benefits in the new jurisdiction.

## 5.6 Project Viewpoint

The “NHSIA Core” Concepts artifact identifies several MPI services as core IT services. Core IT services provide foundational capabilities. The Jurisdiction’s Guide to NHSIA Implementation recommends that jurisdictions plan to implement MPI services early in their alignment with NHSIA.

Some jurisdictions have already implemented some aspects of the NHSIA MPI concepts. Virginia’s Master Data Management approach uses IBM’s Initiate Master Data Service to implement MPI services. Maryland’s CRISP (Chesapeake Regional Information System for our Patients) also uses IBM’s Initiate product for a federated MPI.

## 6 Summary

This section summarizes the essence of the MPI and MPI services

- The US is covered completely by non-overlapping MPI regions.
- Every region has an MPI and a standard set of associated services.
- The standard set of services includes those outlined here. Standardized interfaces support the services.
- Authorized users (e.g., workers and software applications) may find a person anywhere in the US with a single query to a service that looks first in the local MPI and then in others until a potential match is found or all regions have been queried.

## 7 References

National Committee on Vital and Health Statistics, Reports and Recommendations. Available online at <http://www.ncvhs.hhs.gov/reptrecs.htm>.

Connecting for Health Common Framework, Record Locator Service: Technical Background from the Massachusetts Prototype Community, T6, 2006. Available online at [http://www.markle.org/sites/default/files/T6\\_RecordLocator.pdf](http://www.markle.org/sites/default/files/T6_RecordLocator.pdf).

IBM® Initiate® Master Data Service®, A Platform for Master Data Management to Help You Know and Trust Your Data, IMW14345-USEN-00, 2010. Available online at <http://www-01.ibm.com/software/data/infosphere/master-data-service/>.

IBM® Software White Paper, How Master Data Management Serves the Business, IMW14570-USEN-00, 2011. Available online at <http://www-01.ibm.com/software/data/infosphere/industry/healthcare/>.

Maryland Health Care Commission and Health Services Cost Review Commission, CRISP (Chesapeake Regional Information System for our Patients), A Plan for a Citizen-Centric Statewide Health Information Exchange in Maryland, February 2009. Available online at [http://mhcc.dhmdh.maryland.gov/hit/hiePolicyBoard/Documents/CRISP\\_FinalReport.pdf](http://mhcc.dhmdh.maryland.gov/hit/hiePolicyBoard/Documents/CRISP_FinalReport.pdf)

CRISP Response to the Request for Application for a Consumer-Centric Health Information Exchange for Maryland, May 2009. Available online at <http://mhcc.dhmdh.maryland.gov/hit/hiePolicyBoard/Documents/CRISP.pdf>

John Radcliffe, Magic Quadrant for Master Data Management of Customer Data, Gartner RAS Core Research Note G00206031, 4 October 2010.

Health Information Technology (HIT) Policy and Standards Committees, Patient Protection and Affordable Care Act Section 1561 Recommendations, Core Data Analysis (Appendix B), adopted with slight edits by HHS Secretary Kathleen Sebelius, September 17, 2010. Available online at <http://healthit.hhs.gov/ACA/section1561>

American Health Information Management Association, Fundamentals for Building a Master Patient Index/Enterprise Master Patient Index (Updated), Journal of AHIMA (Updated September 2010). Available online at [http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1\\_048389.hcsp?dDocName=bok1\\_048389](http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_048389.hcsp?dDocName=bok1_048389)

David Loshin, Practical Fundamentals for Master Data Management, a DataFlux White Paper, 2011. Available online at

<http://www.sas.com/offices/europe/denmark/pdf/WP075-Practical-Fundamentals-for-Master-Data-Manag.pdf>

<http://www.sas.com/whitepapers/index.html>

## Appendix A – Patient Cross-Referencing Profile

IHE (Integrating the Healthcare Enterprise®) defines a Patient Identifier Cross-Referencing Profile in their group of IHE IT Infrastructure Profiles.

“The Patient Identifier Cross Referencing (PIX) Integration Profile supports the cross-referencing of patient identifiers from multiple Patient Identifier Domains by:

- Transmitting patient identity information from an identity source to the Patient Identifier Cross-reference Manager.
- Providing the ability to access the list(s) of cross-referenced patient identifiers either via a query/ response or via an update notification.”

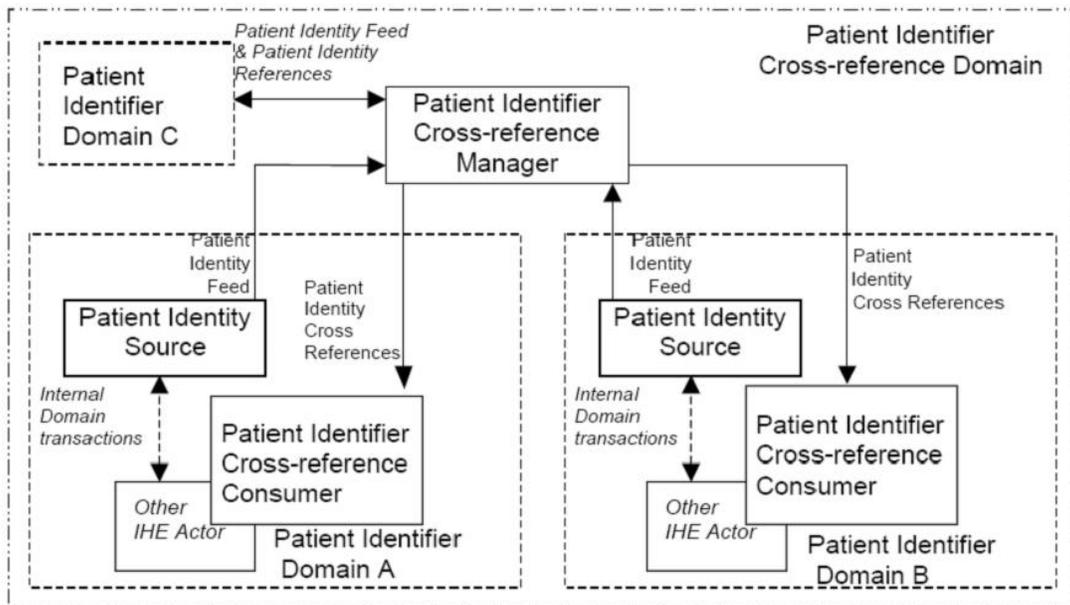


Figure A-1. Process Flow

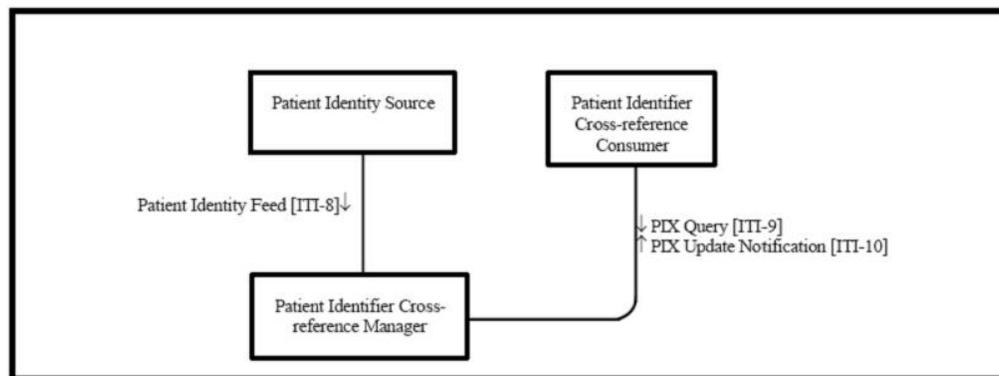


Figure A-2. Actors and Transactions

Source: [http://wiki.ihe.net/index.php?title=Patient Identifier Cross-Referencing](http://wiki.ihe.net/index.php?title=Patient_Identifier_Cross-Referencing)

The Patient Identity Feed (ITI-8 in Figure A–2 above) is defined in the IT Infrastructure Technical Framework, Volume 2a, Transactions Part A Sections 3.1-3.28, Revision 7.0, August 2010. Table A–1 shows the fields in the Patient ID (PID) segment:

**Table A–1. Patient ID (PID) Segment Elements in ITI-8**

Element Name	OPT
Set ID - Patient ID	O
Patient ID	O
Patient Identifier List	R
Alternate Patient ID	O
Patient Name	R
Mother's Maiden Name	R+
Date/Time of Birth	R+
Administrative Sex	R+
Patient Alias	O
Race	O
Patient Address	R2
Country Code	O
Phone Number - Home	R2
Phone Number - Business	R2
Primary Language	O
Marital Status	O
Religion	O
Patient Account Number	O
SSN Number- Patient	R2
Driver's License Number - Patient	R2
Mother's Identifier	O
Ethnic Group	O
Birth Place	O
Multiple Birth Indicator	O
Birth Order	O
Citizenship	O
Veterans Military Status	O
Nationality	O
Patient Death Date and Time	O
Patient Death Indicator	O

Source: [http://www.ihe.net/Technical\\_Framework/upload/IHE ITI TF Rev7-0\\_Vol2a\\_FT\\_2010-08-10.pdf](http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF_Rev7-0_Vol2a_FT_2010-08-10.pdf)

OPT Codes used in Table A–2 are defined in IHE IT Infrastructure Technical Framework, Volume 2x, Volume 2 Appendices, Appendix C (HL7 Profiling Conventions). Table A–2 defines the codes.

**Table A–2. OPT Codes**

<b>OPT Code</b>	<b>Meaning</b>
R	Required
R2	This is an IHE extension. If the sending application has data for the field, it is required to populate the field. If the value is not known, the field may not be sent.
R+	This is an IHE extension. This is a field that IHE requires that was listed as optional within the HL7 standard.
O	Optional
C	Conditional

Source: [http://www.ihe.net/Technical\\_Framework/upload/IHE ITI TF Rev7-0\\_Vol2x\\_FT\\_2010-08-10.pdf](http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF_Rev7-0_Vol2x_FT_2010-08-10.pdf)

The Healthcare Information Technology Standards Panel (HITSP) built on the IHE transaction in their [HITSP Patient ID Cross-Referencing Transaction Package](#).

This page is intentionally blank.

## Appendix B – Accessibility Appendix

This section contains accessible versions of figures and tables in this document. Table and figure numbers that appear here correspond to versions that appear earlier in this document.

Figure 2-1. Master Person Index Concept (shows likely entries in an MPI)

- MPI identifier
- Index entry status
- Identifying Information:
  - Name
  - Gender
  - Date of birth
  - Place of birth
  - Mother's maiden name
  - Previous surname
  - Basic contact info (address, phone, email)
  - Primary IDs (e.g., SSN, driver ID, ID used by agency/program X, etc.)
- Optional additional Identifying Information:
  - Additional contact info (e.g., alternate contact info; emergency contact person and their contact info)
  - Alias(es)
  - Height
  - Eye color
  - Race
  - Citizenship
  - Marital status
  - Other IDs

- Pointer Information
  - Source System AA-A
    - Pointer to record 1
    - Pointer to record n
    - ...
  - Source System AA-Z
    - Pointer to record 1
    - Pointer to record n