##### 

Perceptive Reach

Integrated Reach Database System

(IRDS)

Interface Design Specification



Department of Veterans Affairs

August 2015

Version 3.0

Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Version | Description | Author |
| 8/14/2015 | 3.0 | Final Review | Monica Mohler/Paul Bradley |
| 8/10/2015 | 2.9 | Peer Review | Radina Ivanova |
| 8/9/2015 | 2.8 | Added | Bill Balshem |
| 7/13/2015 | 2.7 | Final Review | Monica Mohler/Paul Bradley |
| 7/8/2015 | 2.6 | Peer Review | Radina Ivanova |
| 7/8/2015 | 2.5 | Added updates to CDW setup. Add info about the HealthIndicators.gov data source | Bill Balshem |
| 6/12/2015 | 2.4 | Final Review | Monica Mohler/Paul Bradley |
| 6/11/2015 | 2.3 | Peer Review | Radina Ivanova |
| 6/10/2015 | 2.2 | June Update | Andrew Smith/Andal Fequiere |
| 5/12/2015 | 2.1 | Final Review | Monica Mohler/Paul Bradley |
| 5/11/2015 | 2.0 | Peer Review | Radina Ivanova |
| 5/11/2015 | 1.9 | May Update | Andrew Smith/Andal Fequiere |
| 4/13/2015 | 1.8 | April Submission: no revisions | Monica Mohler/Paul Bradley |
| 3/11/2015 | 1.7 | Final Review | Monica Mohler/Paul Bradley |
| 3/3/2015 | 1.6 | Peer Review | Radina Ivanova |
| 3/2/2015 | 1.5 | Added contect on:  SSIS processes running against the CDW server ;  Details about VA risk model output being accessed by the IRDS system | Bill Balshem |
| 2/12/2015 | 1.4 | Final Review | Monica Mohler/Paul Bradley |
| 2/4/2015 | 1.3 | Peer Review | Radina Ivanova |
| 02/03/2014 | 1.2 | Updated CDW Content | Bill Balshem |
| 01/14/2014 | 1.1 | Final Review | Monica Mohler/Paul Bradley |
| 12/12/2014 | 1.0 | Final Review | Monica Mohler/Paul Bradley |
| 12/11/2014 | .03 | Peer Review | Radina Ivanova |
| 11/26/2014 | .02 | Added dashboard content | Robert Snelling, Andal Fequiere, Paul Bradley |
| 11/07/2014 | .01 | Added initial database content | Bill Balshem |
| 04/25/2012 | .00 | First Draft | Paul Bradley |

Table of Contents

1. About this document 5

2. Introduction 5

2.1. Scope 5

2.2. Assumptions 6

2.3. References 6

2.4. Acronyms 6

2.5. System Identification 8

3. Interface Definition & Requirements 8

3.1. System Overview 8

3.2. Interface Overview 8

3.2.1. External Data Sources 9

3.2.2. External Interfaces 11

3.3. Data Transfer 13

3.3.1. SQL Server Integration Services (SSIS) 13

3.3.2. Remote Procedure Calls (RPC) 14

3.4. Communications Methods 15

3.4.1. IRDS Surveillance Dashboard 15

3.5. Performance Requirements 16

3.6. Security 17

4. Interface Verification 18

4.1. Data Source Imports 18

4.1.1. Batch Imports 18

# About this document

This document is a “work in progress.” The interfaces will be designed through a series of sprint cycles in collaboration with VA stakeholders and users. The document will be populated with content as the design evolves with each sprint, including all potential data options, configurability rules and options. The early versions of the document will focus on the conceptual design, key business requirements, and the primary end-user features related to the design. Future versions will add additional detail as it becomes available.

# Introduction

VA is seeking to expand suicide prevention to include upstream approaches, designed to reduce initiation or escalation of a risk factor. Upstream suicide interventions target individuals or groups who exhibit biological, psychological, or social risk factors that are more prominent among high-risk groups than among the larger population. Understanding the unique needs of our nation’s Veterans and the military culture as it relates to stigma and mental health is important for early intervention. The goal of the Perceptive Reach Integrated Reach Database System (IRDS) innovation is to promote the general health of the Veteran population and effectively intervene in issues before they escalate in crisis.

The IRDS solution innovates the current process of risk data collection, analysis, and use in effective intervention strategy. The solution will harness the power of large and diverse data stores to aggregate, analyze and identify risk onset as well as reveal previously unidentified at-risk individuals and populations as a holistic and integrated approach.

The IRDS innovation will serve to bolster the three major components of VHA’s Strategic Plan for Suicide Prevention: surveillance, risk and protective factors, and prevention interventions. The IRDS innovation will target antecedent events specific to Veteran populations prior to the onset of risk to mitigate the development of risk.

## Scope

The IRDS Interface Design Specification document describes the relationship between IRDS and each of the external systems connected to it in terms of data items, protocols, and timing of events.

This Interface Design Specification will describe what data will be transferred between the IRDS and its input sources and output destinations such as:

* Input: Suicide Data Repository (SDR)
* Input: Veterans Benefits Administration Data
* Input: Corporate Data Warehouse (CDW)
* Input: VistA Data
* Output: VA Suicide Prevention Coordinators
* Output: Other VA and non-VA Outreach and Intervention Coordinators & Clinicians
* Output VA Suicide Prevention stakeholders

This document should be read in conjunction with the IRDS System Design Document (SDD).

## Assumptions

The assumptions for the Interface Design Specification have not been determined. The document will be updated as soon as the information becomes available.

## References

The following Perceptive Reach IRDS documents may be referenced in tandem with the information recorded here:

* Contractor Project Management Plan (CPMP)
* System Design Document (SDD)

## Acronyms

Table : Acronyms

| Acronym | Term |
| --- | --- |
| Army STARRS | Army Study to Assess Risk and Resilience in Service members |
| BIRLS | Beneficiary Identification Records Locator System |
| CD | Compact Disk |
| CDC | Center for Disease Control |
| DoD | Department of Defense |
| ETL | Extract, Transform, Load |
| GB | Gigabyte |
| ICD | International Classification of Diseases |
| IM/IT | Information Management/Information Technology |
| IRDS | Integrated Reach Database System |
| IT | Information and Technology |
| NDI | National Death Index |
| ODBC | Open Database Connectivity |
| OIT | Office of Information and Technology |
| OMHS | Office of Mental Health Services |
| PII | Personally Identifiable Information |
| PHI | Protected Health Information |
| PMP | Project Management Plan |
| REST | Representational State Transfer |
| RPC | Remote Procedure Call |
| RSD | Requirements Specification Document |
| RTM | Requirements Traceability Matrix |
| SAS | Statistical Analysis System |
| SDCD | State Death Certificate Data |
| SDR | Suicide Data Repository |
| SFTP | Secure File Transfer Protocol |
| SMITREC | Serious Mental Illness Treatment Resource and Evaluation Center |
| SPAN | Suicide Prevention Applications Network |
| SQL | Structured Query Language |
| SSIS | SQL Server Integration Services |
| SSN | Social Security Number |
| TB | Terabyte |
| UI | User Interface |
| VA | Department of Veterans Affairs |
| VBA | Veterans Benefits Administration |
| VCL | Veterans Crisis Line |
| VHA | Veterans Health Administration |
| VLER | Virtual Lifetime Electronic Record |
| VSSC | VHA Support Service Center |

## System Identification

The system identification for the IRDS have not been determined. The document will be updated as soon as the information becomes available.

# Interface Definition & Requirements

## System Overview

The Perceptive Reach development and field pilot proposes to combine technology, outreach and clinical support to realize a clinically based data-driven early intervention and treatment solution aimed at suicide prevention. IRDS is a solution for analyzing multiple and integrated data sets with cutting-edge data analytic techniques and visualizations to identify at-risk individuals and populations and provide proactive and secure notifications of these results to Veteran support services. The Perceptive Reach project proposes to expand the capabilities of the Suicide Data Repository to include new interfaces to clinical data sources, integrated data analytics capabilities, a surveillance dashboard, and secure messaging.



Figure : IRDS System Overview

## Interface Overview

The primary ingress interfaces will use SQL Server Integration Services (SSIS) to retrieve data from the designated data sources such as the SDR. The SSIS package will be configured to pull data from the designated data source at a specified interval (daily/weekly/monthly/annually).

The primary egress interface will be the Direct Messaging interface used to notify VA Suicide Prevention Coordinators (SPC) and other VA or non-VA Outreach and Intervention coordinators and clinicians. The IRDS will leverage the VLER Direct messaging service utilizing a RESTful interface to transmit message content. Another interface utilized by VA Staff and SPCs will be the IRDS Surveillance Dashboard which will provide browser based visualizations of critical data to identify Veterans at high risk for suicide.



Figure : IRDS Interface Overview

### External Data Sources

The external data sources to be imported into the Reach database of the IRDS have been identified as:

1. VA Suicide Data Repository (SDR) - The SDR contains VA and DoD suicide and mortality data. The data store is Microsoft SQL Server. Periodically data is imported into the system from

* Mortality search results from the National Death Index (NDI)
* State Death Certificate Data (SDCD)
* Veterans Crisis Line (VCL)
* Suicide Prevention Applications Network (SPAN)
* VA Beneficiary Identification Records Locator Subsystem (BIRLS) Death File

1. Veterans Health Information Systems and Technology Architecture (VistA) - It is assumed that some VHA data will be retrieved directly from the VA Vista system. VistA uses an M data platform. VistA data will be imported into the Reach database using custom and/or currently existing RPC calls.
2. Corporate Data Warehouse (CDW) - CDW warehouses VHA data (mostly VistA data) in SQL Server format. VistA data stored in CDW will be imported into the Reach database. The automated data import solution (in the form of SSIS packages) will be deployed on an ETL server provided by CDW. The SSIS packages will:
3. Make SQL connections to both the CDW server that contains the Patient data, and the IRDS production server
4. Pull data from the CDWWork database on the CDW server into the VACI\_IRDS database on that server
5. Perform the appropriate data transformations in the VACI\_IRDS database
6. Push the post transformed data from the VACI\_IRDS database into the Reach database on the IRDS server



**Figure 3: CDW/IRDS databases**

1. Enterprise Data Warehouse (EDW) – The Veterans Benefits Administration (VBA) store Veteran data in SQL format in an Oracle database. Some data elements from VBA (financial status, marital status) may be pulled into the Reach database and linked to VHA records for those individuals by matching on elements such as SSN.
2. HealthIndicators.gov - The web site is a source for statistical information. The IRDS system imports data pertaining to suicide death rates for specified demographics via a call to a web service. The data is returned in XML format. That data is stripped from the XML tags and stored in a table in the Reach database, which is queried by the dashboard against Patient data that it is integrated with.
3. Other Non-VA data sources - The VA has expressed an interest in importing and analyzing data that is external to VA data sources. Some of the sources being considered are LexisNexis and PACER. The requirements team is currently reviewing possibilities with the VA.
4. MedSAS- Vista data is collected in SAS data sets in the MedSAS system. The risk model applied to the IRDS system was developed by the VA against the MedSAS data set. When the IRDS analytics team ran the model against the same data in the CDW, the model was no longer applicable. The IRDS analytics team developed a hybrid risk model that uses VistA data from CW, but uses MedSAS data for demographics of patients (DOB, Race, Marital Status, Urban/Rural, Service Connected, Region).The Reach database will contain a table that contains patient data with the date elements listed above and their Scrambled SSN (scrssn) as a unique identifier. When patients are imported from CDW to the Reach database in the IRDS system, that data is linked to the Medsas table via scrssn. If demographic data for a patient is available in MedSAS the medsas values overwrite the CDW values for the patient in the Patient and PatientRiskFactors tables.

### External Interfaces

The external interfaces for the IRDS solution include the following:

1. IRDS Surveillance Dashboard – The dashboard will provide a browser based user interface featuring representations of the analyzed consolidated data sources. These are processed and organized into visualizations, which will assist SPCs and other VA personnel in their duties.
   * This dashboard will reflect data updated on a daily basis and may include charts, tables, maps, animations, other graphics and visual technology.
   * The dashboard will have configurable features based on user roles as defined by the stakeholders. This will provide for distinct views to meet each role’s business needs.



Figure : Surveillance Dashboard

1. VLER Direct Messaging - The Direct Messaging component will integrate into the existing VLER Direct Messaging architecture. The Direct Messaging will create unique messages based on new and existing data annotated in the IRDS.
   * To send messages, the Perceptive Reach Application will utilize the VLER Direct Messaging API using authentication and conforming to a valid web service call based on the Direct Secure Messaging ICD.
   * The Direct Messaging will utilize public and private keys in creating the authorization header on each web service request to the VLER DM API.
   * The Perceptive Reach Application has registered to use Direct Secure Messaging and the authentication credentials have been provided. The Perceptive Reach application has the ability to send a request to use Direct as a Service web services.
   * The Perceptive Reach Application will authenticate with the VLER DM API using hash-based message authentication code (HMAC).
   * Recipient of Direct Messages
     1. Individuals can receive Direct Messages using the Direct Secure Messaging Webmail Portal, which requires a Personal Identity Verification (PIV) for authentication using certificates to login to the system. Users will have the default user role and will be able to send and receive messages.
     2. An external organization can request access to VLER Direct as a Service (DaaS), which allows edge applications to receive and send Direct messages. The information is encrypted and securely sent to the External Partner, as a Direct Message using the API.
2. SDR SSIS Interface - SSIS will be the primary tool for importing data from SDR into the Reach database.
3. CDW SSIS Interface – CDW has given the IRDS development team a SQL Server project database (VACI\_IRDS), located on a CDW server. The database will be used:
4. To import VistA data from CDW work tables into new tables in the VACI\_IRDS, for the purpose of validating the VA Risk Model and then enhancing it.
5. As work space for the automated import process, built in SSIS, to filter and transform the CDW data before pulling it over the network onto the IRDS server and into the Reach database.
6. CDW SaS Data – The IRDS Analytics team will have access to SaS data files located on the CDW server as part of their validation of the VA Risk model. These files will be accessed through R software located on an authorized VA machine. For the initial run of the risk model in the IRDS production environment, the VA risk model output files will be used to access the Suicide Completion indicator field as part of the calculations. The VA has uploaded these files, in SaS format to the SaS Grid. The files will be imported into SQL tables in the VACI\_IRDS database located on the CDW server VHACDWA01.
7. EDW SSIS Interface – An SSIS packages will be developed to create a SQL connection to the VBA EDW Oracle database. The details for that interface are currently being worked out between the IRDS management team and their VBA points of contact.
8. VistA RPC Interface (tentative) - The VA uses the Veterans Health Information Systems and Technology Architecture (Vista) system, for managing Veterans health data. Data may be imported into the IRDS system directly from VistA using RPC calls when immediate access to recently updated clinical records is retired. As VistA is modernized additional interface protocols may be used in place of RPCs. This is to be determined.
9. Active Directory – The VA uses Active Directory, to manage user accounts within the VA Network. The IRDS application will make HTTP requests to the Active Directory instance located within the VA network for user authentication and account information.

## Data Transfer

### SQL Server Integration Services (SSIS)

SSIS will be the primary tool for importing external data sources into the Reach database

For a specific data import, an SSIS package will be developed to

1. Make a connection to the source (SQL table, text file, other)
2. Import the data into a staging area
3. Make the appropriate data transformations (cleaning, standardization)
4. Load the transformed data into the appropriate Reach data store tables

The execution of SSIS packages (.dtsx files) can be automated by scheduling them as a Windows process via SQL Server Agent.

The following SSIS processes will be developed and run in the IRDS against the CDW server VHACDWA01 (below is a description summary of each process, for more detail please refer to the Database Design Document):

* Risk Model Calculation - A sample of Veterans will be selected to calculate the risk model to be used in the IRDS system. For these Veterans data will be pulled from the CDWWork database into the IRDS project database (VACI\_IRDS). Risk factors will be calculated for these Veterans and an aggregate list of risk factors with scoring coefficients will be imported into a table in the Reach database of the IRDS server.
* Surveillance Model Data Import - A sample of Veterans will be selected on a regular basis to be given a risk score against the risk model and identified as being high risk or not. In the initial run of this import, the Veterans in the sample will have their risk factors calculated, in the same manner they were in the Risk Model Calculation. Each Veteran’s demographic data and risk model calculations will be imported into the IRDS Reach database. Each time the Surveillance Model Scoring process is run, the Surveillance Model Data Import will be run right before, to update the risk factors for each Veteran as necessary, due to any new activity found for that Veteran in the CDWWork database (such as new prescription of inpatient visit).
* Surveillance Model Scoring - On a regular basis, risk scores will be calculated for each Veteran, by summing their risk factors against the risk model coefficients. For a Veteran that is identified as high risk during that run, data on that individual that may be viewed in the IRDS dashboard, such as Emergency Contact information, will be pulled from the CDWWork database on the CDW server and copied into the Reach database on the IRDS server.

### Remote Procedure Calls (RPC)

Data may be imported into the IRDS system directly from the Vista using RPC calls. VistA data is stored against a MUMPS back end, which uses text-based files for data storage. For each source of VistA data imported from into IRDS:

1. Either a custom RPC will be written (in M) or a currently existing one will be leveraged
2. An automated java process will execute the RPC and return the query results in text format
3. Those results will be stored in a flat file on the IRDS server to be imported into the Reach database via a SSIS package (using the steps listed in the section above)



Figure : IRDS Data Import Process Flow

## Communications Methods

### IRDS Surveillance Dashboard

The REST APIs provide programmatic access to read only IRDS data that is customized to provide data for visualization components and widgets. Web service APIs that adhere to the REST architectural constraints are called RESTful. HTTP based RESTful APIs are defined with these aspects:

* Base URL, such as <http://example.com/resources/>
* An Internet media type for the data. This is often JSON but can be any other valid Internet media type (e.g. XML, Atom, microformats, images, etc.)
* Standard HTTP methods (e.g., GET, PUT, POST, or DELETE)
* Hypertext links to reference state

This allows the dashboard to have real-time data responses available in JSON, a lightweight data-interchange format. The dashboard architectural approach implements all data requests in an asynchronous and non-blocking way, which allows for multiple data connections. To address safety and security the REST API will be managed through a session token or API key which will provide for privileged based access and will prevent unauthorized use.

## Performance Requirements

These are pilot (proof-of-concept) performance requirements. Once this innovation is proven, the production requirements will be determined by VA.

Table : Pilot Performance Requirements

| Requirement | Specification |
| --- | --- |
| Availability | There is not a high-availability requirement for this innovation pilot. |
| Capacity | The IRDS pilot solution will be verified with a 500GB storage capacity. |
| Backup Capacity | There is not a data or system backup requirement for this innovation pilot. |
| Number of Users | System Administrators: 2  Tool Administrators: 2  Developers: 8  Testers: 4  End Users: 32 |
| Maximum Number of Simultaneous Users | 46 |
| Maximum Downtime | There is not a hig-availability requirement for this innovation pilot. |
| Network Interface(s) | 1 Gigabit Ethernet (or greater) network interface will be required to connect the IRDS pilot solution to the VA intranet. |
| Database Interface(s) | Data imports will be run during non-business hours and be completed before the next business day. (i.e. SDR imports will run 1st of month at 11PM EST and be completed before 6AM the next day). |
| Help Desk Support | To be determined once IRDS pre-production environment is identified. |

## Security

The IRDS solution uses Personally Identifiable Information (PII) and Protected Health Information (PHI). This data will be secured and protect per the following regulations and directives:

* Health Insurance Portability and Accountability Act (HIPAA) of 1996
* U.S.C. § 3541, “Federal Information Security Management Act (FISMA) of 2002”
* U.S.C. § 552a, as amended, “The Privacy Act of 1974”
* VA Directive 6500, “Managing Information Security Risk: VA Information Security Program,” September 20, , 2012
* VA Handbook 6500, “Risk Management Framework for VA Information Systems – Tier 3: VA Information Security Program,” September 20, 2012
* VA Handbook 6500.1, “Electronic Media Sanitization,” March 22, 2010
* VA Handbook 6500.2, “Management of Data Breaches Involving Sensitive Personal Information (SPI)”, January 6, 2012
* VA Handbook 6500.3, “Assessment, Authorization, And Continuous Monitoring Of VA Information Systems,” February 3, 2014
* VA Handbook, 6500.5, “Incorporating Security and Privacy in System Development Lifecycle” March 22, 2010
* VA Handbook 6500.6, “Contract Security,” March 12, 2010
* VA Directive 6508, VA Privacy Impact Assessment, October 3, 2008
* VA Directive 6300, Records and Information Management, February 26, 2009
* VA Handbook, 6300.1, Records Management Procedures, March 24, 2010
* OMB Memorandum M-07-16, Safeguarding Against and Responding to the Breach of Personally Identifiable Information, May 22, 2007
* NIST SP 800-63-2, Electronic Authentication Guideline, August 2013

The system security will be maintained accord to VA policy and will include obtaining an Authority to Operate (ATO). This will include the creation of security artifacts:

* Security Management Plan,
* System Security Plan,
* Privacy Impact Assessment,
* Configuration Management Plan,
* Disaster Recovery Plan,
* Risk Assessment and Security Certification Checklist.

The Governance, Risk and Compliance (GRC) tool (RiskVision) shall act as the management tool for the Assessment and Authorization (A&A) process, and systems shall be assessed in RiskVision by an Office of Cyber Security (OCS) representative for an accreditation recommendation to be submitted to the OIS Chief Information Security Officer (CISO) and VA Chief Information Officer (CIO) for final review and determination.

# Interface Verification

## Data Source Imports

### Batch Imports

Batch imports will be run periodically using SQL Server integration services (SSIS). One output of an SSIS package run will be a completion report that contains information on the run including:

1. No of records imported(inserts, updates)
2. Any errors that occurred during the run
3. Any data integrity issues identified per specific records( i.e. a field that should be numeric contained the letter ‘a’)

These completion reports should be reviewed and verified after an import is run.