

System Design Document

Genomic Information System for Integrated Science 2
(Genesis2) Technical Services

Build 8



January 2018

Document Version 5.0

Department of Veterans Affairs

Document Revision History

Date	Revision	Description	Author
03/03/2016	0.1	Initial version	Engility
04/07/2016	0.2	Updated draft	Engility
06/16/2016	0.3	Updated draft	Engility
08/16/2016	0.4	Updated version	Booz Allen Hamilton
08/29/2016	0.5	Adding design detail	Booz Allen Hamilton
09/11/2016	0.6	Added design detail	Booz Allen Hamilton
04/14/2017	0.7	Made modifications to design detail	Booz Allen Hamilton
05/12/2017	0.8	Updated Version after STAT Review	Booz Allen Hamilton
06/16/2017	0.9	Updated for Build 3	Booz Allen Hamilton
09/15/2017	1.0	Updated for Build 4 new Data Request features	Booz Allen Hamilton
11/03/2017	2.0	Updated for Build 5 Terminology Services additions	Booz Allen Hamilton
12/15/2017	3.0	Updated for Build 7 Terminology Services additions	Booz Allen Hamilton
12/22/2017	4.0	Updated for Build 6 new Data Request features	Booz Allen Hamilton
01/19/2018	5.0	Updated for Build 8 Terminology Services additions	Booz Allen Hamilton

Artifact Rationale

The System Design Document (SDD) is a dual-use document that provides the conceptual design as well as the as-built design. This document will be updated as the product is built to reflect the as-built product.

When to Complete Each Section of the SDD

Section	Completed on or Before PMAS Phase	Rationale
1 – Introduction	MS 0 Review; updated thereafter	Conceptual design should inform evaluation of investments
2 – Background	MS 0 Review; updated thereafter	Conceptual design should inform evaluation of investments
3 – Conceptual Design	MS 0 Review; updated thereafter	Conceptual design should inform evaluation of investments
4 – System Architecture	MS 0 Review; updated thereafter	Conceptual design should inform evaluation of investments
5 – Data Design	MS 1 Review; updated thereafter	Design details should be elaborated upon during PMAS Planning phase and prior to development
6 – Detailed Design	MS 1 Review; updated thereafter	Design details should be elaborated upon during PMAS Planning phase and prior to development
7 – External System Interface Design	MS 1 Review; updated thereafter	Design details should be elaborated upon during PMAS Planning phase and prior to development
8 – Human Machine Interfaces	MS 1 Review; updated thereafter	Design details should be elaborated upon during PMAS Planning phase and prior to development
Attachments	MS 1 Review; updated thereafter	Design details should be elaborated upon during PMAS Planning phase and prior to development

A product's system design should be defined conceptually prior to the allocation of personnel and resources that occur at project initiation. This gives the enterprise an opportunity to evaluate IT investments before project teams are stood up and funding is allocated. Sections 1- 4, which discuss the high-level design, should be completed prior to MS 0. All sections should be completed and updated before MS 1. Projects will need to address all SDD approval constraints prior to the MS 2 review. In addition, the SDD should reflect the as-built product going into the MS 2 review.

Activity	New Capability (1)	Feature Enhancement (2)
Field Deployment (A)	Yes	Yes
Cloud/Web Deployment (B)	Yes	Yes
Mobile Application (C)	No	No

Table of Contents

1. Introduction.....
1.1. Scope.....	1
1.2. User Profiles.....	4
2. Background.....
2.1. Overview of the System.....	6
2.2. Overview of the Business Process.....	6
2.3. Overview of the Significant Requirements.....	7
3. Conceptual Design
3.1. Conceptual Application Design Principles	9
3.2. Conceptual Application Design	11
3.3. Conceptual Data Design.....	22
3.4. Conceptual Infrastructure Design	38
4. System Architecture.....
4.1. Hardware Architecture.....	45
4.2. Software Architecture.....	46
4.3. Network Architecture.....	51
4.4. Service Oriented Architecture / ESS	51
4.5. Enterprise Architecture	52
5. Data Design.....
5.1. Human Machine Interface.....	54
5.2. DBMS Files	60
5.3. Non-DBMS Files	61
5.4. Data View.....	61
6. Detailed Design.....
6.1. Hardware Detailed Design.....	62
6.2. Software Detailed Design	63
6.3. Network Detailed Design	71
6.4. Security and Privacy.....	71
6.5. Service Oriented Architecture / ESS Detailed Design.....	72
7. External System Interface Design

8. Attachment A – Approval Signatures	73
A.1. Identification of Technology and Standards.....	74
A.2. Constraining Policies, Directives and Procedures.....	74
A.3. Requirements Traceability Matrix	74
A.4. Packaging and Installation	74
A.5. Design Metrics	74

1. Introduction

The Genomic Information System for Integrated Science 2 (Genesis2) Technical Services application provides a central infrastructure to support the multiple Department of Veterans Affairs (VA) genomic medicine studies for facilitating recruitment and enrollment of Million Veteran Program (MVP) participants, automating most study-related logistics, including study enrollment, capturing clinical study data, consent, blood-sample tracking, and genomic-data storage. Genesis2 also provides the secure, analytical infrastructure necessary to conduct robust genomic and bioinformatics-related data management and data analysis. Genesis2 includes the planning and analysis of the VA-hosted computing cluster, as well as VA-approved cloud (or other) computing infrastructure, planning and analysis of the Internet 2 (or other WAN connectivity) for genomic data transmission, storage, and planning and analysis of the “big data” tools and systems.

The requested enhancements to the “As Is” functionality is needed to support the anticipated rapid expansion of the MVP.

Genesis2 is the name of the project. Builds are named Build 1, Build 2, Build 3 and so forth.

1.1. Scope

The Genesis Requirements Specification Document (RSD) defines the project scope; however, the initial requirements documented in the RSD have been updated in the Genesis2 project Requirements Traceability Matrix (RTM) and will be maintained in the Rational tool.

The following sections outline the scope for Genesis2.

1.1.1. Overall Scope of Genesis2 to Date

Figure 1 provides an overview of the overall scope of Genesis2 Builds 1 through 4.

	Target Release	Development Work Stream	Epics and User Stories	Key Features Available at the End of the Release
Base Year Releases	1	User Interface	E2.1, E2.2, E2.3, E3.1, E3.2	Requestors/ Data Managers can create a request Data Managers can approve, return, or reject Data Request with reasons specified
		Backend		
		Security Docs/Training	E5, E6	
	2	User Interface	E2.7, E2.8, E4.4	Requestor/Data Managers can add comments with or without advancing the workflow. Data Managers can copy tables from Source to Destination Databases. Support for PIV Card/Single Sign On using Windows Authentication Develop reusable Data Operations utilities.
		Backend	E4.1, E4.2	
		Security Docs/Training	E5, E6	
	3	User Interface	E4.5, E4.6	Support for Data Destination Manager (DDM) and Data Source Manager (DSM). The DSM prepares the data requested in one or more SQL server tables and informs the DDM. The DDM copies the tables over to a Genesis Destination Database including Schema Names Email notifications sent out.
		Backend	E4.3	
		Security Docs/Training	E5, E6	
Option Year 1 st Release	4	User Interface	E4.8, E2.1	General and Management Comment Support and Visibility Requestors or Data Managers can Copy or Cancel a Request Integration with Genesis OpenLDAP for User Management
		Backend		
		Security Docs/Training	E5, E6	

Figure 1: Genesis2 Scope Builds 1 - 4

Buils 5 and 7 Terminology Services Scope: A Researcher uses Terminology Services to locate the data elements they need by navigating a set of Concepts from many Ontologies like SNOMED-CT, HP, and NCIT. The goal of Terminology Services is to enable Researchers to start their searches with the clinical and bio-medical language they are familiar with, and arrive at concepts in these Ontologies. The concepts are attached to specific data elements in Genomic, Survey and Clinical data. Terminology Services provides a guided search capability to data elements of interest for research use. More details are provided in Section 1.1.2.

Build 6 Data Requests Scope: Build 6 builds on Data Request types supported in Build 4. First, Data Destination Managers are renamed Genisis Data Managers, and Data Source Managers are renamed as VINCI Data Managers. In Build 4, requests for data were sent to VINCI Data Managers and they fulfilled them (**VINCI Data Requests**). In Build 6, Genisis Data Managers and Researchers can send VINCI Data Managers data as sometimes needed (**Transfer to VINCI Requests**). More details are provided in Section 1.1.2.

Build 8 Terminology Services Scope: This release builds on the features included in Build 7. It includes a Tree View navigation pane for all the Ontologies included so far, starting with the Root node of each Ontology. It also includes a number of minor User Interface tweaks such as highlighting the changes suggested by a Researcher or a Terminologist, and System Administrator features such as Concept Card customization (different Ontologies having different sets of attributes displayed in the Concept Card) and Command Line Interface (CLI) additions and improvements. More details are provided in Section 1.1.2.

1.1.2. Release Scope

Table 1 provides the Scope of Genisis2 builds to date.

Table 1: Scope of Each Release to Date

Date	Revision	Description	Project Name	VA Department
02/2017	Build 1	<ul style="list-style-type: none"> Enable Researcher / Data Manager to create a request for Clinical, Genomic or Survey Data. Enable Data Managers to approve, return, or reject Data Request with reasons specified. 	Genomic Information System for Integrated Science 2 (Genisis2)	VA ORD (Office of Research and Development)
05/2017	Build 2	<ul style="list-style-type: none"> Enable Researcher / Data Manager to add comments with or without advancing the workflow. Enable Data Managers to copy tables from Source to Destination Databases. Develop reusable Data Operations utilities. 	Genisis2	VA ORD

Date	Revision	Description	Project Name	VA Department
06/2017	Build 3	<ul style="list-style-type: none"> Support for two kinds of Data Managers: Genisis Data Manager and VINCI Data Manager: <ol style="list-style-type: none"> The VINCI Data Manager prepares the data requested in one or more SQL server tables and informs the Genisis Data Manager. The Genisis Data Manager copies the tables over to a Genisis Destination Database. Email notifications are sent. 	Genisis2	VA ORD
09/2017	Build 4	<ul style="list-style-type: none"> General Workflow Additions: <ol style="list-style-type: none"> Copy Request Function – Ability to create a Draft Request from another a request that already exists. Cancel a Request. Ability to add Comments that are visible to everyone (General Comments) and only to management (Operational Comments). System Administrator Functions: <ol style="list-style-type: none"> Ability to add users to the system by synchronizing with the Genisis2 OpenLDAP directory and searching for a specific user. System Administrator Dashboard functions: Summary of Number of Users by Role, Request Metrics – Raw Numbers of Requests in Draft, Approved, Fulfilled, etc., and drilling down to list of requests of any one category, down to the details of any one request. 	Genisis2	VA ORD
11/2017	Build 5	<ul style="list-style-type: none"> Microservice APIs – Support the Terminology Services Programmer to define 4 microservice APIs to support the Search, Concept, Bookmark and Mapping services. Cohort Definition Functions – Support Researcher to (1) search for a concept and view the search results, (2) Concept Card, (3) bookmarked and un-bookmark concepts, and (4) label and un-label bookmarked concepts. Terminology Creating - Support Terminologist with Create, Modify and Delete functionalities through CLI. 	Genisis2	VA ORD

Date	Revision	Description	Project Name	VA Department
11/2017	Build 7	<ul style="list-style-type: none"> Cohort Definition Functions: Support Researcher with useful Search Results and a simplified Concept Card for a Data Element to define cohort. Multiple Roles Connection Functions: Support Researcher with Single Sign-On (SSO), iFrame, Data Elements Integration, and Baseline Survey Data Element List Display. 	Genisis2	VA ORD
12/2017	Build 6	<ul style="list-style-type: none"> Ability to Create a Data Transfer to VINCI Request by a Researcher or a Genisis Data Manager. Genisis or VINCI Data Managers to be able to Cancel a Request or Deny a Request. When the request is accepted, ability to do a Copy Table of the data to VINCI. Ability to update the Approved Studies database. Ability to execute a simple QC pipeline on data. Convert a table into a Pipe Separated Value (PSV) formatted file, do a data Quality Assessment Report, and do Date of Birth Validation. 	Genisis2	VA ORD
<u>01/2018</u>	<u>Build 8</u>	<ul style="list-style-type: none"> <u>Microservice APIs - Update Concept Card API and create Batch API service.</u> <u>Terminology Editing – The system displays the tab an edit was performed in following submission of the edit. A visual indicator showing that suggested edits to a Concept Card have been added by MVP.</u> <u>CRIA Process - Display the added information in the Concept Card.</u> <u>System Administration - Support Terminology Services Administrator with CLI Improvements, Concept Card Customization and User Role Authorization.</u> <u>Multiple Roles Connection - Graphical Display, Online Help for</u> 	<u>Genisis2</u>	<u>VA ORD</u>

1.2. User Profiles

The primary intended user population will be Researchers who are experienced in performing analysis on computational clusters.

The Genisis2 application is designed for its current users to provide, collect, manage, and analyze genomic data for research operations. Table 2 identifies the users and their responsibilities.

Table 2: User Profiles

Name of Users	Description	Responsibilities
Genisis System Administrator (Data Request application)	Performs system administration functions for the Genisis2 application	<ul style="list-style-type: none"> Assumes system administrator privileges to maintain Genisis2 application and network interfaces Monitors and tracks user access privileges for all users Imports relevant data from non-VA /third-party sources Performs application updates Inserts content changes, including documentation Deletes obsolete content Configures/reconfigures the computing environment Provides on-demand updates and content changes Manages job queue to create and configure environments Ensures that Genisis2 is maintained as an enterprise-level operation servicing the needs of MVP
Genisis Data Manager	Approves data request, new users, and provisions data	<ul style="list-style-type: none"> Performs system updates. Ensures Genisis2 complies with regulatory and security changes.
VINCI Data Manager	Prepares the data being asked for by the Researcher and delivers the data	<ul style="list-style-type: none"> Addresses questions and clarifications about the data being requested. Prepares the data in one or more SQL server tables. Informs the Genisis Data Manager that the data is ready for use.
Researcher	Request data; performs analysis	<ul style="list-style-type: none"> Uses Genisis2 Terminology Services to identify data elements they need using Ontologies/Concepts. Creates, modifies, and cancels data requests. Performs research on Genisis2 clinical, genomic, and survey data.
Administrator (Terminology Services)	Approves changes suggested by Researchers or Terminologists	<ul style="list-style-type: none"> Performs backend functions such as loading Ontologies, and adjusting the look and feel of individual ontologies through a Command Line Interface (CLI).

Name of Users	Description	Responsibilities
	<u>(equivalent to Genisis Data Managers above) to Ontology</u>	<ul style="list-style-type: none"> • <u>Approves or rejects changes suggested to Ontology attributes or alternative definitions preferred for MVP use.</u>

2. Background

2.1. Overview of the System

A Genisis2 Researcher uses other systems to query on data collected from various VA locations and public data sources to help form evidence for research proposals. They then formulate relevant data evidence and use it to submit a proposal to the Genisis2 governance process and any applicable Internal Review Board (IRB). **Genisis2 Data Request sub-system** is a workflow system used by the Researcher after obtaining this approval(s) to create and submit a request for detailed clinical, genomic and survey data. Genisis2 also provides a central repository for integrating clinical and genomic data collected from various VA locations and public data sources perform data cleaning operations on this data and use it for research.

The Genisis2 Data Request sub-system is supplemented by the **Terminology Services** sub-system that allows Researchers *to search for the data they need among the data that exists* using language that they are familiar with (Ontologies), navigate to data of interest and request them specifically.

A detailed set of requirements are specified in the Requirements Traceability Matrix (RTM) in the form of Epic stories that constitute the desired functionality for Genisis2. These requirements are decomposed to actionable items with acceptance criteria to produce the target solution design. The design specifications outlined in this document are based on the approved requirements and acceptance criteria and will serve as the guiding framework for the design and development for the Genisis2 system. Design elements in this document will also be mapped to requirements within the RTM to demonstrate how each requirement will be implemented.

The major design elements included in this document range from the automation of the Researcher request for clinical data, approval of the request, system usage and monitoring by system administrators and regulatory and security compliance associated with these operations.

2.2. Overview of the Business Process

Figure 2 provides a high-level view of the Genisis2 business processes. As depicted in the flow, Genisis2 is a recipient of external data from multiple sources, including clinical data from CDW/VINCI, consent forms and personal surveys from MVP sites, sequenced genotype data from MAVERIC, and other external sources, such as National Death Index (NDI) and Centers for Medicare and Medicaid (CMS).

The Honest Broker extracts Personally Identifiable Information (PII) from the data, and then stores the de-identified data in a secure database schema within the Genisis2 data warehouse. De-identified data is then integrated with MVP genotype data, survey data, and consent data into a schema, and moved into study marts for Researchers. This study mart also provides

Researchers a virtual workspace to perform analysis and run computations, and are only given access to the data that has been approved for their studies.

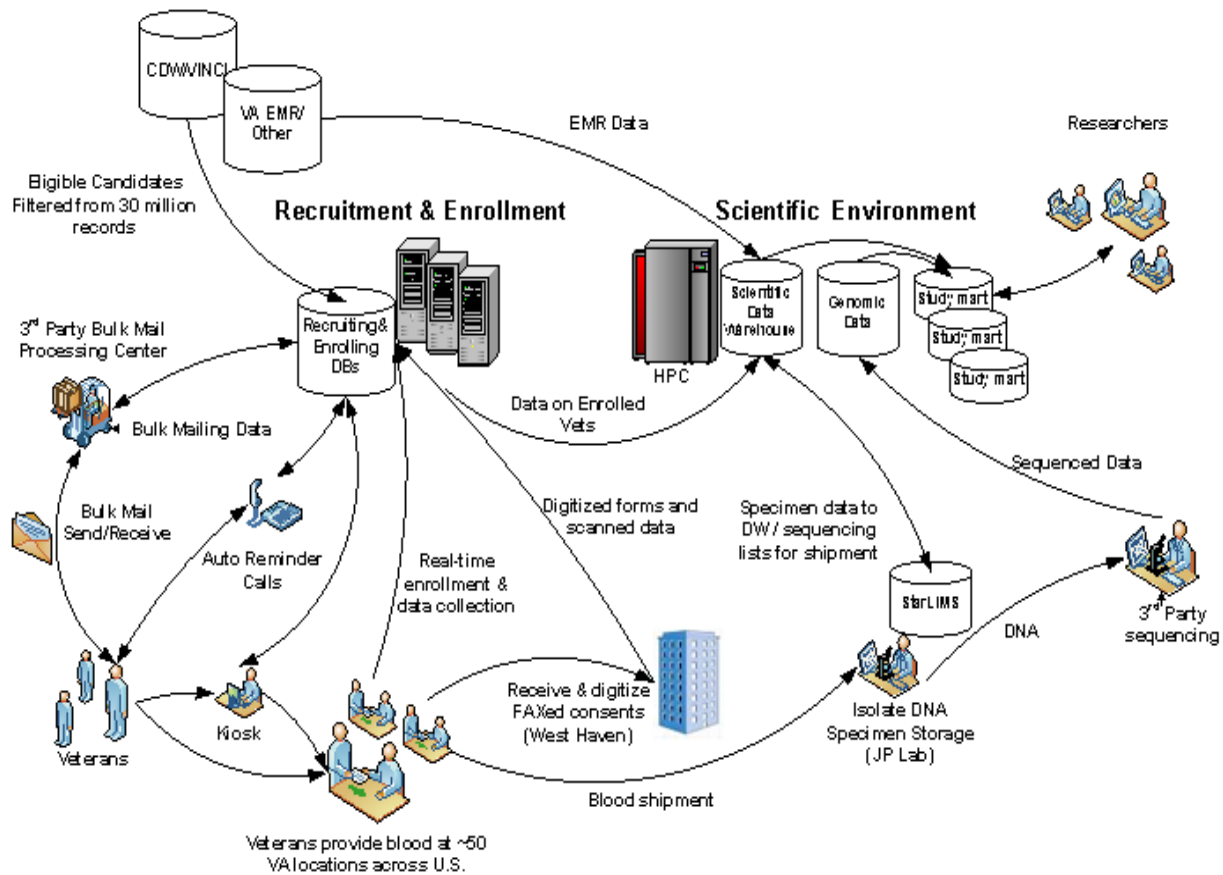


Figure 2: High-Level Genisis2 Business Process Flow

2.3. Overview of the Significant Requirements

The requirements identified in this section are not to replace the existing RSD, nor serve as the basis for the RTM. This section offers a high-level perspective to individuals who are not familiar with the requirements; and therefore, may need more background information in order to better understand the basis for the design.

These requirements are explained in more detail in the Genisis2 RSD, with preliminary requirements in the Genisis2 BRD and [Genisis2 RTM](#).

The Genisis2 project will execute the requirements described in the following sections.

2.3.1. Unified Genisis System Administrator/Researcher Dashboard

The Genisis2 requirements propose enhancements to the existing dashboard to provide a unified dashboard for Researchers and Genisis System Administrators. The enhanced dashboard will help Researchers request study data and track the status of their study data request throughout its lifecycle, and allow System Administrators to monitor the Genisis2 application.

2.3.1.1. Genesis System Administrator Improved Functionality and Dashboard

Enhancements for Genesis System Administrators are addressed in the Genesis2 requirements to support the rapid ramp-up of demand for system resources, user access, and setup of new study marts. A new Genesis System Administrator dashboard complements the improved functionality and supports the following features:

- Supporting Personal Identification Verification (PIV) authentication for all users to access the Genesis2 application. (This might be achieved through Windows Authentication through the Browser when someone has used a PIV card to gain access to the VA environment.)
- Providing role-based access for each Researcher account.

2.3.1.2. Genesis Data Manager Dashboard

Genesis Data Managers will have their own set of permissions that are lesser than those of Genesis System Administrators. They will not be able to add users or modify some Genesis2 application parameters such as workflow definitions and approval chains. They will have all the permissions needed for reviewing and approving Researcher data requests, performing data management activities, copying data, and performing quality control on them.

2.3.1.3. VINCI Data Manager Dashboard

VINCI Data Managers will have their own set of permissions that are not as extensive as Genesis Data Managers. They will not be able to add users or modify some Genesis2 application parameters like workflow definitions, approval chains, etc. They will have the permissions needed for reviewing and approving data requests coming from Genesis Data Managers, preparing the data asked for, and providing them in SQL server tables.

2.3.1.4. Researcher Dashboard

Researchers will have a new dashboard to submit requests for clinical, genomic or survey data, track and submit change requests, track status updates about their study, access VINCI tools, and receive notifications when there are updates to their requests.

2.3.2. Security Enhancements

Security enhancements are required to meet Federal mandates, for example, the addition of PIV for controlling access to VA systems by all VA users. This is achieved through Windows Authentication through the Browser when someone has used a PIV card to get into the VA environment in the first place.

In addition, new functionality is added to improve provisioning of ACLs to manage Researcher access to study marts. The new security enhancements will meet the two-factor authentication required for Authority to Operate (ATO) certification.

Genesis2 security will be reviewed for approval to upgrade from its current Temporary Authority to Operate (TATO) to the ATO certification through the Assessment and Authorization (A&A) process. A&A approval is a prerequisite for ATO certification. Genesis2 will undergo a series of scans and assessment tests to validate security of the software code, the database, and the firewall. The discovery of any “Critical” or “High” risks will be mitigated, and subsequent scans

will validate that the risks have been remediated. In addition, Genesis2 will undergo a level of penetration testing as needed by the requirements for an ATO. For Genesis2, Service Level Agreement Modification (SLAM) #16-051 is now in place to engage the support of Enterprise Operations (EO). EO will facilitate the ATO process and collaborate with Genesis2 management to ensure that Genesis2 achieves its ATO certification.

3. Conceptual Design

This section of the SDD provides details about the following topics:

- Conceptual Application Design Principles
- Conceptual Application Design
- Conceptual Data Design
- Conceptual Infrastructure Design

3.1. Conceptual Application Design Principles

These are the general principles used in the Conceptual Application Design:

- **Genesis2 Application Design needs to be Extensible and Flexible**
 - ***Flexibility*** – Genesis2 envisions a single workflow process handling Researchers’ requests and fulfilling them. As research efforts scale in number and variety of what they study, there may be many variations of these processes needed depending upon who the Researchers are and the purposes of their studies. Modular design that separates the workflow functions from the data handling functions will go a long way in achieving this needed flexibility.
 - ***Extensibility*** – Genesis2 Application design needs to be extensible so that it can work with other new applications, other new sources of data, other infrastructure back-ends, and capacity that may be used in the future for the MVP project. Modular design of the Genesis2 applications and loosely coupled components will help achieve this extensibility in the future. Genesis2 will support a RESTful Application Programming Interface (API). The Genesis2 User Interface will use this API for its implementation. Other VA systems that want to talk directly to Genesis2 can use the same approach. Documentation is provided and updated with each release of Genesis2.
- **Separate “Genesis2 Request Workflow” functions from Data Operations Functions**
 - ***Genesis2 is a Workflow Application.*** Workflow steps involve:
 - The creation of requests by Researchers.
 - Modification or cancellation of these requests.
 - Verification and validation of these requests by Genesis Data Managers (*Data Destinations*).
 - Receipt of data from VINCI/CDW or other *Data Sources*.
 - Data Quality Control of this data.
 - Copying the data to the Researcher’s Study Marts

- Informing the Researchers that their requests have been reviewed, approved/denied, or fulfilled (i.e., their data is ready for analysis.)

In that sense, the Genesis2 application needs capture the above steps in a workflow and is a *General Workflow Enabler*.

- ***Data Operations Management Enabler*** – Genesis Data Managers perform ETL Functions and data transformation functions for quality. The Genesis2 application will capture metadata about these operations but leave the actual cleansing of data to the Genesis Data Managers.
- ***Keep the Workflow and Data Operations Business Logic Separate*** – The application design will keep the request handling workflow business logic separate from the data operations business logic. As research efforts scale, this is critical if we want to keep the application flexible and extensible. This will help us accommodate new kinds of data operations in the future, especially those we do not foresee today.
- ***Associate and archive Data Management Logic with Study Mart Requests*** – The application design must let Genesis Data Managers do their work manually for a Researcher query or study request. However, the design must have a way to capture all the steps involved in that request along with Meta data about what the inputs of that data operation were, inputs, outputs, etc. The application should not care about what is used in ETL or what the query is, but allow archival of those utilities with descriptions of what they do, how they do it, and accommodate a library of such tools. For example, cleansing the Date of Birth field may involve a query/update script while removing PII using a script that talks to the Honest Broker system may be another utility. Also, allow the documenting of the sequence of data operations use. This will form a secure audit trail for analysis if need be later on.
- **Provide for a Semi-Automated Workflow with Manual Checkpoints**
 - ***Integrated Workflow/Document Management tool***. The application design will consider an ***Integrated Workflow tool*** to implement Genesis2 so that implementing the first type of workflow along with possible future ones are easy.
 - ***Manual Workflow with Automated Steps***. As of today, not the entire request handling workflow steps or data operations can be fully automated end-to-end. The Genesis Data Manager executes manual checkpoints to say, checking on study approvals before approving a data request. The approval and the workflow steps will be automated with data operation scripts created and libraries of those archived over time. In the future, these operations can be automated so that the workflow proceeds smoothly and stops only when a manual checkpoint is absolutely needed.
 - ***Tracking and Reporting of Workflow Progress for Transparency and Process Improvement*** – The application design will enable the tracking and reporting of the progress of a request since it is captured in a workflow at any time. This way, Researchers can have transparency in knowing exactly where their request for a query mart or a study mart is at any time. You can compare expected time intervals between steps with actual time taken for reporting and analysis and use it to improve processes.

A high-level flow of the system is represented in Figure 4, showing the pathways a request for data can take.

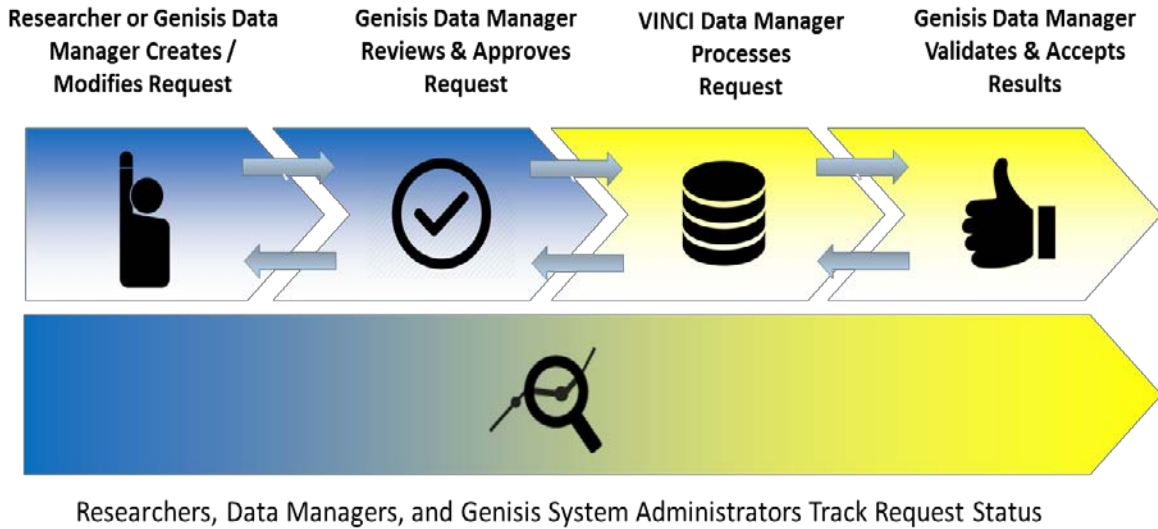


Figure 4: High-Level Data Request Flow Process

The key workflow that is central to the implementation of the high-level flow in Figure 4 is represented in the processes illustrated by Figure 5.

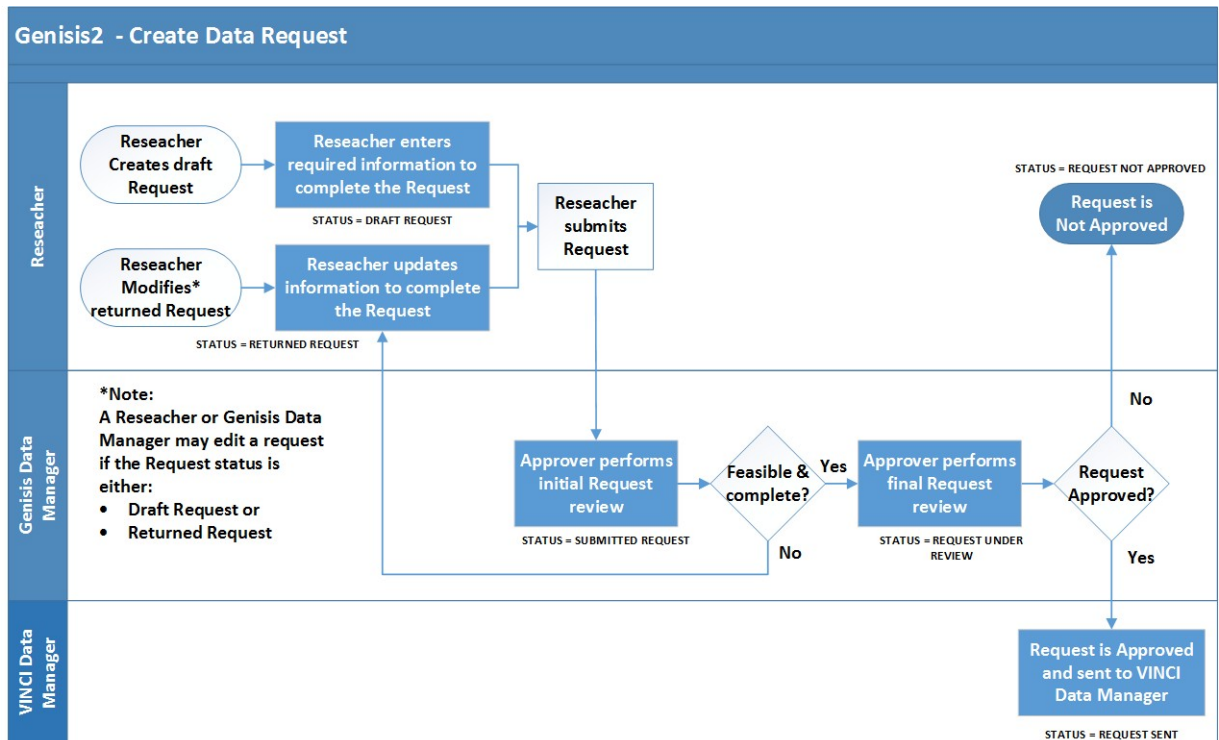


Figure 5: Data Request Flow

3.2.2. Genisis2 Terminology Services Process Flow

A high-level process flow for the Genisis2 Terminology Services is as shown in Figure 6. It represents the workflow for a Researcher to start searching for Data elements in Ontologies, in

terms that they are familiar with, ending up with data they need. Once they have put together the Data Elements that makes sense to them they go forward with executing queries that involved them.

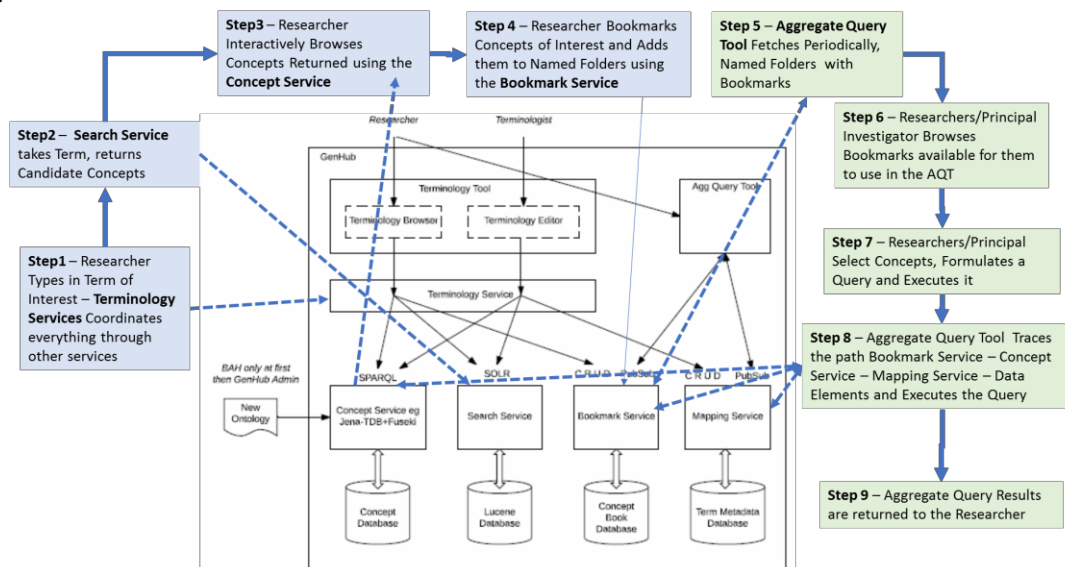


Figure 6. Genesis2 Terminology Services Process Flow

3.2.3. Application Context

Figure 7 depicts the Application Context Diagram (ACD) for the Genesis2 application.

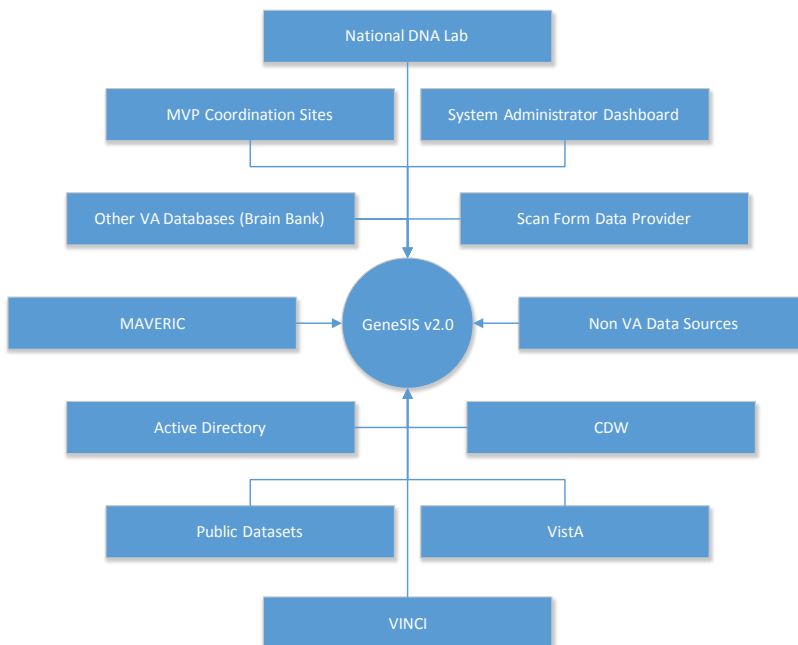


Figure 7: High-Level Application Context Diagram

Table 3 describes the information in the Application Context Diagram in four sections. Note that a single object (in the center of Figure 7) represents the system for which this design applies. Therefore, it is not referred to in Table 3.

Please note that this section outlines all the Data Sources that Genisis2 may interface with in the future, not necessarily within the scope of this project now.

Table 3: (Grouping) Application Context Description

ID	Name	Description	Interface Name	Interface System
Scan Form Data Provider	Scan Form Data Provider	Scan form data providers process form data from study participants and present forms in tabular form. VA Westhaven Research Lab and Synovate, Inc. are two current scan form data providers.	Form	Genisis2
CDW	VA Corporate Data Warehouse	The VA Corporate Data Warehouse presents data through relational database views.	CDW	Genisis2
VistA	Veterans Health Information Systems and Technology Architecture (VistA) electronic health records system	VistA is the VHA electronic health record system.	VistA	Genisis2
Public Datasets	Public Data sets	Genomic annotations and ontologies are publicly available and necessary for Genomics research.	Public Datasets	Genisis2
Active Directory	VA Active Directory	The VA Active Directory provides authentication services to Genisis2.	VaAd	Genisis2
Brain Bank	Brain Bank	The Brain Bank is a bio specimen repository maintained by the VA.	BrainBank	Genisis2
National DNA Lab	VA National DNA Laboratory	The VA National DNA Laboratory is the repository for bio specimens and DNA collected from the VA Genomic Medicine Program.	DnaLab	Genisis2
VINCI	Veterans Informatics and Computing Infrastructure	VINCI is a partner with the Corporate Data Warehouse (CDW) and hosts all data available through CDW as well as some unique data.	VINCI	Genisis2.

ID	Name	Description	Interface Name	Interface System
MAVERIC	The Massachusetts Veterans Epidemiology Research and Information Center	MAVERIC is an interdisciplinary research and development organization with the goal of creating a learning healthcare system within VA through application of research resources and methodologies to important clinical problems.	MAVERIC	Genisis2
Non-VA Data Sources	Non-VA Data Sources	National Death Index (NDI), Centers for Medicare and Medicaid Services (CMS) and other national data.	Non-VA Data Sources	Genisis2
MVP Coordination Sites	Million Veteran Program Coordination Sites	MVP Surveys and Consent Forms.	MVP	Genisis2
Genisis System Administrator Dashboard	Genisis System Administrator Dashboard	Dashboard allows Researchers to request study data and track the status of their study throughout the lifecycle, and allow system administrators to monitor the Genisis2 application.	Genisis System Administrator Dashboard	Genisis2

Interfaces External to OI&T

ID	Name	Related Object	Input Messages	Output Messages	External Party
N/A	Public Datasets	Public Data Sets	Data records encoded in structured text. The data is publicly available reference data, which is cached in Genisis2 for use in genomic analysis.	N/A	National Institutes of Health, Universities, Government-sponsored research institutions.

Interfaces Internal to OI&T

ID	Name	Related Object	Input Messages	Output Messages	External Party
N/A	Form	Scan Form Data Provider	Scan form data is presented as a simple database table with one row per subject.	N/A	VA Westhaven Research Lab

ID	Name	Related Object	Input Messages	Output Messages	External Party
N/A	CDW	CDW	CDW subset data is presented as database views from the Austin Information Data Center.	N/A	AITC
N/A	ActiveDirectory	Active Directory	LDAP query responses. This is part of VA Windows Authentication System.	LDAP queries to lookup users to the Genesis2 application. This is part of VA Windows Authentication System	VA Active Directory Infrastructure
N/A	BrainBank	Brain Bank	The Brain Bank inventory records is presented as a set of relational database views.	SQL queries to look up bio specimens stored within the Brain Bank	VA Brain Bank
N/A	DNALab	VA National DNA Lab	The National DNA Lab presents bio specimen data as a set of relational database tables.	SQL queries to look up bio specimens stored within the National DNA Lab	VA National DNA Lab
N/A	VINCI	Veterans Informatics and Computing Infrastructure	VINCI data sets are CDW extractions from VistA, MedSAS in SAS and SQL, and DSS in SAS and SQL.	SQL queries to look up VINCI data sets	AITC
N/A	MAVERIC	The Massachusetts Veterans Epidemiology Research and Information Center	Genotype data.		VA Boston Healthcare System

ID	Name	Related Object	Input Messages	Output Messages	External Party
N/A	Non-VA Data Sources	Non-VA Data Sources	National Death Index (NDI), Centers for Medicare and Medicaid Services (CMS) and other national data.		NDI, CMS and other
N/A	MVP Coordination Sites	MVP Coordination Sites	MVP Surveys and Consent Forms.	N/A	Nationwide
N/A	Genisis System Administrator Dashboard	Genisis System Administrator Dashboard			Virtual

Externally Shared Data Stores

ID	Name	Data Stored	Owner	Access
N/A	N/A	N/A	N/A	N/A

3.2.4. High-Level Application Design

The High-level Application Design depicted in the Figure 8 identifies the major components of the application and the relationships of the major application components to each other and to surrounding applications.

Please note that the following descriptions are providing only the CONTEXT for the Genisis2 system. The other applications such as Honest Broker, Investigator, Governance, Metadata and Administrator are all under the scope of other development projects, many of them in development as of now. That is the reason for Not Applicable (NA) or to be Determined (TBD) entries in the table, since their exact status as of this update is not clearly known at the time of this release.

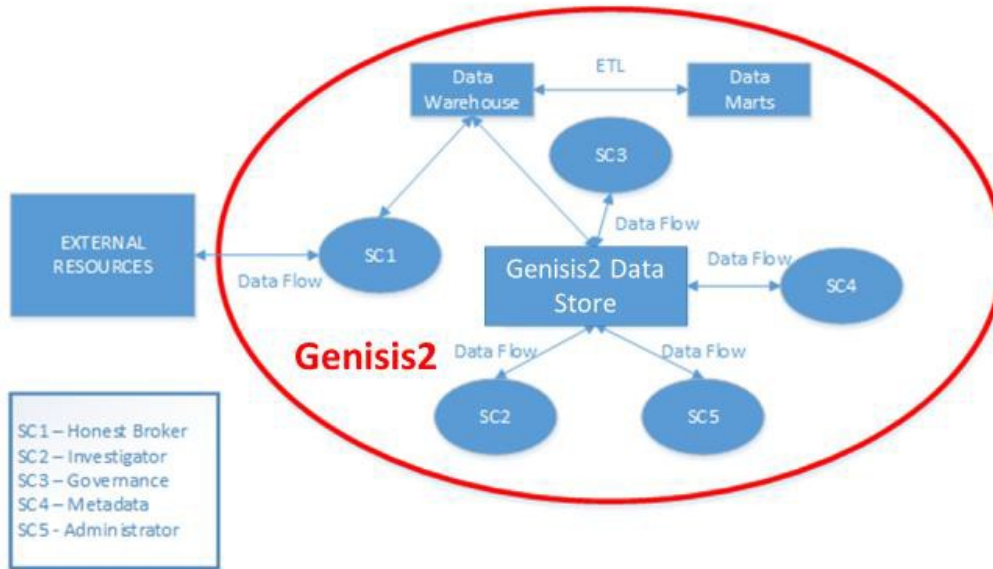


Figure 8: High-Level Application Design

Tables 4 and 5 provide information regarding these components.

Table 4: Objects in the High-Level Application Design

ID	Name	Description	Service or Legacy Code	External Interface Name	External Interface ID	Internal Interface Name	Internal Interface ID	SDP Sections 1&2
N/A	External Resources	External Data	N/A	Legislative Information Management System (LIMS) Backend, West Haven, Synovate, CDW, VistA, VINCI, MAVERIC, MVP Sites, Public Datasets	TBD	N/A	N/A	N/A
N/A	Data Warehouse	Survey & bio-specimen data	N/A	N/A	N/A	N/A	N/A	TBD

ID	Name	Description	Service or Legacy Code	External Interface Name	External Interface ID	Internal Interface Name	Internal Interface ID	SDP Sections 1&2
N/A	Data Mart(s)	Study specific survey & bio-specimen data	N/A	N/A	N/A	N/A	N/A	TBD
SC 1	Honest Broker	The Honest Broker receives external data and uploads it to the data warehouse	N/A	N/A	N/A	Honest Broker	SC1	TBD
SC 2	Investigator	Interface to create and manage Genisis studies	N/A	N/A	N/A	Investigator	SC2	TBD
SC 3	Governance	Maintain Study States	N/A	N/A	N/A	Governance	SC3	TBD
SC 4	Metadata	Metadata	N/A	N/A	N/A	Metadata	SC4	TBD
SC 5	Administrator	Provides System Administrator functionality	N/A	N/A	N/A	Administrator	SC5	TBD

Table 5: Internal Data Stores

ID	Name	Data Stored	Steward	Access
Genisis Data Store	Genisis	Stores all data to support all Genisis2 application components.	Genisis Web Applications	Create, Read, Update, and Delete

3.2.5. Application Locations

Table 6 specifies the locations where the Genisis2 application components are hosted and Table 7 provides the users.

Table 6: Application Locations

Application Component	Description	Location at Which Component is Run	Type
LIMS Backend	MVP bio sample, genomic data	DNA Repository Lab at MAVERIC, Boston, MA	Business Logic / Data Logic / Interface Code
Consent Form	MVP Consent forms	VA West Haven, CT	Business Logic / Data Logic / Interface Code
Synovate	MVP Survey Form Data	VA West Haven, CT	Business Logic / Data Logic / Interface Code

Application Component	Description	Location at Which Component is Run	Type
CDW	Clinical data	VA Austin, TX (AITC)	Business Logic / Data Logic / Interface Code
VistA	TBD	VA Austin, TX (AITC)	Business Logic / Data Logic / Interface Code
Data Warehouse	Survey & bio-specimen data	VA Pittsburgh Information Technology Center	Data Logic
Data Mart(s)	Study specific survey & bio-specimen data	VA Pittsburgh Information Technology Center	Data Logic
Honest Broker	The Honest Broker receives external data and uploads it to the data warehouse	VA Pittsburgh Information Technology Center	Interface Code
Investigator	Interface to create and manage Genisis studies	VA Pittsburgh Information Technology Center	Presentation Logic / Business Logic
Governance	Maintain Study States	VA Pittsburgh Information Technology Center	Business Logic
Metadata	N/A	VA Pittsburgh Information Technology Center	Interface Code
Administrator Dashboard	Provides System Administrator functionality	VA Pittsburgh Information Technology Center	Presentation Logic / Business Logic

Table 7: Application Users

Application Component	Location	User
Recruitment & Enrollment applications	50+ locations across the country. All major VA hospitals	Site Coordinators, Lab techs, MVP staff, report writers
Scientific Environment	Across the US	Scientists, MVP staff, Researchers

3.2.6. Preliminary Architecture

3.2.6.1. Architecture for Genisis2 Data Workflows

Figure 9 presents the architecture for handling Genisis2 Data Requests Workflows and the other external components with which the application interacts.

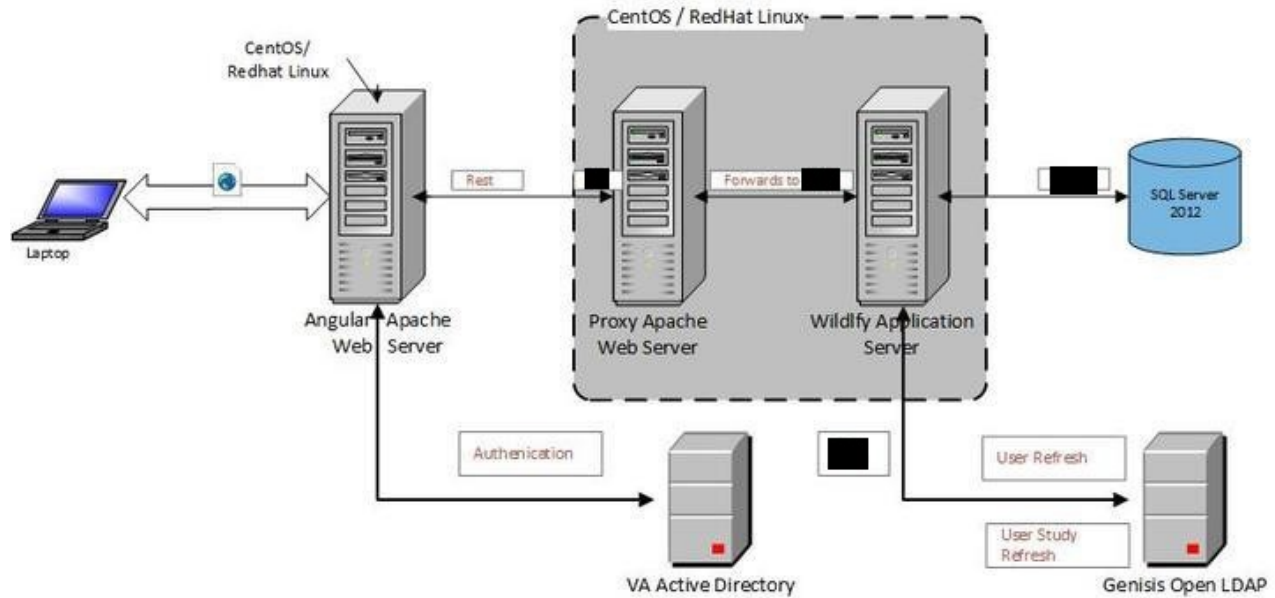


Figure 9: Server Architecture for Genesis2 Data Request Workflows

The Genesis2 architecture consists of the following components:

- **Genesis2 Core Application** – The Genesis2 core application will be implemented using an integrated Workflow tool such as Activiti. This core application will run in the Web/ Application Server and will consist of:
 - **Genesis2 Workflow Process Manager** – This will implement the functions needed by the different types of users of this system – Researcher, Genesis Data Manager, VINCI Data Manager and Genesis System Administrator. Each user will have their own dashboard and access to functions that they are authorized to perform. The dashboard will be their login page showing status of their requests or workflow steps, tasks they need to perform and Notifications from various activities. The workflow tool selected will have the appropriate workflow design and business rules implementation capabilities within it. Genesis2 will use these capabilities to design and implement the initial workflow and business rules in force today.
 - **Notifications Management:** – Requirements that need email or mobile/text message notifications will be handled by existing VA communications management capabilities. Genesis2 will integrate programmatically with these communication mechanisms to create and send out these notifications or text messages.
- **Data Operations and Tools** – As outlined in the Conceptual Application Design Principles section, Genesis Data Managers would be performing their data cleanup, approval and movement operations using scripts from command line interfaces. *Genesis2 will allow them to be started from the application as needed.*

3.2.6.2. Terminology Services Workflows

Figure 10 presents the architecture for handling Terminology Services workflow and the other external components with which the application interacts.

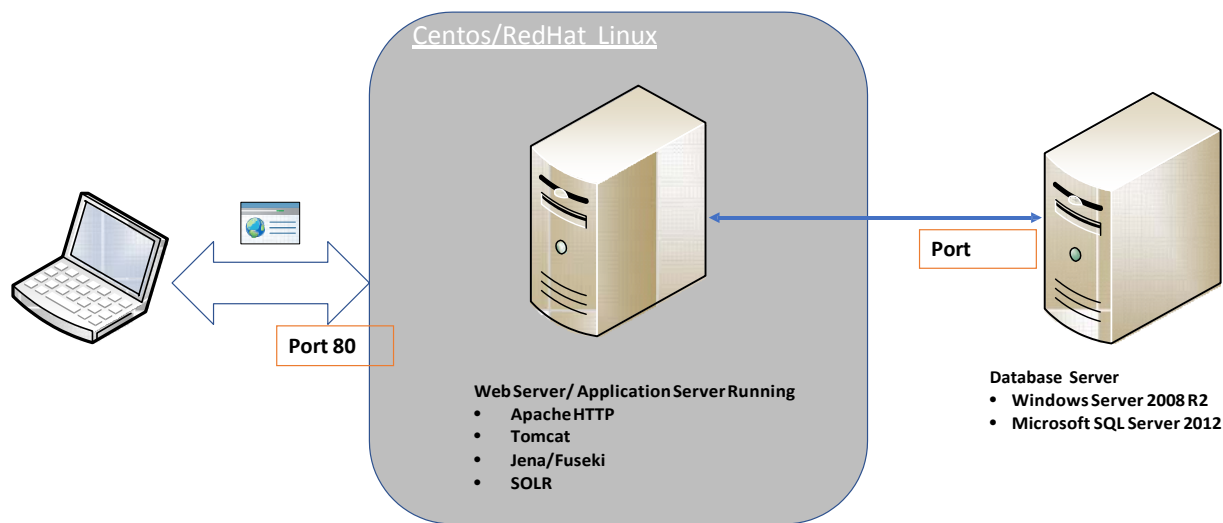


Figure 10: Terminology Services Server Architecture

The Terminology Services architecture consists of the following components:

- **Terminology Services Application** – The Terminology Services core application is implemented using SOLR search services and Jena/Fuseki Triple Store and Query capabilities.
 - **Notifications Management:** – Requirements that need email or mobile/text message notifications will be handled by existing VA communications management capabilities. Genesis2 will integrate programmatically with these communication mechanisms to create and send out these notifications or text messages.
- **Database Server** – All database storage and querying needs will be implemented using a Microsoft SQL Server 2012 software component running on a server which has the Microsoft Windows Server 2008 R2 software as the Operating System.

3.3. Conceptual Data Design

3.3.1. Project Conceptual Data Model

The conceptual data model for the Genesis2 application is based on the central entity called Request, as shown in Figure 11. Each Request has attributes for the person who created the request (Researcher), the type of request (either aggregate data or detail data), when the request was created, when it was fulfilled, etc. Each request can also have pointers to data tables or flat files within the data file repository (e.g., either the Genesis landing Zone or a Researcher Study Mart) that satisfied the request. Importantly, the status history of the request is tracked as the request progresses through its workflow. Each transition, (e.g., from submitted by Researcher to approved by Genesis Data Manager) is recorded (along with the person who made the transition), when the transition occurred, and any additional comments or notations. The request may have comments or notations added without a status change; the person that added the notation and the date the notation was added would also be recorded.

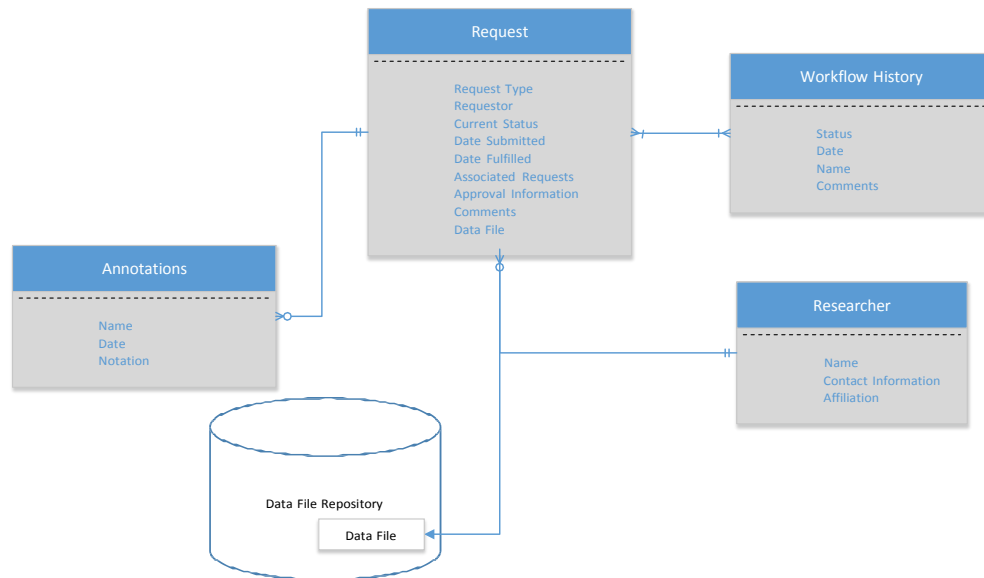


Figure 11: Project Conceptual Data Model

3.3.2. Database Information

Table 8 lists the existing database inventory.

Table 8: Database Inventory

Database Name	Description	Type	Steward
Genesis_Activiti_DB	All workflow related data is handled in tables created by the Activiti workflow system. Contains workflow data only.	Relational Database	ORD
Genesis_DB	The Main Genesis Database that handles all data not handled by the Workflow System above (like Comments table).	Relational Database	ORD
Jena/Fuseki	The Main Terminology Services Triple Store database that stores all Ontologies / Concepts that have been loaded to date.	Triple Store	ORD

Database Name	Description	Type	Steward
TS_DB	The Terminology Services Database that deals with User Specific activities such as Bookmarks, Labels, etc.	Relational Database	ORD

3.3.3. User Interface Data Mapping

Table 9 provides the Database Mapping.

Table 9: Database Mapping

User Interface Type	Description	Database Mapping	Additional Comments
Researcher/Genisis Data Manager Landing Page/ Request Screens	Researcher or Genisis Data Manager looking at a status of all their requests, drilling down into one request or creating a new request for data.	Genisis_DB/ Genisis_Activiti_DB	All workflow related data is handled by the Genisis_Activiti_DB automatically while all other data go into tables in the Genisis_DB
Data Table Copy Statuses	When Research Data is ready and is being copied from one Data Source to the Genisis Databases, statuses of the copy are logged.	Management_Table in the Genisis_DB Database	Number of Rows being copied and a Checksum calculated both BEFORE the copy and AFTER the copy and stored in this log. If these do not match, the table is dropped with a Table not copied error message and also logged in this table.
Terminology Services User Specific (Researchers, Terminologists) data such as Bookmarks, Labels, etc.	When users bookmark concepts for their use and create labels to classify them.	TS_DB is the database that handles storage and retrieval of this data	Bookmarks, Labels

3.3.4. Application Screen Interface

The following sections provide high-level Application Screen designs as of Build 86. Where features are not yet implemented, mockups at the level of detail currently available are provided.

3.3.4.1. Genesis Data Request Workflows

3.3.4.1.1. Researcher Landing Page/Dashboard

Figure 12 provides a screenshot of the Researcher Login/Dashboard. From this view, a Researcher can track the progress of their requests.

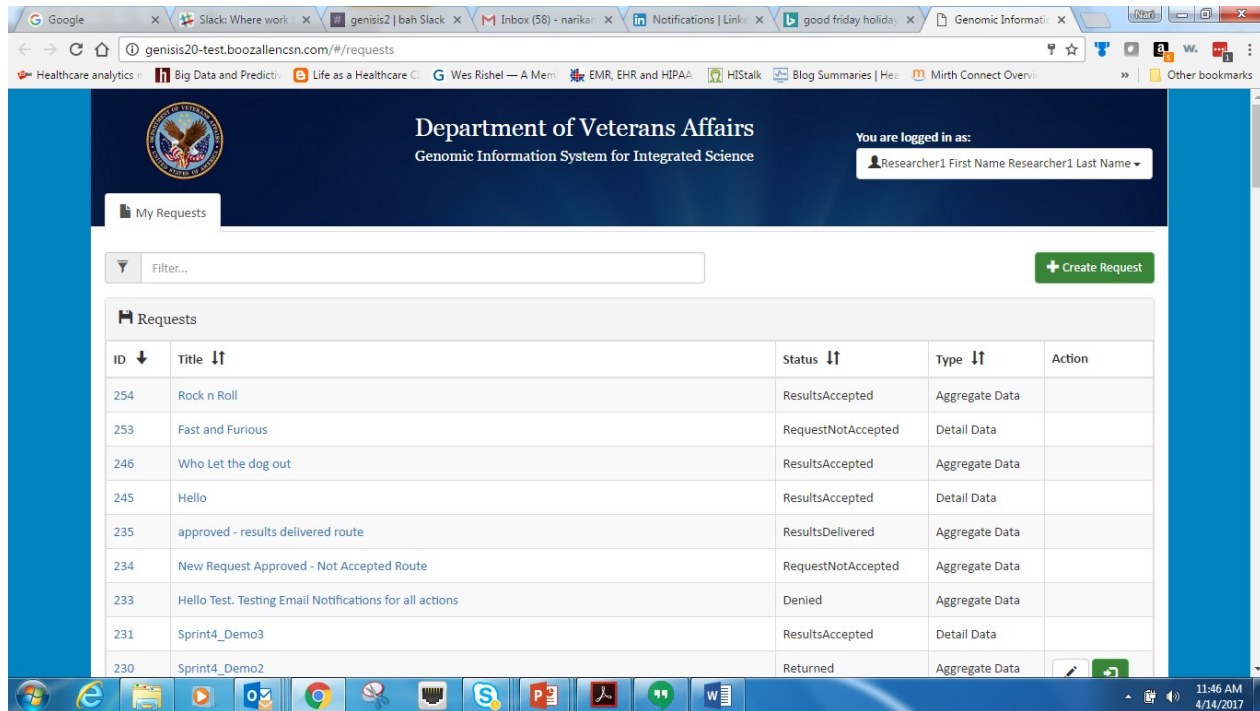


Figure 12: Researcher Landing Page/Dashboard Design

Researcher Dashboard: The Researcher Dashboard is also the landing page when a Researcher logs in using 2-factor authentication. It shows in various widgets the Tasks that are waiting for the Researcher to perform, request status changes and notifications from Genesis2, and other external components integrated with Genesis2.

Request Operations Tab: The Request Operations tab leads to options that can help a Researcher create a new request for data, modify a submitted request (internal state should be modifiable; for example, if a request has already been submitted to VINCI, it may not be modifiable) or cancel a request. **(Future build feature: will be part of this screen, but not part of Build 86.)**

Data Operations Tab: This tab allows a Researcher to open a current request and store information about any data operations they may be performing. For example, if they have developed a script to query a study mart or a query mart for some information or cleaning up some data, they can associate this script or query with a request, store additional metadata such as description, inputs, outputs, or processing changes intended. **(Future build feature: will be part of this screen, but not part of Build 86.)**

3.3.4.1.2. Genisis Data Manager/VINCI Data Manager Landing Page/Dashboards

Figure 13 is a screenshot of the Genisis Data Manager Login/Dashboard. From here, Genisis Data Managers can follow up on the status of the requests they are managing or perform actions such as move the workflow to the next step in the process. Alternatively, they can copy data from remote databases to the Genisis2 Database (Landing zone) for further processing before sending them off to Researcher study marts.

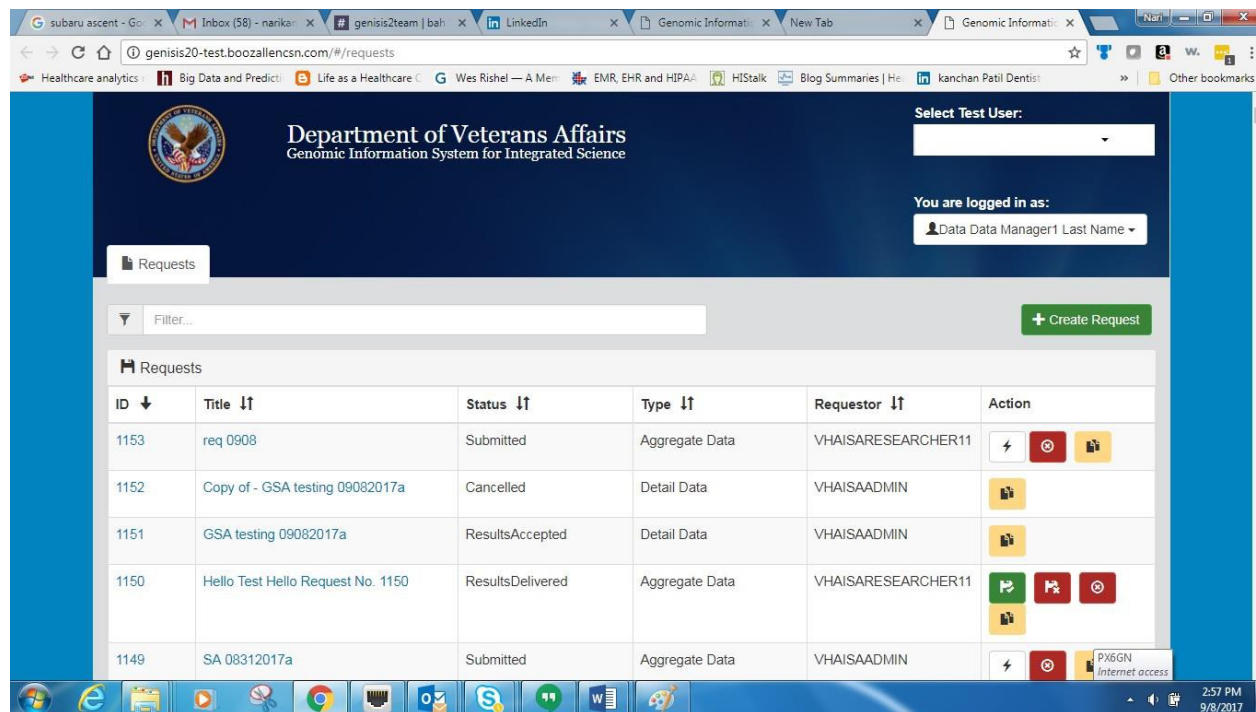


Figure 13: Genisis Data Manager Landing Page/Dashboard Design

Genisis Data Manager Dashboard: This is also the landing page when a Genisis Data Manager logs in using 2-factor authentication. This shows in various widgets, the Tasks that are waiting for the Genisis Data Manager to perform, request status changes and notifications from Genisis and other external components integrated with Genisis. The VINCI Data Manager Dashboard is very similar to this; however, the VINCI Data Manager does not have the same access as the Genisis Data Manager so some of the buttons will not be available to the VINCI Data Manager.

3.3.4.1.3. Genisis System Administrator Landing Page/Dashboard

Figure 14 is a *preliminary mock-up/wireframe of the Genisis System Administrator Landing Page/Dashboard. Once the preliminary design is approved and formal design is complete, this section will be updated again.*

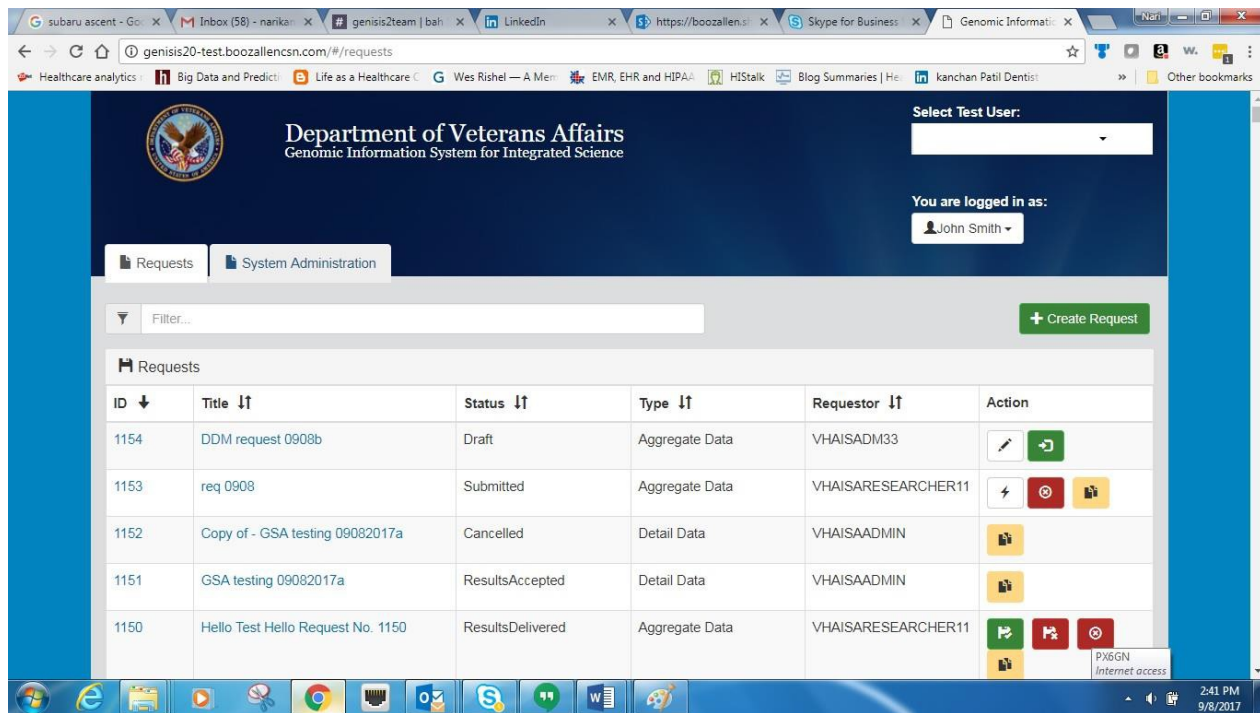


Figure 14: Genisis System Administrator Landing Page/Dashboard

Genisis System Administrator Dashboard: This is also the landing page when a Genisis System Administrator logs in using 2-factor authentication. This landing page has two main tabs – Requests and System Administration. The Request tab enables all of the functions as that of any of the other roles – Researcher, Genisis Data Manager or the VINCI Data Manager can perform. As a super user, a Genisis System Administrator is able to perform all of these activities. The Systems Administrator tab will lead to two other tabs – User Management (Statistics about how many users are currently registered, broken down by role) and Request Metrics (How many Requests are in each stage of the workflow with the ability to drilldown to any one request and see all its details).

Manager Workflow Tab: This tab leads to features that allow a Genisis System Administrator to enable additional workflows. This feature may be seldom used after the initial implementation, but is available to the Genisis System Administrators only, if needed. **(Future Release feature: will be part of this screen, but not part of this Release.)**

3.3.4.2. Terminology Services Workflows

3.3.4.2.1. Terminology Services Landing Page

Figure 15 is a screenshot of the Terminology Services landing page when a Researcher or a Terminologist accesses that page.

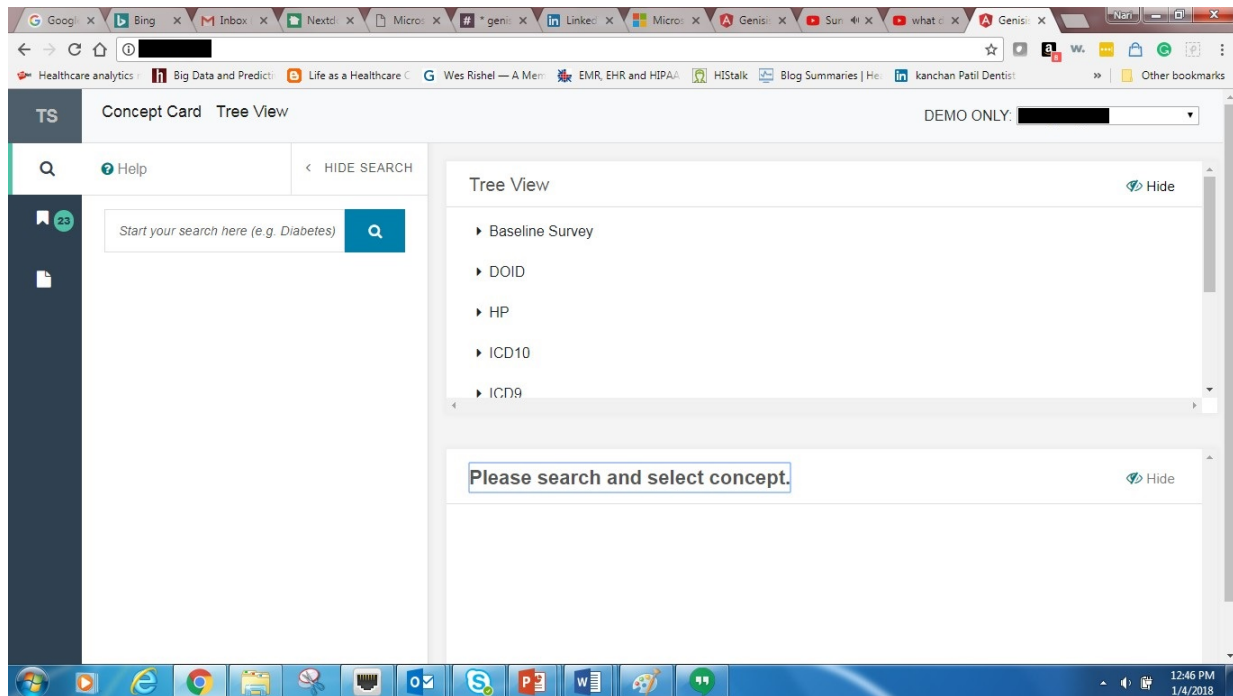


Figure 15: Terminology Services Landing Page

Terminology Services consists of the backend loading a number of standard Ontologies like SNOMED-CT, NCIT, and DOID. These Ontologies make available to the user a number of Clinical and Genomic Concepts for them to start their search for Data Elements they might be interested in. For example, if their research deals with Type I Diabetes, they start searching for Diabetes and navigate their way to data elements that may be attached to concepts that they are interested in. Figure 15 shows the landing page when users access Terminology Services.

Terminology Services users can navigate using the icons in the left-hand pane, representing functionality to Search Concepts (the magnifying glass), Bookmark Concepts (the bookmark icon) and display a list of Batch Changes (lower text-list icon – suggested changes to Ontology attributes that make more sense in the context of the MVP project).~~Users type in the terms they are interested in within the search box at the top left side of the landing page.~~

3.3.4.2.2. Search Results Page

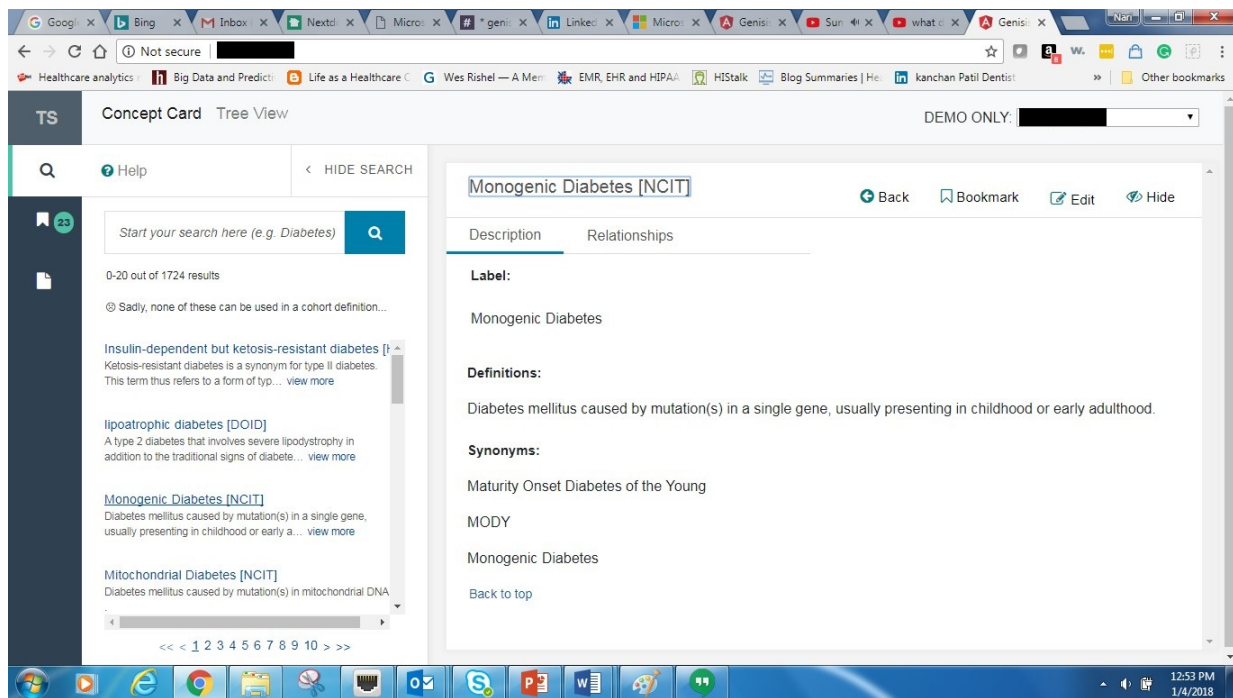


Figure 16: Search Results Page

Figure 16 shows a user typing in “Diabetes” in the Search box and concepts from all ontologies that deal with Diabetes are shown in the Search Results box on the left. Each result is shown with the Concept Title and some number of characters in the definition. The Ontology from which a concept was picked up is also displayed next to the Concept Name like [NCIT] or [DIOD].

The user can choose a concept that they want to explore further and click on the Concept name. This takes them to the Concept Card display in the next screenshot.

3.3.4.2.3. □ □ □ □ □

Concept Card Details

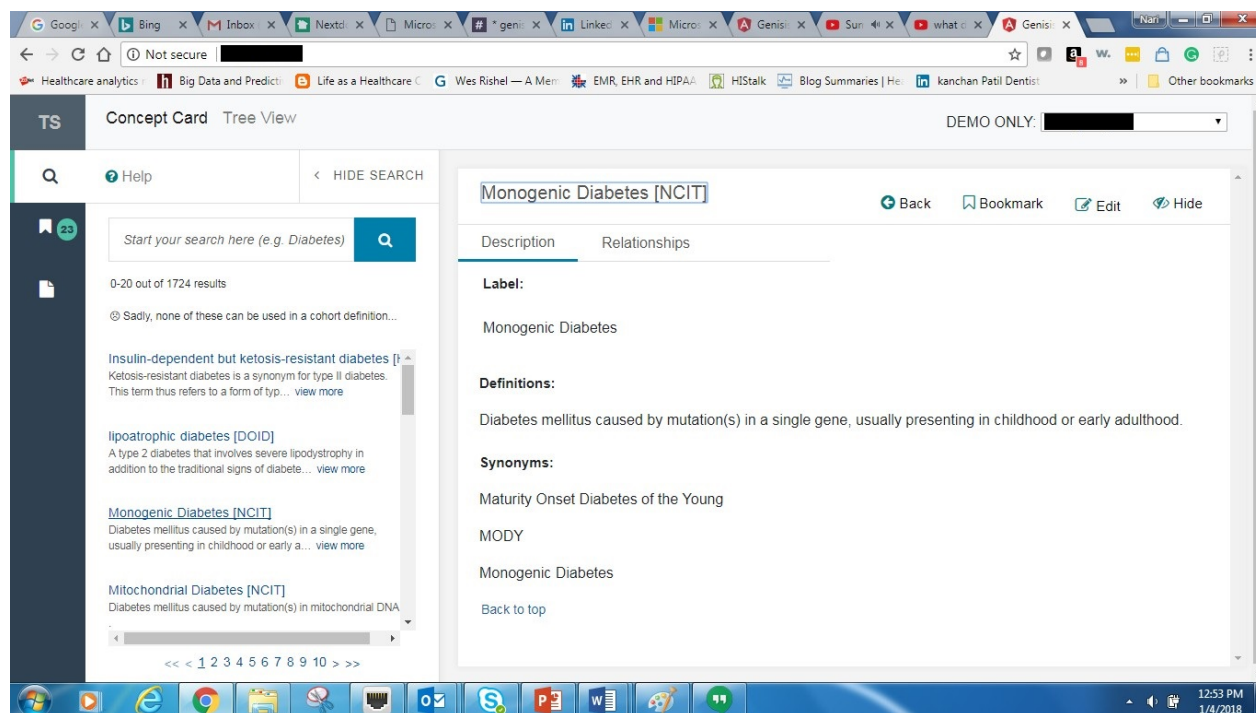


Figure 17: Concept Card Page

Figure 17 shows the Concept Card of one of the concepts selected for detailed display. It shows various properties of this concept under different tabs such as Description, Relationships and Metadata. The user navigates these tabs and to determine if this concept is the one they are looking for. If so, then the user can Bookmark it as shown in the next screenshot.

Relevant information regarding a Concept is displayed on the Concept Card (right pane). In this example from the NCIT ontology, the “Description” tab displays the Concept’s label and definition. A Terminology Services user can click on the “Relationships” tab to navigate to additional information.

Please note that Concept Cards have a unique tabbed format based on the information associated with a given ontology. In this example, the “Relationships” tab is configured to display a list of clickable hyperlinks representing hierarchically related ontological concepts (e.g., “Super” or “Sub” classes of the Concept displayed).

3.3.4.2.4. ☐ ☐ ☐ ☐ ☐ ☐

Bookmark a Concept

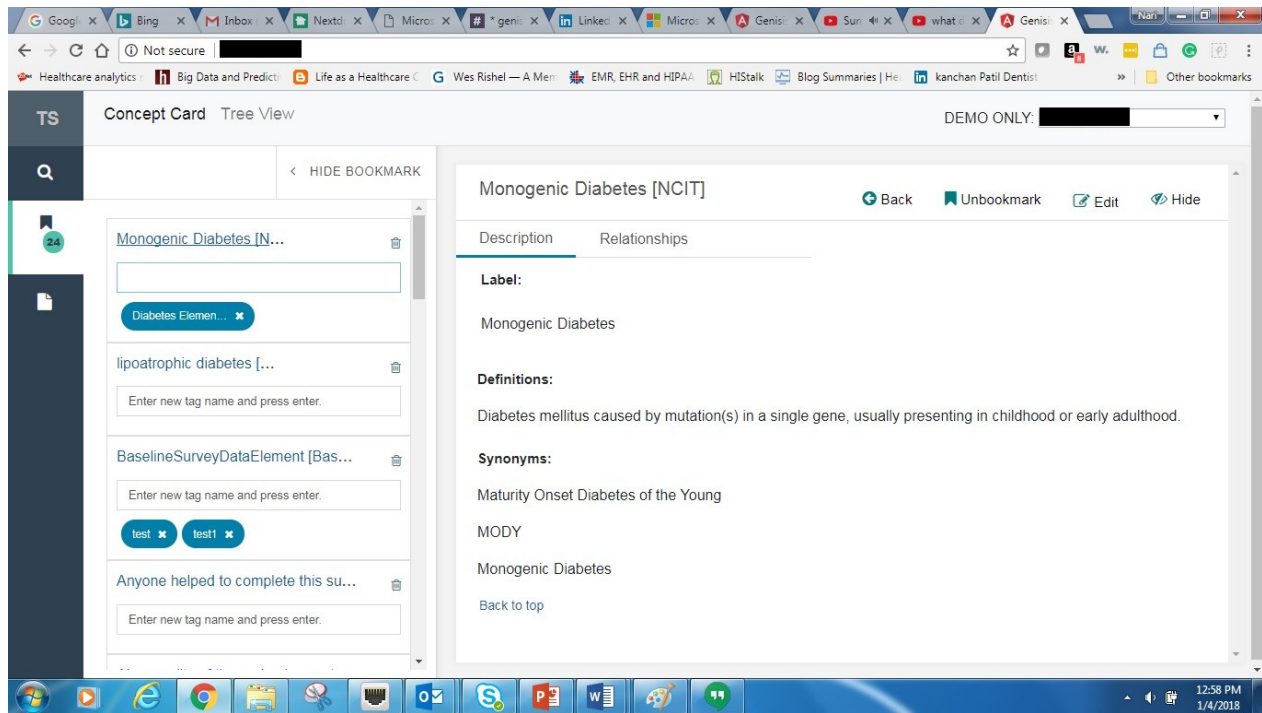


Figure 18: Bookmark a Concept Page

Figure 18 shows a page where the user bookmarks a Concept that they find of interest. If Data elements from MVP Clinical, Survey and Genomic data are attached to this concept, users can utilize the collection of bookmarked concepts that are of interest to them to query the data further for their research. The Bookmark Icon on the black bar on the left enables them to Bookmark the selected Concept. It also enables them to attach Labels to their bookmarks (shown in dark ovals in the middle panel). Labels enable a Researcher to classify their bookmarks and search for them at a later time.

3.3.4.2.5.□□□□□□

Search Bookmarks by Label

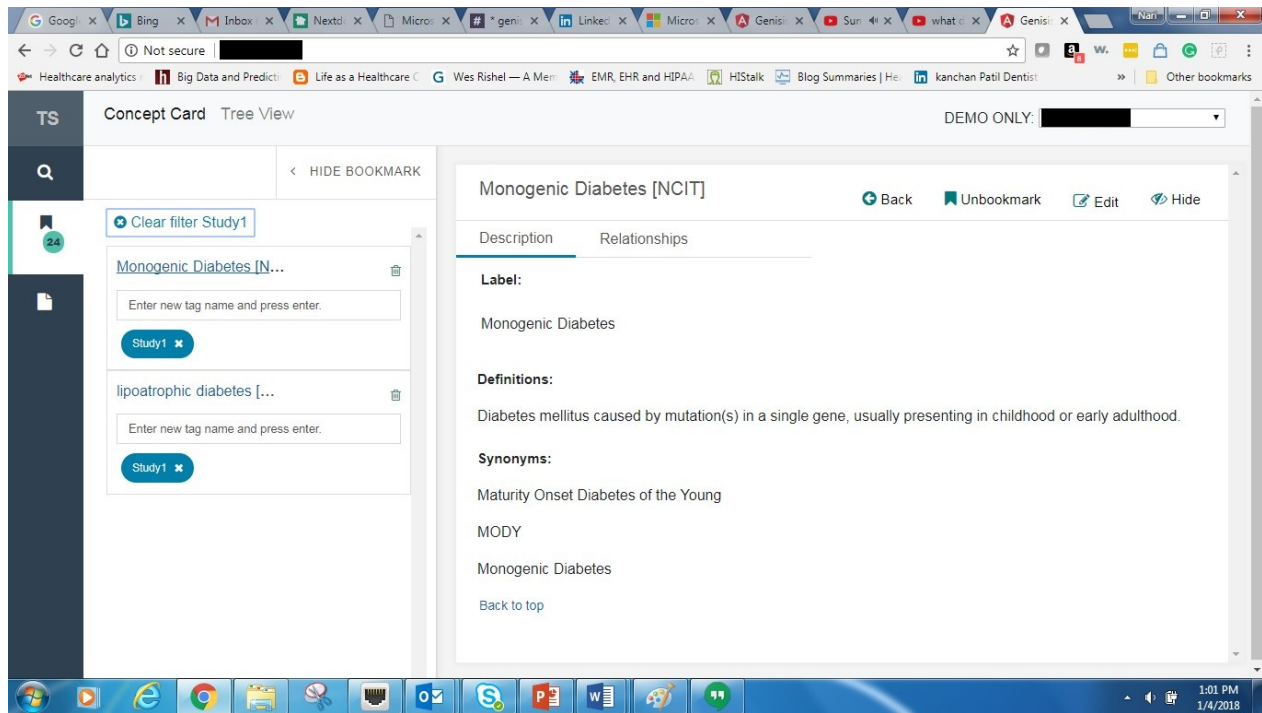


Figure 19: Search Bookmarks by Label

Figure 19 shows how a user can search for a bookmark by label. By clicking on a label of interest (“Study1” in the figure above), all bookmarks that have that label are filtered and shown in a list on the left as above. The user can select any one of the bookmarks and see the details. The button “Clear Filter Study1” when clicked takes the user back to seeing all of the bookmarks they made.

3.3.4.2.6. ■■■■■ **Browsing Data Elements**

The user can click on the List Icon on the left panel. This displays the data elements list in the central panel as shown in Figure 20. Selecting any one of the Data Elements displays the Concept Card in the right panel showing all the details of that data element. Each data element could be potentially mapped to one or more Concepts in the various Ontologies loaded into the system. By starting the search for Concepts by typing in Terms in the Search Box, the user can navigate down to Concepts that have Data Elements attached to them. They can bookmark these Concepts and use them to create a Cohort of VA MVP Volunteers for their research.

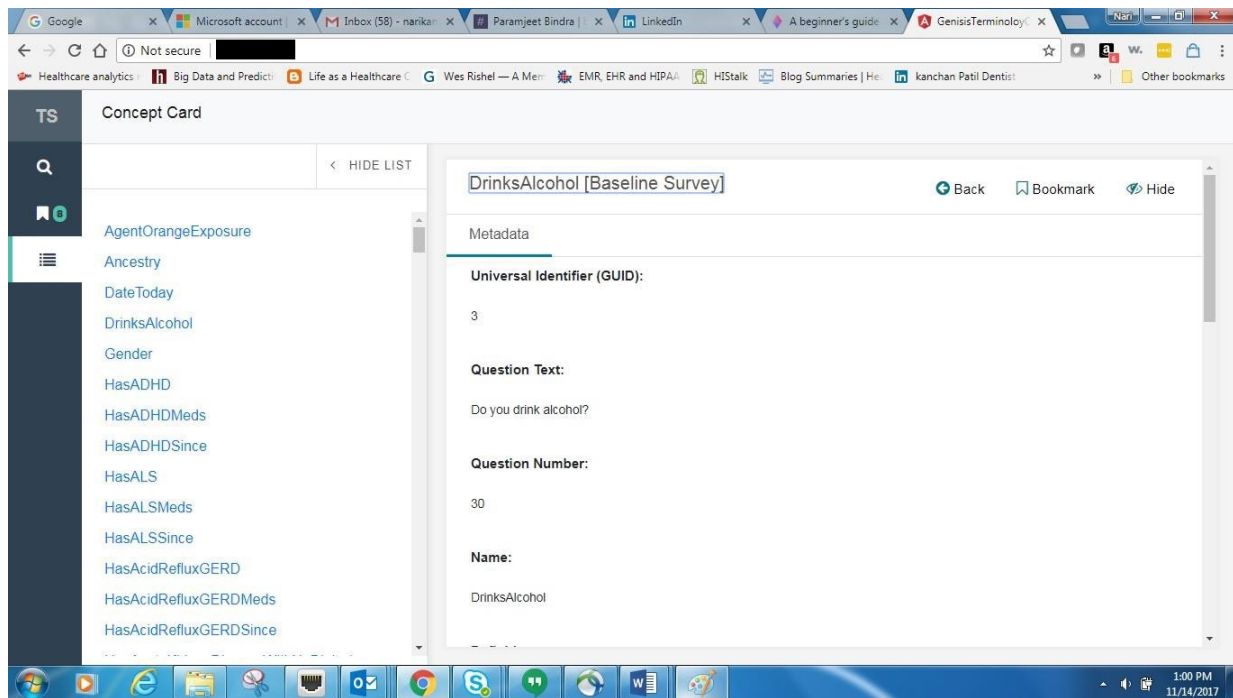


Figure 20: Browsing Data Elements

3.3.4.2.6 **Browsing Ontologies**

Users can navigate ontologies graphically using the Tree View (upper right pane) and click on Concepts in the Tree View to display their Concept Cards (lower right pane) as shown in Figure 20. To display the Tree View if it is hidden, click the "Tree View" link in the upper left-hand corner.

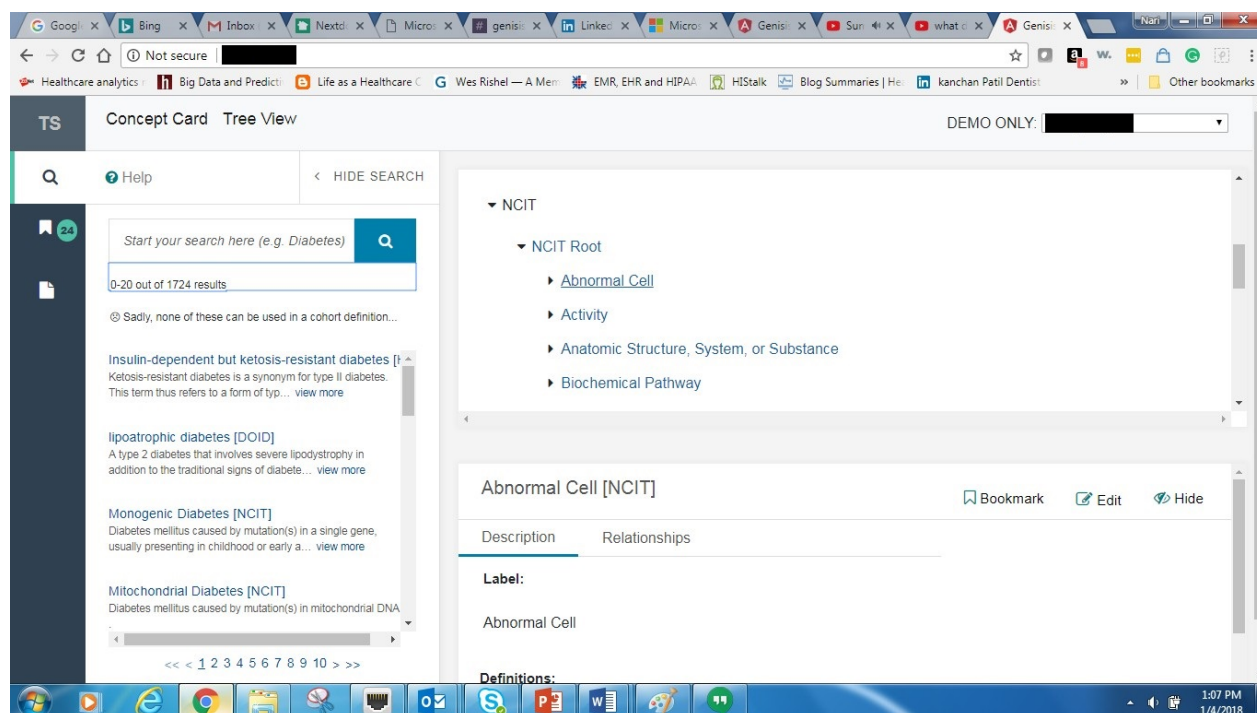


Figure 20: Browsing Ontologies

3.3.4.2.7 Suggesting, Batching, and Submitting Changes for Approval

Figure 21 shows how a Researcher or an Administrator can suggest changes to attributes of Concepts in Ontologies. These changes are like annotations on top of existing Ontologies/Concepts that make better sense for MVP researchers.

Users can click on the Edit/Editing button of a Concept Card to suggest changes. Then Edit mode is turned on. Clicking on the green plus signs opens up a dialog box for adding information. Clicking on the red “x” removes previously suggested information. All additions or removals are considered changes that need to be batched together with a name and submitted for approvals.

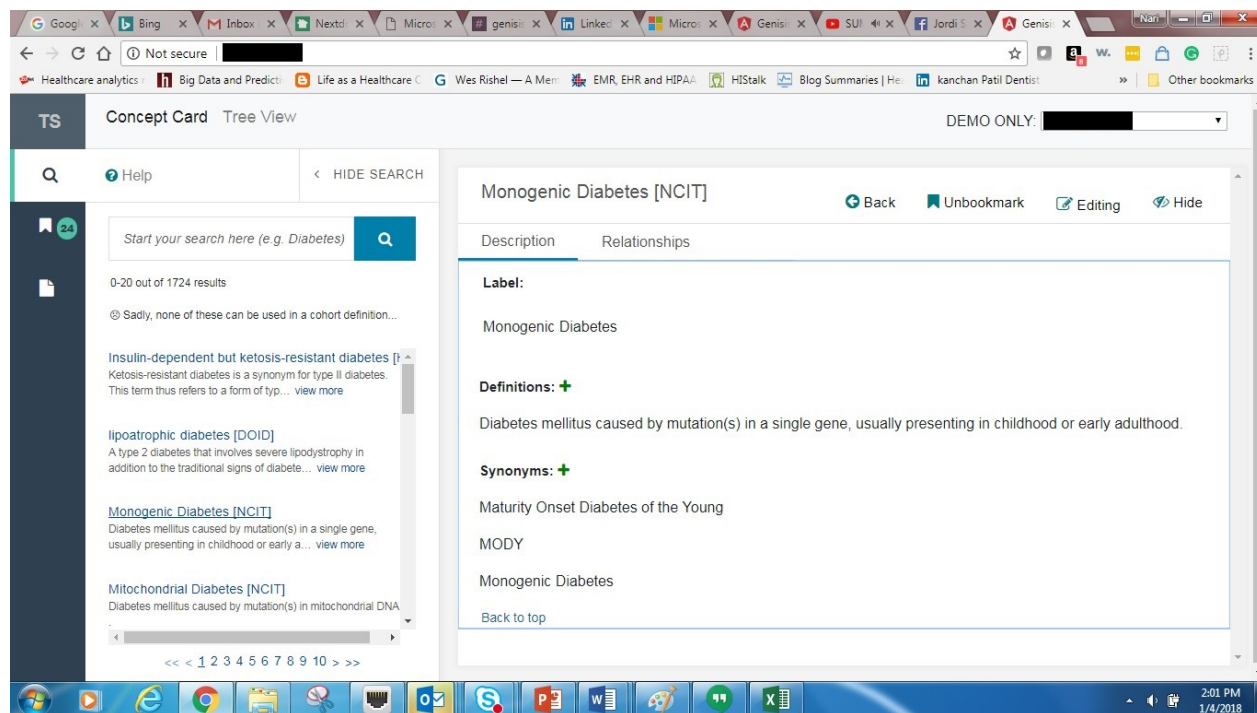


Figure 21: Suggesting Changes – Edit Mode

3.3.4.2.8 Creating Batches of Changes

Figure 22 shows how a change to a definition of Diabetes Mellitus can be submitted as a Batch of Changes. Batches of changes are given a name (Test batch 1 in this example) and submitted for approval.

Any batch of changes can be viewed by clicking on the batch name as shown in Figure 23. It also shows a single change in the Batch Description pane. The corresponding change (underlined) is in the Concept Card below. The **Submit for Approval** button at the top right can be used for submitting a batch of changes.

Users can select the Batch Change menu to view Batches of suggested changes (left pane). For a selected Batch, users can view detailed change information in the “Batch Details” pane and add a comment for each batch that will be viewable by an Administrator. The Concept Card for containing each change will be displayed in the lower right- hand pane, with each change indicated (an underline for new information, a strikeout for removed information).

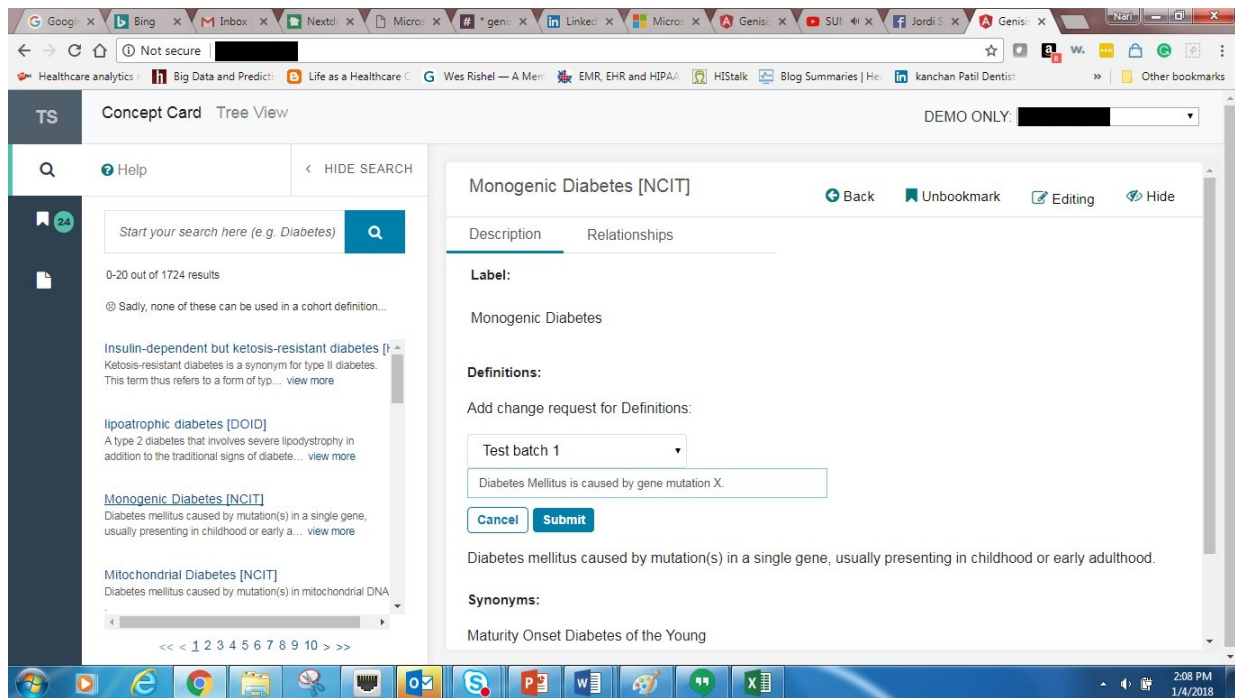


Figure 22: Suggesting Changes

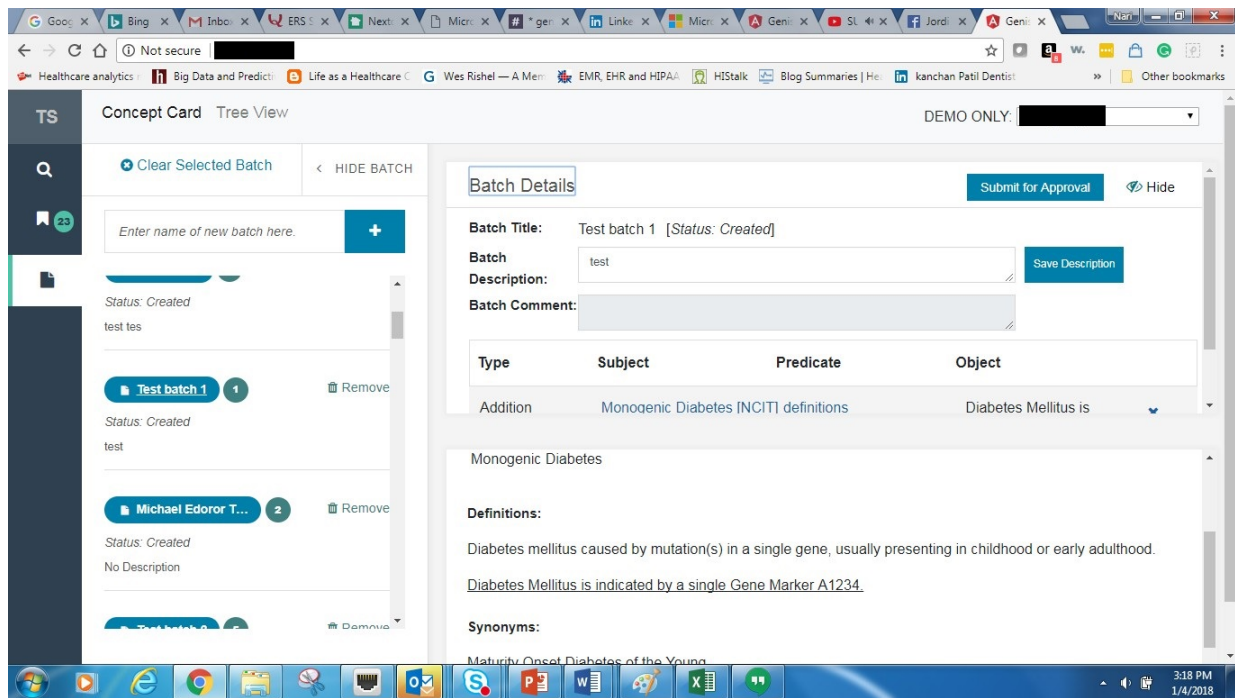


Figure 23: Viewing a Batch of Changes and Submitting it Batches of Changes

3.3.4.2.9 Approving or Rejecting a Batch of Changes

Approvers can view Batches of suggested changes to Approve or Reject in the Batch menu (upper right of the Batch Details pane) as shown in Figure 24.

If rejecting, an Approver must include a Batch Comment that will be sent back to the Researcher. Details of suggested changes can be viewed in the Batch Details table and the changes are displayed in the Concept Card they correspond to (lower right pane) as shown in Figure 25.

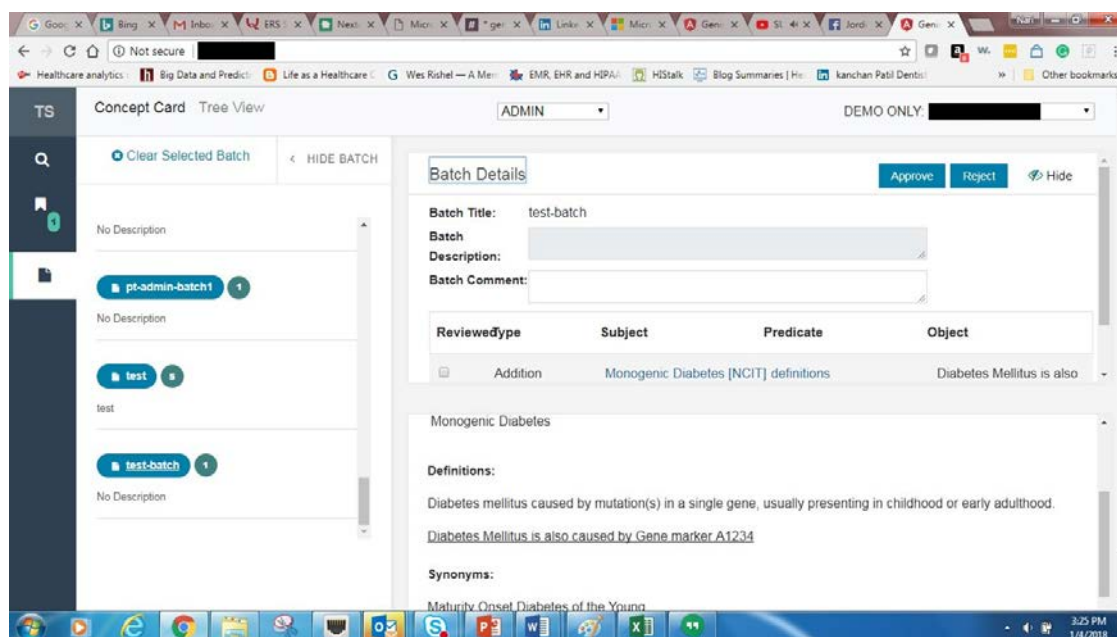
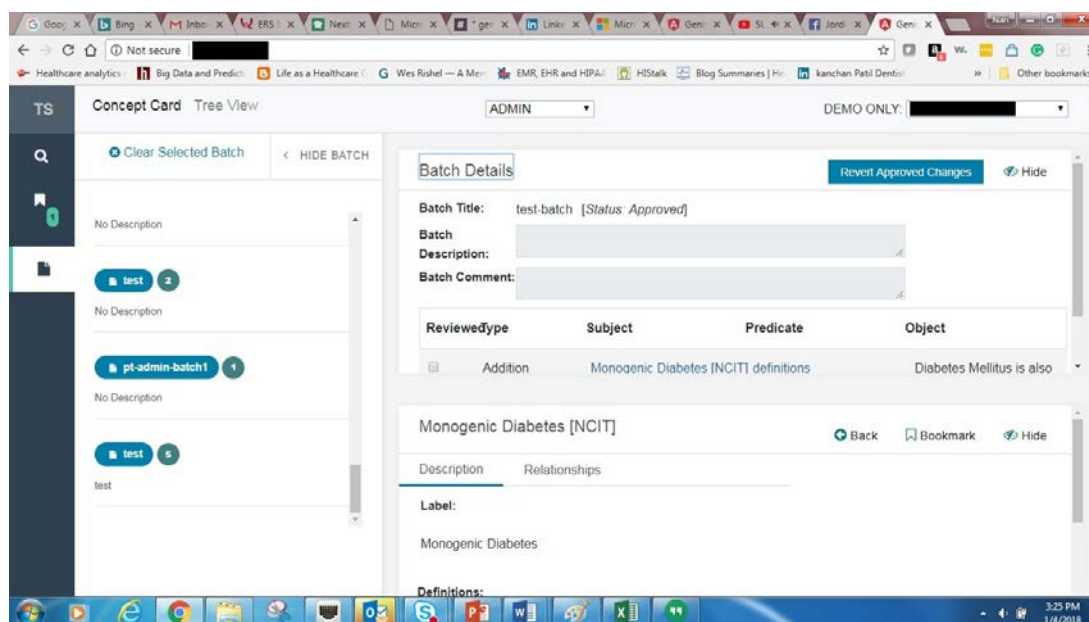


Figure 24: Approve or Reject a Batch of Changes



3.4. Conceptual Infrastructure Design

The conceptual infrastructure design is a high-level overview of the infrastructure used to support the Genesis2 system. *These activities are not in the scope of the Genesis2 project but provided here only for context. It is assumed that all of these are in place already or will be made available by VA on time for the Genesis2 development team.*

The current infrastructure supporting Genesis will be expanded as follows:

- Add Petabyte (PB) storage progressively through the years 2016 to 2018, with a projection of 45 PBs by June 2018.
- Increase the bandwidth by at least 2X (doubling) to add throughput between VINCI and Genesis, and to meet the “big data” traffic demands anticipated with the addition of new studies.
- Add computing capacity by interfacing Genesis to a public/private hybrid cloud with the capability to load-balance during “high burst” periods of intensive large-scale computing, and to provision cloud services for archival storage. Hybrid cloud deployments must be documented in the Federal Risk and Authorization Management Program (FedRAMP) and requires a FedRAMP Authority to Operate (ATO).
- An implementation strategy developed to demonstrate the feasibility of a hybrid cloud in a pilot instance to:
 - Test integration of the hybrid cloud in the pilot instance;
 - Optimize the hybrid cloud for performance and adaptive/perfective maintenance; and
 - Test hybrid cloud capabilities to support load-balancing during “high burst” periods of intensive large-scale computing, and provisioning for archival storage.

Figure 261 provides a high-level illustration of the requested expansion and enhancements for Genesis2.

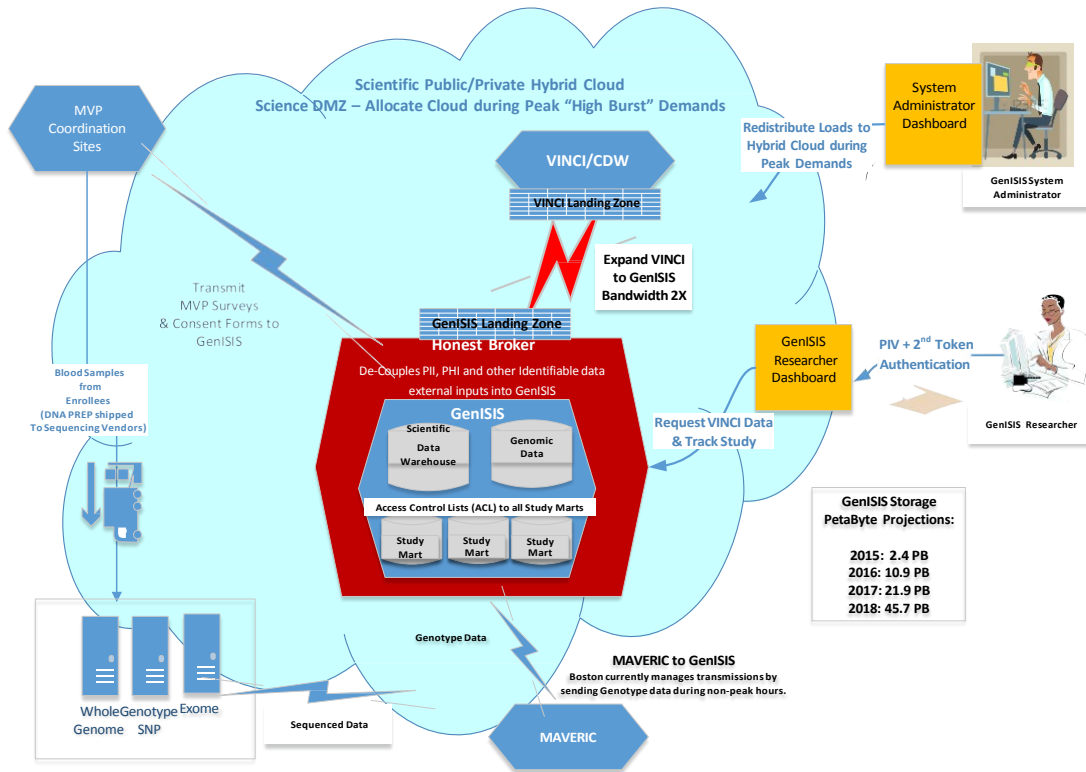


Figure 264: Genisis2 “To Be” With Expanded Bandwidth, Storage, and Hybrid Cloud

3.4.1. System Criticality and High Availability

The Genisis2 application is not mission critical. The Recruitment and Enrollment Platform will support national users during business hours from 8:00am – 8:30pm ET, with 98% uptime. However, the Scientific Platform (HPC cluster and analysis stations) will support Researchers with 100% uptime. The Scientific Platform is used across multiple geographic time zones, and must support concurrent, long-running computational jobs that may be scheduled at any time during a 24x7 period. The Genisis requirements for Availability, Allowable Downtime, Recovery Time Objectives (RTO) (for disaster recovery) and Recovery Point Objectives (RPO) (for disaster recovery) are identified in the Genisis2 RSD.

3.4.2. Special Technology

Special Technology has not been identified for this system. Table 10 is a placeholder for this information should it be identified in the future.

Table 10: Special Technology Requirements

Special Technology	Description	Notional Location	TRM Status
N/A	N/A	N/A	N/A

3.4.3. Technology Locations

This section describes the various technology components that will be used and their locations, as shown in Table 11. The host site is located at the Pittsburgh, PA Information Data Center with

the MAVERIC facilities at the Boston VA Healthcare System identified as the Disaster Recovery (DR), development, and test environment. *N/A below is Not Applicable in general or at least as of this Release. If the status changes, updates will be made as necessary in the future.*

Table 11: (Grouping) Technology Location Details

Technology Component Production 1	Location	Usage
Workstations	All VA research sites	Genisis access via web browsers
System Hardware	VAMC – Pittsburgh, PA	<ul style="list-style-type: none"> • Genisis host site web server • Application server • Database server • Data Mart server
Special Hardware	N/A	N/A
Interface Processors	N/A	N/A
Legacy Mainframe	N/A	N/A
Legacy Application Server	N/A	N/A
Legacy Databases	N/A	N/A
Other	N/A	N/A

Technology Component Production 2	Location	Usage
N/A	N/A	N/A

Technology Component Certification	Location	Usage
N/A	N/A	N/A

Technology Component Education	Location	Usage
N/A	N/A	N/A

3.4.4. Conceptual Infrastructure Diagram

The following diagrams depict a concept for the future Genisis infrastructure, which provides flexible, on-demand access to MVP cloud services. This cloud model is expected to enable implementation of ongoing, changing requirements to expand or contract capabilities as needed.

This includes HPCC and PB storage resources. Figure 272 illustrates how all groups access cloud resources, while Figure 283 focuses on the workflow of the HPC environment. *Note that these are future states envisioned and not part of the project as of this release.*

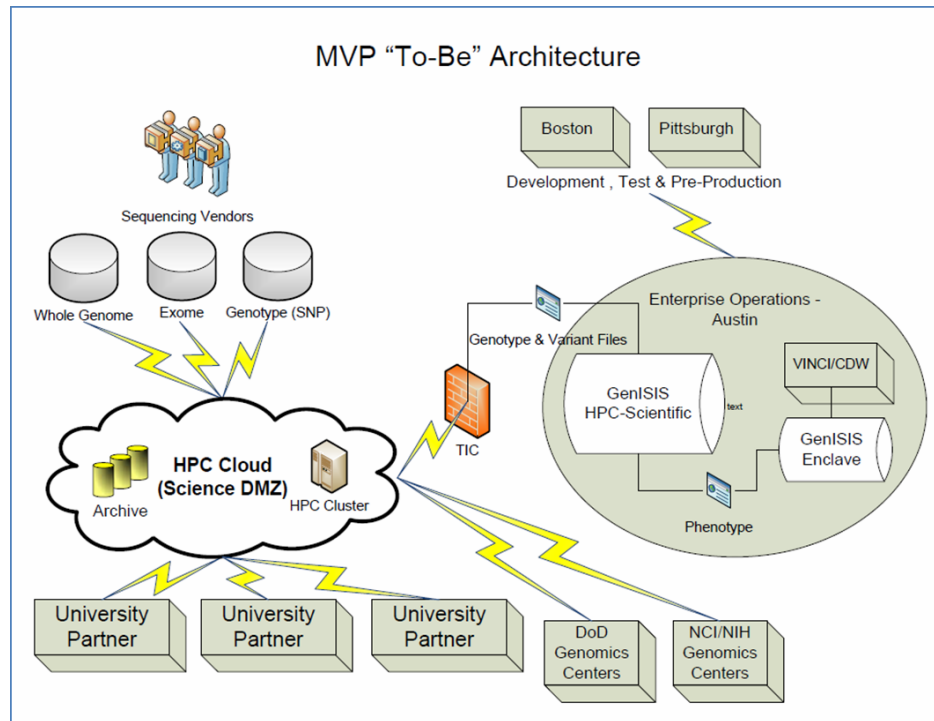


Figure 272: Concept of Cloud for VA Genomics and Precision Medicine

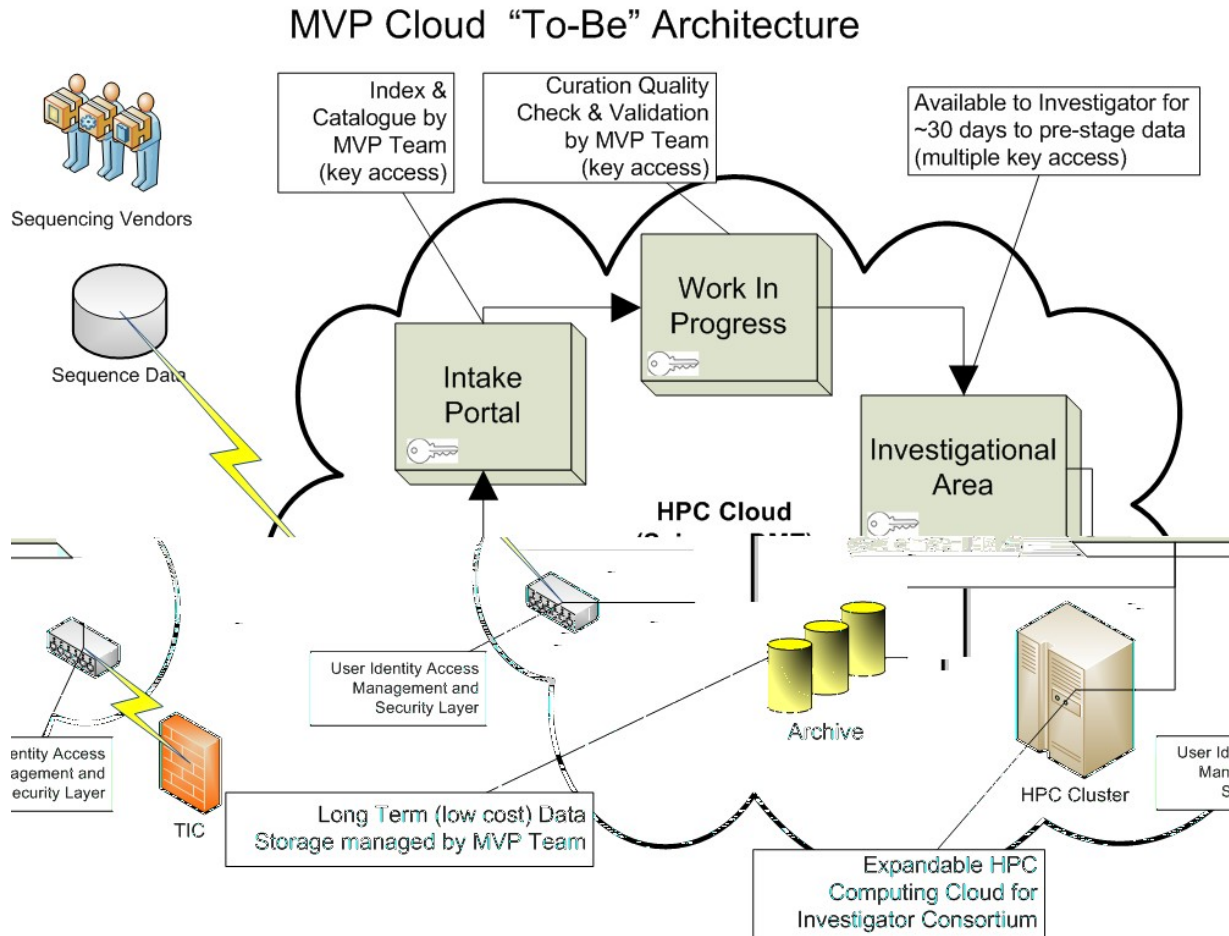


Figure 283: High Performance Computing Cloud as a Science DMZ

3.4.4.1. Location of Environments and External Interfaces

Figure 294 depicts the environments that will be supported, including the local networks to which they will be attached, and the locations at which they will be installed. Each of the external interfaces is shown in terms of where they enter the network.

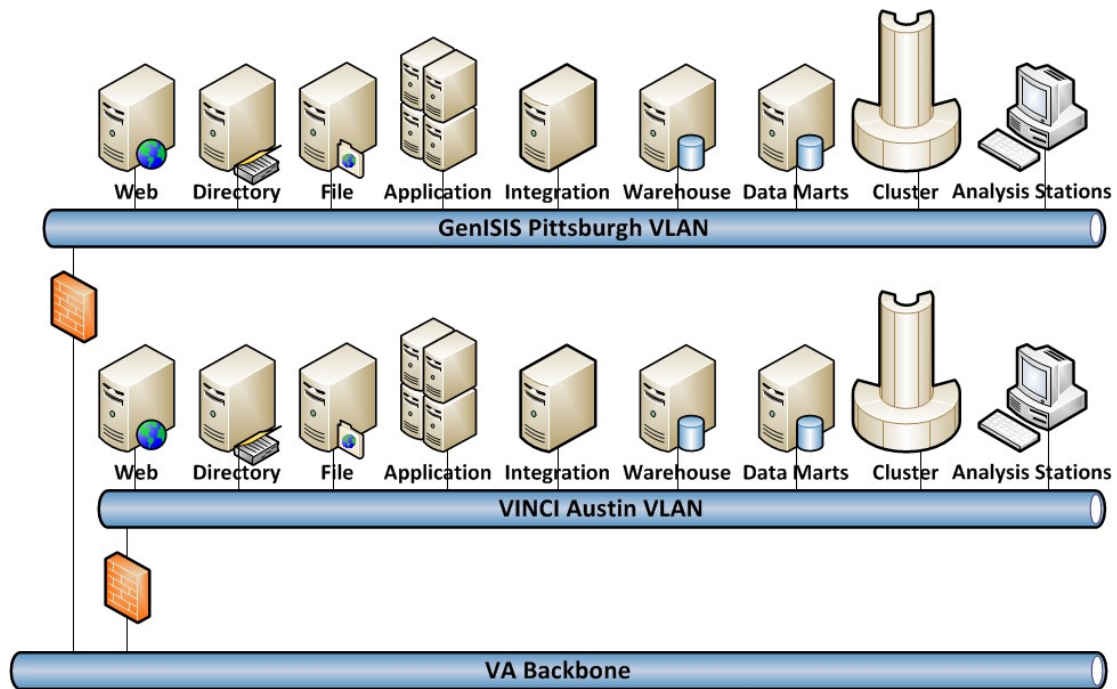


Figure 294: Conceptual Networks and Environments

3.4.4.2. Conceptual Production String Diagram

3.4.4.2.1. Genisis Workflows Production String Diagram

The Conceptual Production Environment for the Genisis Data Request Workflows consists of three tiers as shown in Figure 3025.

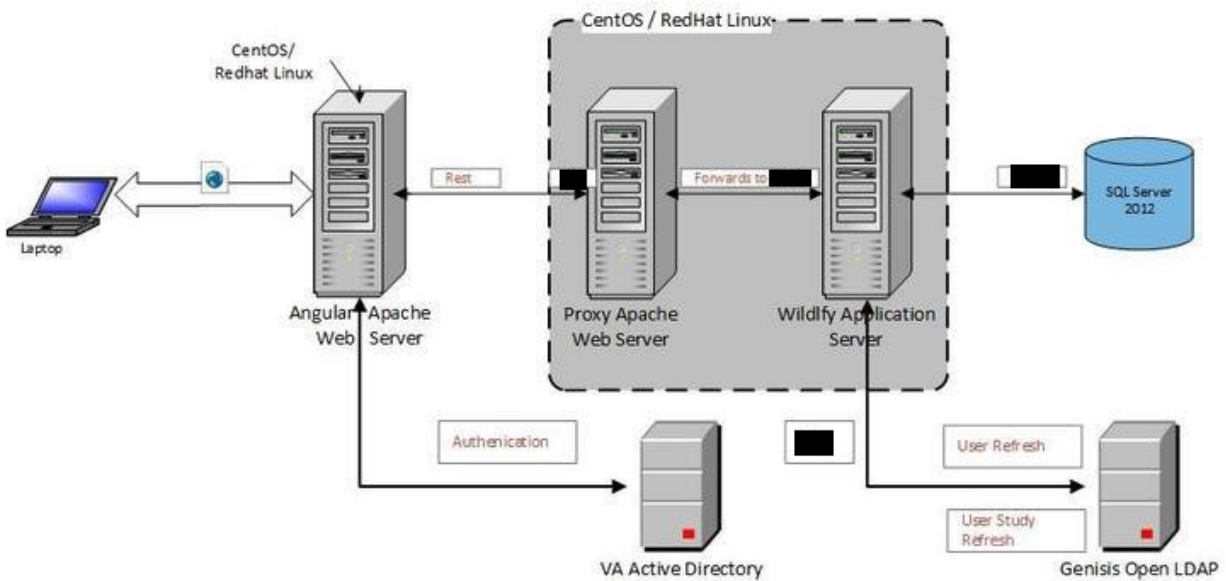


Figure 3025: High-Level System Design for Genisis Data Request Workflows

Genesis2 architecture for the Data Request Workflows consists of the following three tiers:

- **Presentation/Webserver Layer:** Apache Webserver on Linux – does 2-Factor Authentication using a PIV card and Windows Authentication with the VA LDAP system. This webserver handles browser-based access from users. Traffic from this is redirected to a second webserver in the next layer. The second webserver is for access (not used currently but in place for the future) from other systems within the VA network to access Genesis2 application through an API (Application Programming Interface).
- **Application/Business Logic/Workflow Layer:** Wildfly server on Linux– Application code in Java using the Activiti Workflow engine to coordinate the workflow.
- **Database Layer:** Microsoft SQL Server 2012 on a Microsoft Windows 2012 R2 server.

3.4.4.2.2. Terminology Services Production String Diagram

The Conceptual Production Environment for Terminology Services consists of three tiers as shown in Figure 3126.

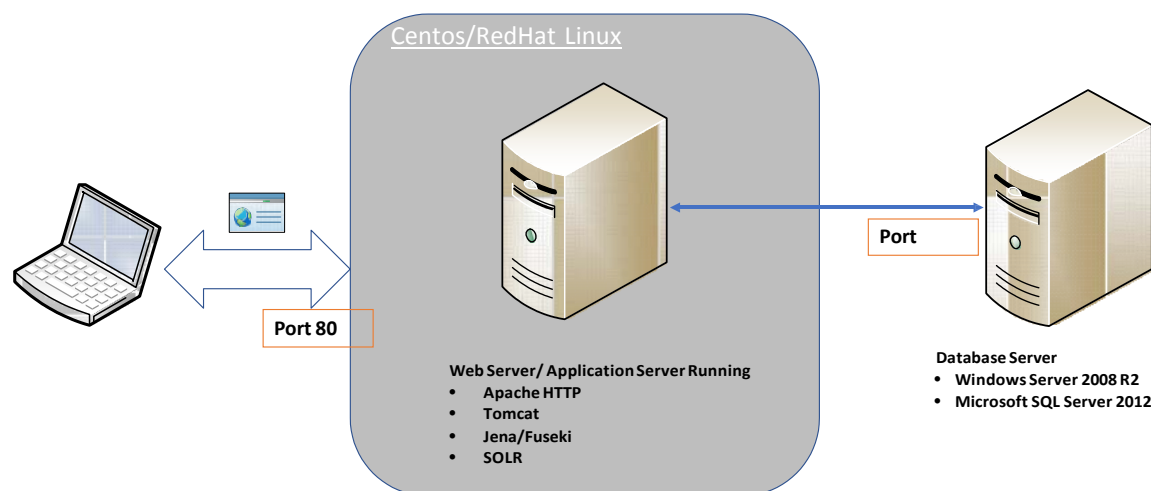


Figure 3126: High-Level System Design for Terminology Services

Genesis2 architecture for Terminology Services consists of the following three tiers:

- **Presentation/Webserver Layer:** Apache Webserver on Linux – This handles browser based access from users and sends them as REST calls to the Tomcat Application Server in the same Virtual server.
- **Application/Business Logic/Workflow Layer:** Tomcat server on Linux– This application server is resident in the same Virtual Machine as the Webserver. Application code in Java implements the logic portion of Terminology Services.
- **Database Layer:** Microsoft SQL Server 2012 on a Microsoft Windows 2012 R2 server.

4. System Architecture

The Genesis2 application is designed to serve as a research computation platform for the VA Genomic Medicine Program. In order to leverage the experience of the worldwide genomic research community, the Genesis architecture follows open source and COTS architecture that is currently in wide use.

Logically, Genesis2 is divided into the following major subsystems:

- Relational Database, which provides the Genomic Data Warehouse and Data Query applications;
- A large capacity file store for storage of large flat files from genomic analysis;
- Web-based applications, which provide governance management, and mediate the data query and the computational job submission applications;
- Virtual analysis workstations for ad-hoc data analysis and job preparation; and
- An HPCC with large-capacity, parallel file system for performing genomic analysis.

4.1. Hardware Architecture

Genesis2 hardware architecture is a centralized design that comprises server and storage components to meet the needs of large-scale genomic analysis. This section and Section 6.1 describe the server and storage components in more detail. Refer to Genesis2 RSD for requirements and projections to increase PB storage capacity from fiscal year (FY) 2016 through FY 2018.

A survey of major centers of genomic research showed that the predominant storage systems used were Isilon clustered file systems connected by GigE networks. The streaming performance characteristics of the Isilon system were used as the basic requirements of the system.

Genesis storage design consists of three classes of storage:

Local Disk: Each computational node is deployed with 300GB of Statistical Analysis Software (SAS) storage, which may be upgraded during a future refresh.

SAN Storage: A 30TB EMC CX4 Storage array is connected to a dedicated 10GigE iSCSI network to provide virtualized storage for the HPCC head node and the virtual analysis workstations.

Clustered File Store: 60TB Isilon and 2.7PB NetApps storage appliance provides a high-bandwidth analysis workspace for genomic analysis. Storage is presented via NFS over the non-blocking 10GigE HPCC data network.

The HPCC consists of commodity Linux servers that are chosen to maximize the performance-to-price ratio. The first iteration of Genesis consists of 114 identical Dell M610 Blade servers (HPCC Compute nodes). Each server has dual hex-core Intel Xeon processors, with 96GB memory, packaged in a dense, half-height blade. Computational servers chosen for future expansion will be based on best performance-to-price ratio available. Figure ~~3227~~ depicts the Genesis hardware architecture.

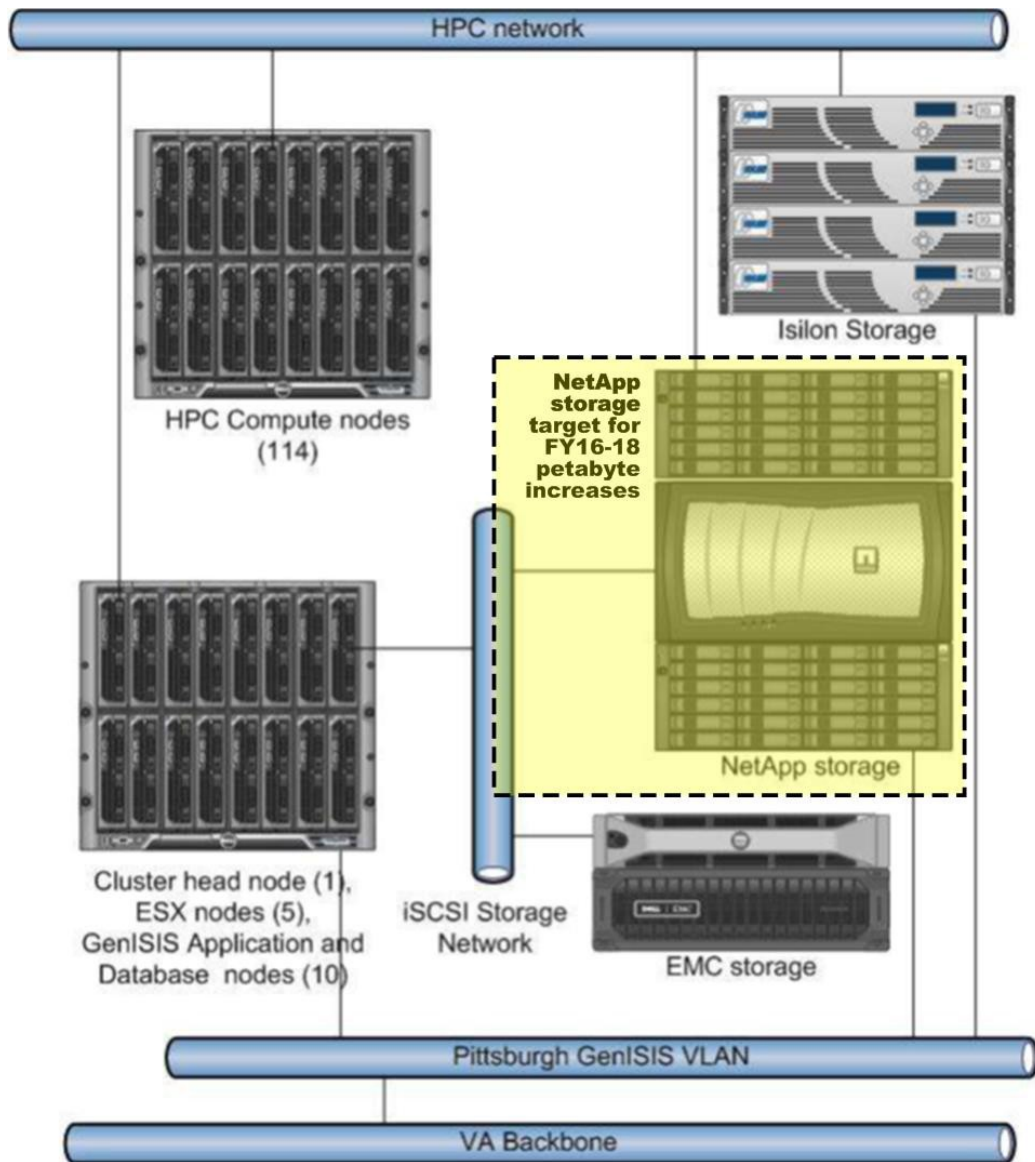


Figure 3227: Genisis Hardware Architecture

4.2. Software Architecture

The Genisis2 front-end is accessed by a VA standard web browser and uses the existing authentication that is obtained from logging into a VA domain with valid user credentials. The VA has multiple domains, each with its own users. The goal for Genisis2 is to implement the new requirement of VA PIV access control for all VA domain users.

All Genisis2 interactions use Hypertext Transfer Protocol Secure (HTTPS), which requires signed server certificates from the VA's security authority. The Apache Web Server will handle all authentications. User VA credentials used to authenticate will be passed to VA LDAP, along with the request; the container called depends on the URL of service call. For security purposes, Apache will only listen on localhost and are limited to Apache JServ Protocol (AJP) connections. Other ports will be disabled, as needed, if not required by the application server.

After the user is successfully authenticated, the user's authorizations must be gathered and evaluated. Since the Active Directory does not store information associated with Genisis, the Genisis database itself is used to provide group membership. Figure [3328](#) depicts the software architecture and context of Genisis2 Data Request workflows.



Step 1 – Researcher Types in Term of Interest – **Terminology Services** Coordinates everything through other services

Step 2 – Search Service takes Term, returns Candidate Concepts

Step 3 – Researcher Interactively Browses Concepts Returned using the **Concept Service**

Step 4 – Researcher Bookmark **Terminology Services** Concepts of Interest and Adds them to Named Folders using the **Bookmark Service**

Step 5 – Aggregate Query Tool Fetches Periodically, Named Folders with Bookmarks

Step 6 – Researchers/Principal Investigator Browses Bookmarks available for them to use in the AQT

Step 7 – Researchers/Principal Select Concepts, Formulates a Query and Executes it

Step 8 – Aggregate Query Tool Traces the path **Bookmark Service** – **Concept Service** – **Mapping Service** – **Data Elements** and Executes the Query

Step 9 – Aggregate Query Results are returned to the Researcher

The architecture includes the following components and data flows:

- Terminology Tool** (containing **Terminology Browser** and **Terminology Editor**) interacts with the **Terminology Service**.
- Terminology Service** interacts with **SPARQL**, **SOLR**, **CQL**, and **Agg Query Tool**.
- SPARQL** interacts with **Concept Service** (e.g., Jena-TDB+Fuseki) and **Search Service**.
- SOLR** interacts with **Search Service**.
- CQL** interacts with **Bookmark Service** and **Mapping Service**.
- Agg Query Tool** interacts with **Bookmark Service** and **Mapping Service**.
- Concept Service** interacts with **Concept Database**.
- Search Service** interacts with **Lucene Database**.
- Bookmark Service** interacts with **Concept Basic Database**.
- Mapping Service** interacts with **Term Metadata Database**.

Figure 3429: Software Architecture and Context of Genesis2 Terminology Services

4.2.1. Software Development Approach Analysis

4.2.1.1. Genesis2 Data Request Workflows

The following approach was taken to analyze the options available for software development based on the architectural principles outlined in Section 3.1. Our team:

1. Identified **Custom Software Development** as one of the options available to us for the development of Genesis2 application.
2. Identified a number of candidate **Commercial Off-the-Shelf (COTS)** and **Open Source** software products that could be possible candidates for *alternatives for Custom development*.
3. Identified key features that the software products should have for consideration as Genesis2 implementation ready-made features, or features that can be modified readily for use.
4. Identified the status of the above software products according to [the One-VA Technical Reference Model v17.2, Technology Standard List](#). Chosen for consideration were products **Approved** or **Approved with Constraints** AND for **CY2017** and **CY2018**.
5. In the case of Custom Software Development, the assumption was made that key features considered in the products for comparison could be developed within the software application itself.

In addition to the Genesis2 application, we also considered COTS or Open Source products that can be used in the Data Operations part of the project (Copying files between one Landing Zone to another and between a Landing Zone and a Study Mart or a Query Mart). The following analyses present the development team's findings for these two parts of Genesis2.

Analysis for the Genesis2 Application

The analysis of alternatives results for the Genesis2 Application are illustrated in Figure 350.

Option	TRM Approval or Approval with Constraints	Version Approved for CY2017	Cost Model	General Notes	Cost	Ability to Customize	Real-time Alerts and Notifications	Community Adoption	Maturity	General Integration Capabilities	Workflow/Process Flow	Active Directory Integration
Custom Development	NA	NA	NA	Necessary Features Can be Built	3	4	4	NA	NA	4	3	3
Microsoft SharePoint Portal	3	2013 and Above	COTS	Intranet Focused, Content Management based Workflows	3	3	3	3	3	2	2	3
LifeRay Portal	3	7.0 and Above	Only COTS version is Viable	Intranet Focused, Content Management based Workflows	2	2	2	1	1	2	2	2
Alfresco One/Activiti Workflow	3	201605 GA Version	Only COTS version is Viable	Intranet Focused, Content Management based Workflows	2	2	2	2	2	2	2	2
Nuxeo	0											
Plone	0											
Open Atrium/Drupal	0											
Portofino	0											
MindTouch	0											
	Prohibited for use											

Figure 350: Summary of Analysis of COTS Alternatives for the Genesis2 Application

Salient highlights of the analysis results:

- We did not analyze further if a COTS or an Open Source product did not have TRM approval (e.g., Nuxeo, Plone, etc.) or if it was prohibited for use within VA by OIT (Trac). They merit an empty circle.
- All of the COTS and Open Source products that were approved had workflow designs that were tied to document workflows or content workflows (such as approval of web pages across multiple people, for example). That is the reason they merit only half a circle for workflows in the chart above.
- Products like LifeRay and Alfresco had community editions that had limited functionality and only COTS versions of them had the features needed for Genisis2. This is the reason they both merit the half circle for Cost.

Analysis for the Data Operation Tools

The results of the analysis of COTS alternatives for the Genisis2 Application are illustrated in Figure 364.

Option	TRM Approval or Approval with Constraints	Version Approved for CY2017	Cost Model	General Notes	Real-time Alerts and Notifications	Community Adoption	Maturity	General Integration Capabilities	Active Directory Integration
Titan SFTP Enterprise	3	16.X and above	COTS - \$1950 for Enterprise Edition	Most widely used COTS product for automating Secure File Transfer Uses	3	4	4	3	3
WinSCP	3	5.7.7 and above	Open Source	Custom Code Built Around the Product	2	4	3	2	1

Figure 364: Analysis of COTS Alternatives: Genisis2 Research Data Handling Tools

Salient highlights of the results:

- The development team found many alternatives for transferring files securely.
- The choices were narrowed down to those products that can recognize files when they appear in folders, capable of real-time notifications when transfers are completed and supported SFTP (Secure File Transfer Protocol).
- The choices were further narrowed down to the most widely used products represented by the Community Adoption full circles. WinSCP has a large following in the users of Open Source products but on lighter set of simpler tasks. The Titan SFTP Enterprise is an enterprise-ready tool used widely in large organizations and more mature (V16 vs V5.7 for WinSCP).
- The Titan SFTP Enterprise product has many enterprise-class features and is very widely used in organizations with great reviews, especially for use with Windows platforms.

4.2.1.2. Terminology Services

Component selection for Terminology Services was a much simpler process because of the limited availability of qualified, open-source components. In addition, a number of components

were common with the Genisis2 Data Request workflow system. The final components chosen were:

- **Apache HTTP** – For handling browser requests through this webserver component. The analysis performed for Genisis2 Data Request Workflows was useful for this selection since it was identified as one of the better components.
- **Jena/Fuseki** – This is the Triple Store database that was chosen since it is one of the few Triple Store database software that is open-source, TRM approved, and already in use within the VA. A formal performance evaluation was made to make sure this has the necessary performance characteristics necessary for use.
- **SOLR** – This is the text search engine component chosen. It is also open-source, TRM approved, and in widespread use within the VA.
- **Microsoft SQL Server 2012** – This component was chosen for the Database Server since the analysis for Genisis2 Data Request Workflows already identified it as one of the best options.

4.2.2. Software Development Approach Chosen

The Development team recommends the following approaches based on the analysis of alternatives above and their results:

Genisis2 Application

The development team recommends the Custom Development approach for the Genisis2 application. The reasons are as follows:

- All of the other alternatives, even though they have many needed features at first blush are oriented towards simple Intranet, document-based or content-based workflow applications. Genisis2 automates the data request, preparation and fulfillment process and is, as such unique. Trying to force-fit this workflow on to a tool designed for content management applications may be a mismatch.
- Genisis2 workflows may evolve over a period with many subtle variations between one and the other. Custom development may be the best way to keep these features up-to-date and without bending to the quirks of the product chosen.
- Custom code can be tuned for high performance while if they are implemented with COTS or Open Source products may come with unnecessary performance penalties with respect to features within the product that the Genisis2 application may never use.

Product for Data Operations

The development team recommends the Titan SFTP Enterprise product for Data Operations. The reasons are as follows:

- The Titan SFTP Enterprise product includes features such as recognizing when a file appears in a directory and automating actions following that with scripts that can be invoked.
- This product contains real-time notifications needed by email or mobile SMS/Texts.

- It seems to be used widely in organizations, stable and long standing (version 16 is the one approved with constraints by VA OIT TRM).

Even though this is a COTS product with a nominal cost, the development team thinks that this will help move Genesis2 Data Operations to a higher level of automation, especially when it comes to moving data between landing zones or between landing zones and marts (study or query).

4.3. Network Architecture

Communications between system entities are divided into the following categories:

- Web browsers to web application
- Web application to web service
- Web application/service to database
- HTTP server (Apache) to Application Server (Apache)
- Web application/service to LDAP server
- HTTP server (Apache) to Active Directory Domain Controller

All communications will be encrypted using Transport Layer Security (TLS) version 1.2 or higher. A signed server certificate will be assigned to the system hosting the Apache HTTP server. Any web browser connecting to the Apache HTTP server will be redirected to a TLS connection, if not already using TLS. Hypertext Transfer Protocol (HTTP) 1.1. will be the data exchange protocol. User identities will be authenticated and user authorizations will be enforced.

Connections between web applications and web services will use Representational State Transfer (ReST) mechanisms over HTTP. Data exchanges will be typically Extensible Markup Language (XML) documents. Connections to external web services will be authenticated and authorized.

Web application/service connections to database will use TCP/IP connections. All connections will be authenticated and authorized to protect the contents. TCP/IP is used to support remote database connections.

Connections to LDAP servers will use LDAPs to ensure secure exchanges for user information. Connections will be authenticated and authorized.

Single-sign-on will be supported from the HTTP server using Kerberos authentication with an Active Directory Domain Controller.

4.4. Service Oriented Architecture / ESS

Genesis2 Presentation Layer talks to the Application Layer only through an Application Programming Interface (API), which is documented and presented along with the delivery of the software. This SOA approach can be useful for any other external software system to call Genesis2 the same way its own User Interface calls it.

4.5. Enterprise Architecture

Genesis2 will continue to retain its status as a “Medical Device” program, with the intent to remain exempt from the EA Technology Reference Model (TRM) requirements. Because of the high-computational/scientific computing aspects related to investigational genomics, Genesis is “isolated” and behind a Virtual Local Area Network (VLAN)/ Access Control List (ACL), and does not sit on the VA Backbone; therefore, Genesis2 shall meet the criteria required of a Medical Device Isolation Architecture.

The following is a list of the COTS software planned for the Genesis2 enterprise development environment, with TRM approval status and versions. The development team may update or alter as necessary, but must also abide by Architecture Engineering Review Board (AERB) constraint: “The project will ensure that all the tools in use are compliant with the TRM. If a tool is not compliant, the project will seek a waiver for its use.” Table 12 provides the technology location details.

Table 12: (Grouping) Technology Location Details

#	Product Name	Version	TRM Approved (Y/N)	Description/Use/TRM link
1	Apache HTTP Proxy Server	2.3 + ⁺	Y	Web Server, Authentication https://URL/Services/TRM/ToolPage.aspx?ti_d=5009
2	WildFly	10.0.0-Final	Y	Web Application Container, Genesis Web Apps https://URL/Services/TRM/ToolPage.aspx?ti_d=240
3	Apache Tomcat	9.0	Y	Web Application Container, Genesis Web Apps https:// URL /Services/TRM/ToolPage.aspx?ti_d=5451
4	Apache Jena/Fuseki	3.4.0	Y	Triple Store Database and Query Engine https:// URL /Services/TRM/ToolPage.aspx?ti_d=9476
5	Apache SOLR	6.0.0	Y	Text Search Engine https:// URL /Services/TRM/ToolPage.aspx?ti_d=7005
6	Microsoft SQL Server	2012	Y	Relational DBMS for Genesis Not used for external connections. https:// URL /Services/TRM/ToolPage.aspx?ti_d=5020
7	Red Hat Enterprise Linux	7.x	Y	Operating System https:// URL /Services/TRM/ToolPage.aspx?ti_d=6367

#	Product Name	Version	TRM Approved (Y/N)	Description/Use/TRM link
8	Windows Server	2012	Y	Operating System https://www.oit.va.gov/Services/TRM/ToolPage.aspx?tid=35
9	JDK – Sun Java SE Development Kit	1.8.0_92 (8u92)*	Y	Java Runtime Environment and Utilities https:// URL /Services/TRM/ToolPage.aspx?tid=5161

+The Apache HTTP Server Version 2.3 is in a status of “Divest in the 4th Quarter of Calendar Year 2018” currently according to the VA TRM. We will deliver Builds 9 and 10, re-evaluate all of our software versions, and upgrade to the appropriate versions of all software during a Sustainment release before contract end on June 7, 2018.

*The Java SE Development kit 1.8.0_92 (8u92) is in a status of “Divest through 4th Quarter of Calendar Year 2018” currently according to the VA TRM. We will deliver Builds 9 and 10, re-evaluate all of our software versions, and upgrade to the appropriate versions of all software during a Sustainment release before contract end on June 7, 2018.

5. Data Design

Figure 372 represents the request data status over its lifecycle in Genisis2 Data Request Workflows.

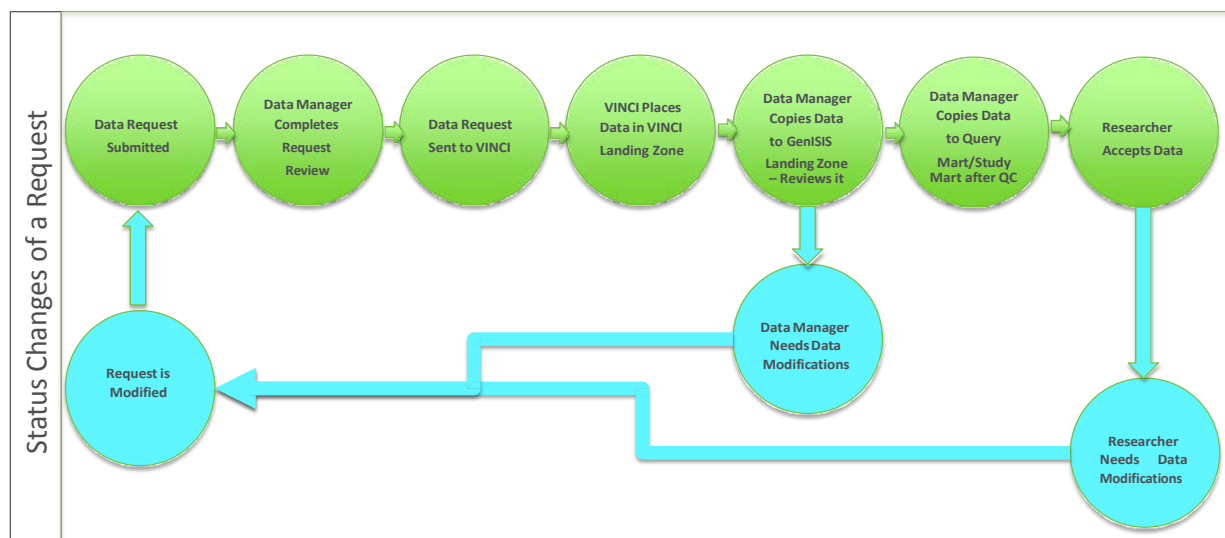


Figure 372: Request Status Lifecycle

Figure 383 shows in more detail the various State Transitions and Status changes.

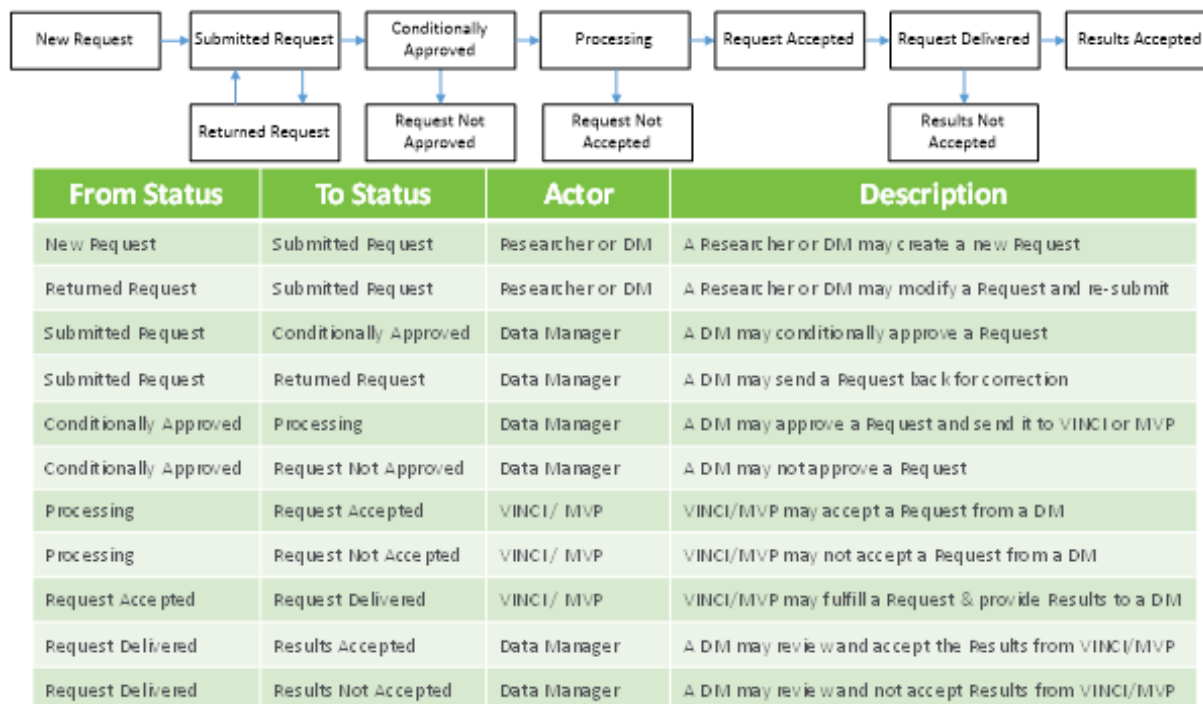


Figure 383: Status Transitions

5.1. Human Machine Interface

5.1.1. Genesis Data Request Workflows

The following provides details of users and what they accomplish using screens in the Human Machine Interface:

- Researchers** – Researchers are typically Researchers who are requesting Genomic Data, Clinical Data and Survey Data collected from Veteran volunteers in the MVP program. They log in using two-factor authentication, type in their request, attach documents explaining what they need in more detail, and submit it to the Genesis Data Manager. Once they get notification that the data is ready and copied to their study mart, they can start their research efforts with the data provided.
- Genesis Data Managers** – Genesis Data Managers are data coordinators that referee requests for data between sources of that data and the Researchers. Some data may not be available and some data may have some clarifications that get passed back and forth before the data can be made available. There may be some data cleaning efforts that is done on the data that is obtained from the source. Genesis2 does not handle those activities but Genesis Data Managers may add a comment or two regarding those activities.
- VINCI Data Managers** – These managers are sources within the VA for Clinical Data and Survey Data. Genesis Data Managers coordinate with VINCI Data Managers in reconciling the data that's available and the data requested by researchers. In some cases,

the same kind of data may be available in more than one location within the Clinical Data Warehouse (CDW), for example. Height and Weight may be recorded in VA Primary Clinics, In Patient Hospitals or VA Specialist divisions, being stored in different parts of the CDW. The Data Source and Destination Managers make sure the right information is extracted and provided to researchers (Researchers in the Genisis2 sense).

- **Genisis System Administrators** – These are super users that can perform all of the functions that a Researcher or a Genisis Data Manager can do. In addition, they can update from the Genisis2 Open LDAP server, the latest set of users. Permissions such as Researcher, Genisis Data Manager, VINCI Data Manager or another Genisis System Administrator are assigned in the Genisis2 OpenLDAP server (central repository for permissions for many other systems including Genisis2).

5.1.2. Terminology Services

The following provides details of users and what they accomplish using screens in the Human Machine Interface:

- **Researchers/~~Terminologists~~**—Genisis2 Terminology Services consists of the backend loading a number of standard Ontologies like SNOMED-CT, NCIT, and DOID. These Ontologies make available to the user a number Clinical and Genomic Concepts for them to start their search for Data Elements. For example, if their research deals with “Diabetes”, they start searching for “Diabetes” and navigate their way to data elements that may be attached to Concepts that they are interested in. The user can then bookmark these Concepts. They can also label them for future grouping and search. Users then use those bookmarked Concepts attached to Data Elements for requesting MVP Clinical, Survey and Genomic data.
- **Administrators** - Approves changes suggested by Researchers or Terminologists (equivalent to Genisis Data Managers above) to Ontology Attributes. They also perform backend functions such as loading Ontologies, adjusting the look and feel of individual ontologies through a Command Line Interface (CLI).

5.1.3. Interface Design Rules

Below are some of the conventions and standards for designing the GUI:

- It should follow the general look and feel of the VA Office of Research and Development (ORD) websites and tools.
- It should follow the general Logo and other Identity related colors that the Veteran’s Administration uses in all of its tools and websites.
- It should provide the standard set of links at the footer of each screen that most of the VA sites and tools have in their footers.
- It should be fairly easy to navigate and use.
- It should be Section 508 compliant.

5.1.4. Inputs

Below are some of the characteristics of the Inputs to Genesis2:

- All inputs to Genesis2 are planned to be by way of the keyboard and mouse clicks only. No forms are scanned in. Documents such as PDF, Microsoft Word, PowerPoint, Excel spreadsheets (all of Microsoft Office formats) may be attached to what they type in on the Genesis2 screens.
- No forms packages are used in the GUI.
- Messages – Positive or Negative are displayed at the Top Left corner of all screens in the Genesis2 application when inputs are wrong.
- Access restrictions are implemented by Roles – Researcher, Genesis Data Manager, VINCI Data Manager or Genesis System Administrator. Different tabs are displayed or taken away depending on who logs in.
- Security is enforced at login time using the VA PIV Card based Two-Factor Authentication. This and the VA email address is used to authenticate users. Authorization is controlled by the Roles that a user has in the system.

5.1.5. Outputs

Below are some of the characteristics of the outputs from the Genesis2 application:

- Genesis2 could transfer data automatically behind the scenes because of actions on the Genesis2 screens.
- Messages – Positive or Negative are displayed at the Top Left corner of all screens in the Genesis2 application when operations are successful or not successful.
- Outputs could also be reports generated as Txt or CSV files in the future when reporting functions are added in a future release. Reports envisioned currently are ad-hoc, on demand generated but could also be scheduled periodically in the future.

5.1.6. Navigation Hierarchy

The navigation hierarchy ~~will be elaborated on as the application becomes more fully developed.~~ is outlined in detail in Sections 3.3.4, 5.1.7 and 5.1.8.

5.1.7. Login Screens

When the user types in the appropriate URL for Genesis2, they reach a screen that looks like this. They need to insert their PIV Card into the Card slot; appropriate certificates all pop up as shown in Figure 394. The user selects the VA PIV certificate.

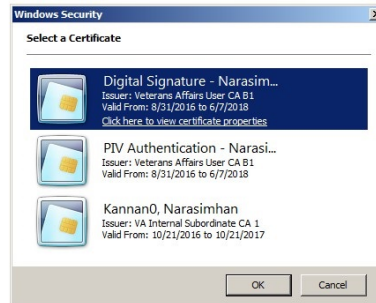


Figure 394: Initial Login

This prompts for the PIV Pin and then logs the user in as shown in Figure 4035.

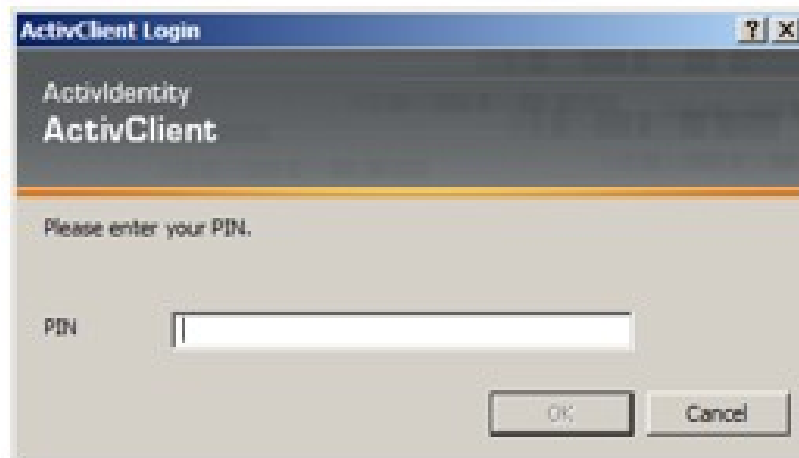


Figure 4035: PIN Prompt

Landing Zone and Navigation Aspects – Genesis Data Request Workflows

Once the user logs in, they are placed in a Landing zone and the various aspects of navigation are as shown in Figure 4136. Annotations are also provided.

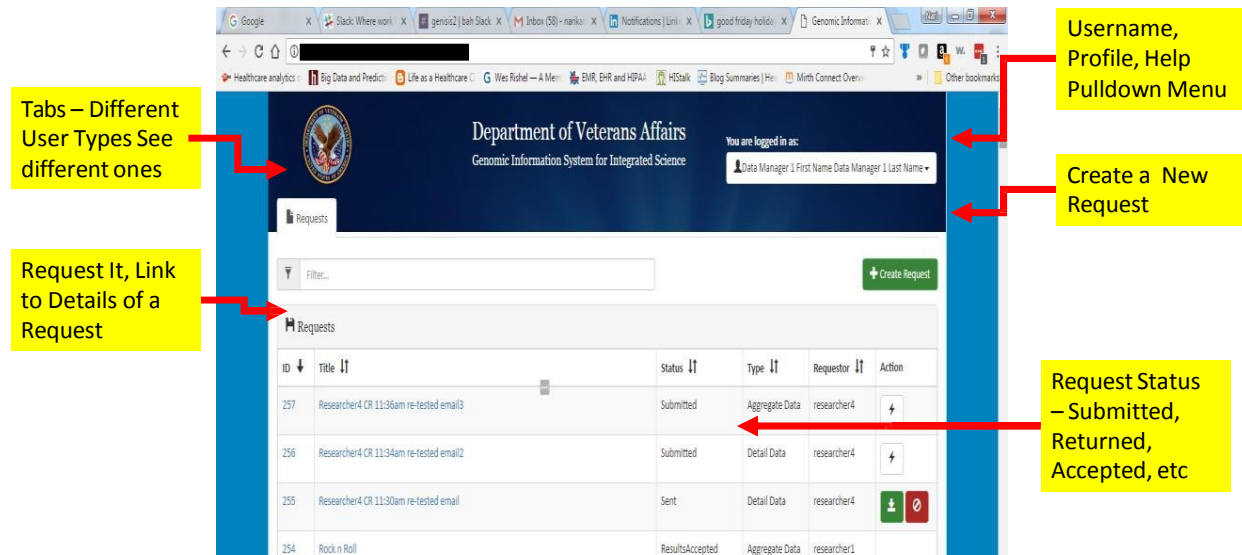


Figure 4136: Landing Zone and Navigational Aspects

Landing Zone and Navigation Aspects – Terminology Services

When the user accesses Terminology Services, they reach a landing page as shown in Figure 4237. This is where the user starts their search for concepts by typing in terms in the Search Box.

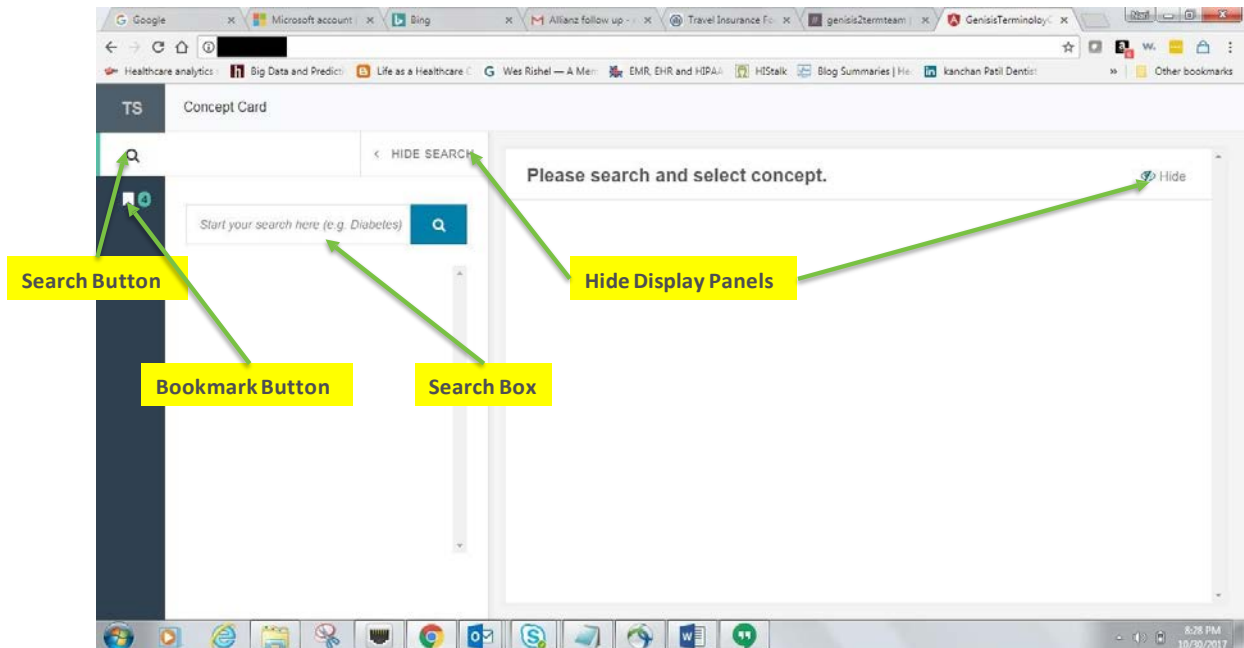


Figure 4237: Terminology Services Landing Zone and Navigational Aspects

Figure 4338 shows the Navigational Aspects when the user has found a Concept of interest and is interested in Bookmarking that Concept

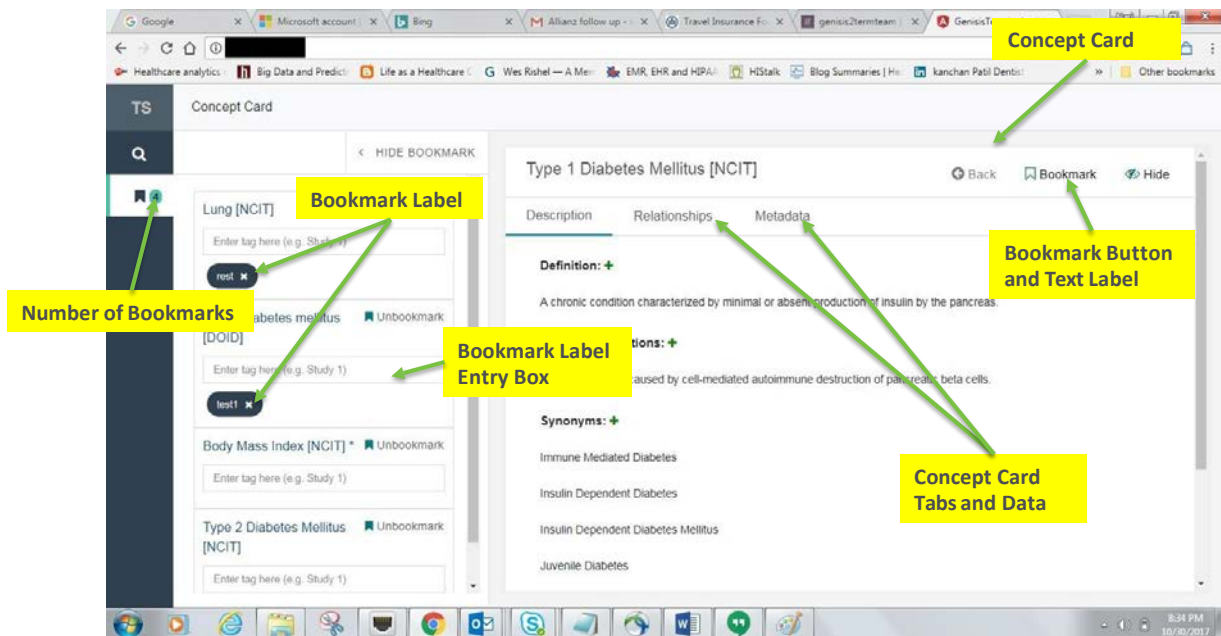


Figure 4338: Terminology Services – Bookmarking Concept

5.1.8. Details of Any One Data Request

Figure 4439 is a screen shot of the details of any one request.

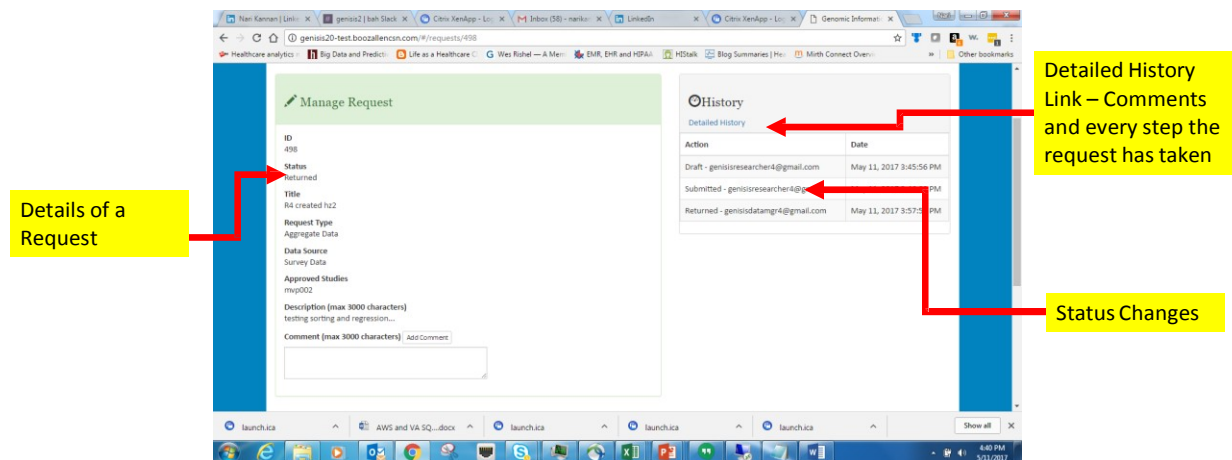


Figure 4439: Navigating a Single Request

5.2. DBMS Files

Table 13 summarizes the Databases involved in the Genisis2 Application.

Table 13: Genisis2 Databases

Database Name	Description	Type	Steward
Genisis_Activiti_DB	All workflow related data is handled in tables created by the Activiti workflow system. Contains workflow data only.	Relational Database	ORD
Genisis_DB	The Main Genisis Database that handles all data not handled by the Workflow System above (such as Comments table)	Relational Database	ORD
TS_DB	The main Terminology Services Database that handles user related data such as Bookmarks, Labels, etc.	Relational Database	ORD

5.3. Non-DBMS Files

Terminology Services creates a Triple Store database that stays in memory after all the Ontologies are loaded. The Ontologies are in RDF format.

5.4. Data View

Table 14 provides a high-level view of the tables in each of the databases involved:

Table 14: Genesis2 Tables in the Genesis2 Databases

Genesis_DB	Genesis_Activiti_db	TS_DB
Comment_History	ACT_EVT_LOG	Bookmarks
Management_Table	ACT_GE_BYTEARRAY	Concept_mapping
Request	ACT_GE_PROPERTY	Concept_mapping_data_elements
Request_History	ACT_HI_ACTINST	Data_element
RequestType	ACT_HI_ATTACHMENT	Data_element_components
Role_Type	ACT_HI_COMMENT	Data_element_source
Source	ACT_HI_DETAIL	Data_element_type
Study_Approval	ACT_HI_IDENTITYLINK	Labels
User_Approver	ACT_HI_PROCINST	Simple_data_element
User_Role_Type	ACT_HI_TASKINST	
User_Type	ACT_HI_VARINST	
Users	ACT_ID_GROUP	
Workflow_Status	ACT_ID_INFO	
	ACT_ID_MEMBERSHIP	
	ACT_ID_USER	
	ACT_RE_DEPLOYMENT	
	ACT_RE_MODEL	
	ACT_RE_PROCDEF	
	ACT_RU_EVNT_SUBSCR	
	ACT_RU_EXECUTION	
	ACT_RU_IDENTITYLINK	
	ACT_RU_JOB	
	ACT_RU_TASK	
	ACT_RU_VARIABLE	

6. Detailed Design

6.1. Hardware Detailed Design

Figure 450 depicts a view of the hardware components, network segments, bandwidth and connectivity that make up the existing Genisis environment. Storage and network components may be modified in increments to align with and implement VA infrastructure expansion plans. Refer to Section 2.6.4 to Expand Genisis2 Storage, Bandwidth, and Computing in the [Genisis2 RSD](#).

GenISIS HARDWARE TOPOLOGY (PITTSBURGH)

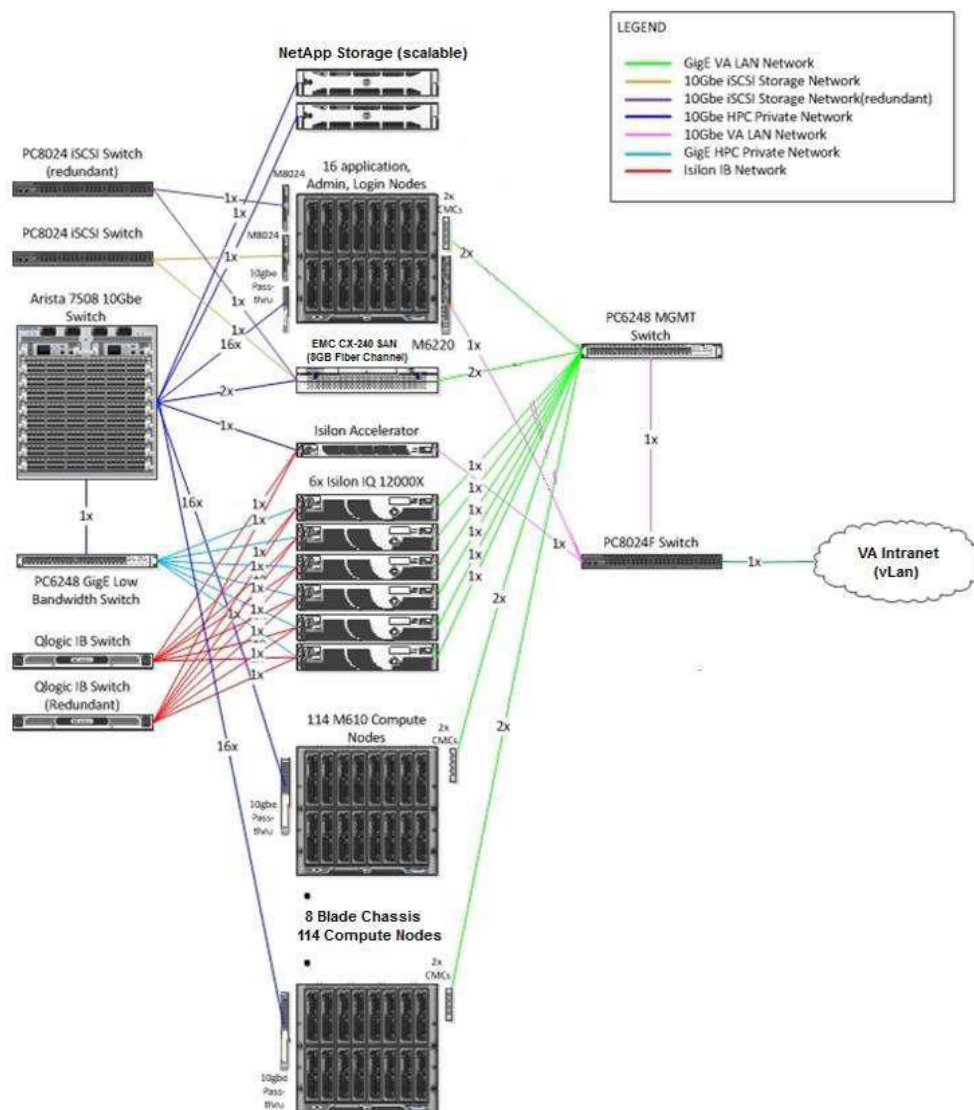


Figure 450: Hardware Topology

6.2. Software Detailed Design

As a platform for Genomic research, Genesis2 consists of Linux and Windows analysis stations and a high performance-computing cluster. Investigators shall be provided with COTS and open source tools used for genomic research. The palette of tools shall be in constant flux reflecting the rapid pace of change in genomic research.

In addition to existing analysis tools, Genesis2 introduces the Unified Genesis System Administrator—Researcher Dashboard, a web application described in further detail in the following sections.

6.2.1. Conceptual Design

The software design is can be divided into two categories:

1. Web Application
2. Web Service

Each web application will follow the same basic design and components, as will web services. Web applications provide connections for web browsers to access web pages implemented in the application. Web services provide connections to custom clients and are not intended for typical user access, only other services and web applications.

Database access will be via JDBC packages that interact directly with the DBMS. Hibernate will be used to provide a higher level of abstraction for managing data stored in the Genesis database. Program-level access to database contents will be provided by the Java Persistence API (JPA) package. JPA functionality will integrate with Hibernate to create a complete path to the database. Using Hibernate and JPA helps alleviate the need to manage every detail of data access. The application can concentrate on the business requirements. These packages manage the creation and update of database tables, along with streamlining data update and retrieval. The JPA mechanisms primarily target the persistence and retrieval of Java-based objects, but JPA can also support complex queries to the database, along with calling stored procedures. These are typically known as native queries and tend to be database specific. Since Genesis does not contain time-critical applications, the performance impact will be minor.

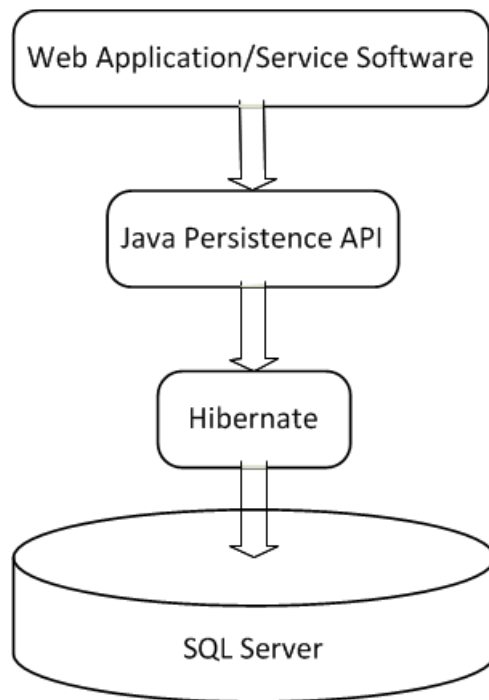


Figure 464: Data Access

All database-related Java objects are defined in a common library that is available to each web application/service. These Java objects are annotated with directives that instruct the framework on how to store and retrieve the objects.

Accounts will be used to prevent unauthorized access to the Genisis database system. Each application/service must present valid credentials. These credentials are stored locally and are protected from external access. Sensitive files will be placed in a location that is unavailable to external access, except for a local system administrator login. Both web applications and services are designed to operate in standard servlet containers. For Genisis, the container will be Wildfly. The primary development language is Java.

The Wildfly server will be configured to limit access from outside of the host computer to increase security of the applications residing on Wildfly. Wildfly will only listen on the localhost address. All outside access will occur via the Apache HTTP Server. Wildfly will interact with the Apache HTTP Server via an AJP protocol-based connection. Applications and services running on Wildfly will assume that identities of all external users have been authenticated. Wildfly will use these trusted identities to authorize access to each web page or service endpoint. Identities will be transferred using the HTTP header (REMOTE_USER).

Web services will follow the ReST model of interfacing. ReST provides a simplified access to a web service via HTTP methods of GET, PUT, POST, and DELETE. Spring Framework MVC supports building RESTful services and associated clients. Using annotations methods are associated with HTTP methods and URLs. The Genisis web services will exchange XML documents defined in a schema.

6.2.1.1. Product Perspective

Each web application is built upon various open source packages. The primary package is a subset of the overall Spring Framework:

- Spring Framework
- Spring Webflow
- Spring Security
- Spring LDAP

Web page implementation is provided by Java Server Faces (JSF) and Primefaces. Primefaces provides a set of widget definitions for common web application controls that either enhance existing JSF controls or implement new functionality. In addition, web pages will contain Hyper Text Markup Language (HTML) and Java Facelets constructs. Asynchronous JavaScript and XML (AJAX) will be used to enhance user interactions with the web pages. XML manipulation is provided by XMLBeans and JAXB packages, which provide XML parsing and generation based on schemas. All web applications follow the same basic architecture. External web page requests are routed through Spring's Model View Controller (MVC) architecture to the appropriate web page, based on the uniform resource locator (URL) entered. Spring Security examines the URL to determine if the user is authorized to access the requested resource. Security authorization will be implemented by each application for each resource it exposes. Authorization is role-based. User roles will be stored in a local Lightweight Directory Access Protocol (LDAP) server. The local LDAP server will be queried by the web application to gather all assigned roles when the user first enters the application. Access to the local LDAP is implemented using the Spring LDAP extension. The web application security will compare the user's roles to the required role for the resource. If the user is not authorized, a page indicating the error will be sent back to the user instead of the requested resource. Each application will be designed to avoid authorization errors by only presenting options that the user is authorized to access. Menu options will be presented based on user roles. After authorizing access, Spring will direct the request to the web flow processing.

Spring Webflow is an extension to the Spring Framework that implements a state-machine architecture that manages views, which are basically web pages. Each view is associated with a flow. The flow defines the transitions between views, which guides a user through an execution of a business task. A flow may have one or more views associated with it. An application consists of one or more flows. Genesis uses Webflow to enforce transitions between views when necessary.

Generally, each web flow is backed by one or more Spring beans. The Spring beans provide the business logic behind the view and access to the data sources. The beans are typically defined in the flow configuration. When the application enters a flow, any defined beans will be created according to the rules of the flow. They will then be available to view within the workflow. These beans exist during the life of the flow. There are other beans, such as data access interfaces, which exist during the life of the application. For example, when information is shown in a table, a "navigation" bean is defined. The navigation bean provides objects, such as data models, to support the table displayed in the view. The model contains the list of data elements shown in the view. When the data sets are potentially large, lazy loading will be implemented to enhance the user's experience.

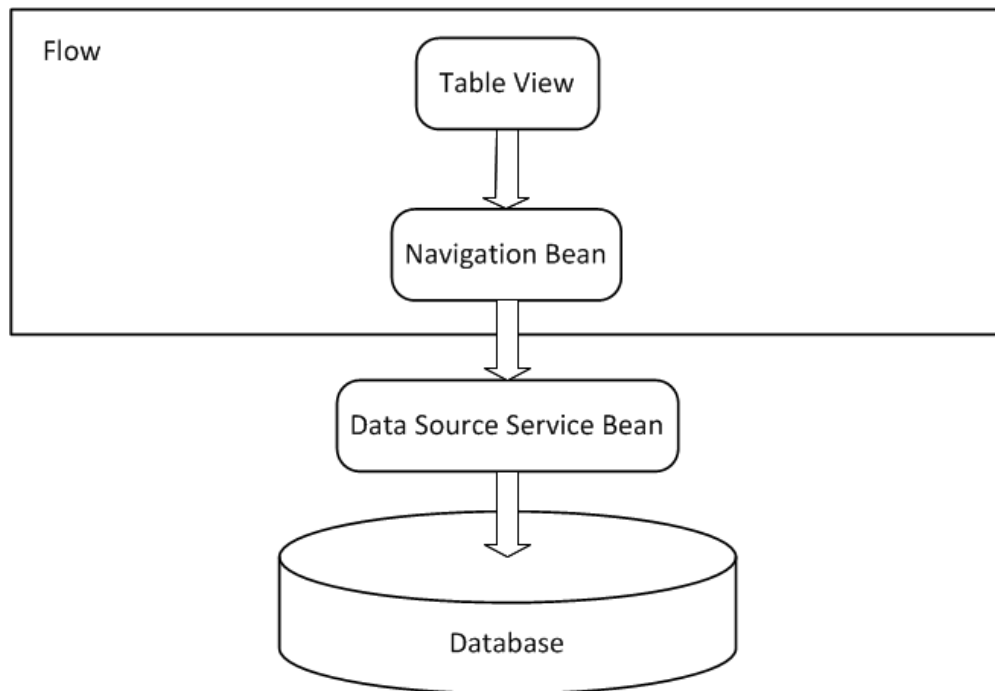


Figure 472: Web Flow

This is a typical architecture. For the case where data is entered, a “management” bean is defined. Again, the purpose is to provide the business logic and data for the view or views. When business flows require more than one view, the flow defines the transitions between views. Based upon user interactions, a view may return a new state. Figure 18 illustrates the concept. When the flow is entered, the management bean is created. The management bean lives during the entire flow. The first view shown to the user is “View 1.” After satisfying the requirements of the first view, the user presses, for example, “Next.” The view defines the “Next” button and results in a “next” state, by returning that value to the flow. The flow has defined the “next” state for “View 1” as “View 2.” Spring Webflow then presents “View 2” to the user. This pattern is followed throughout the flow. After completing the requirements for the flow, the flow is exited, and we return to a top-level flow, which is typically a landing page for the application.

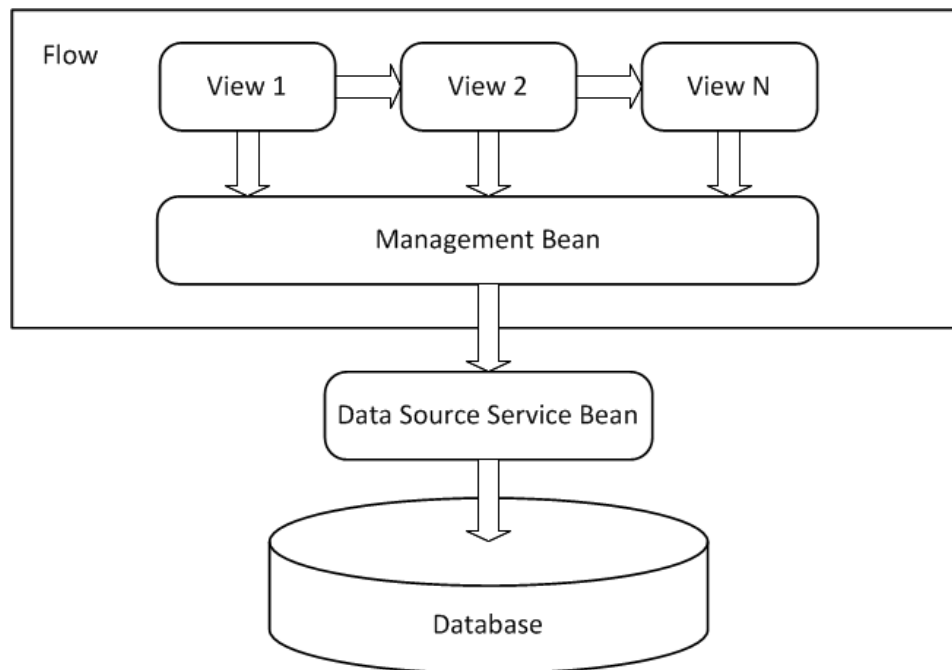


Figure 483: Wildfly Servlet

Each flow extends a parent flow definition. The parent flow defines common transitions for all the child flows.

6.2.1.1.1. User Interfaces

Genesis2 enhances the existing user interface (UI) configuration through the new functions of the Unified Genesis System Administrator – Researcher Dashboard. This enhanced UI serves as a portal for both Researcher and system administrator activities. The dashboard will help Researchers request study data and track the status of their study throughout the lifecycle, and allow system administrators to monitor the Genesis2 application. Diagrams illustrating UI characteristics for the dashboard design are included the following figures in this document as of this Build.

- ~~3.3.4.1.1 Researcher Landing Page Dashboard Design~~
- ~~3.3.4.1.2 Genesis Data Manager/VINCI Data Manager Landing Page Dashboard Design~~
- ~~3.3.4.1.3. Genesis System Administrator Landing Page Dashboard Preliminary Design~~
- ~~3.3.4.2.1 Terminology Services Landing Page~~
- ~~3.3.4.2.2 Search Results~~
- ~~3.3.4.3.3 Concept Card Details~~
- ~~3.3.4.3.4 Bookmark a Concept~~
- 3.3.4.1 Genesis Data Request Workflows
- 3.3.4.2 Terminology Services Workflows

6.2.1.1.2. Hardware Interfaces

There are no hardware interfaces involved with Genesis2. This section is N/A.

6.2.1.1.3.□□□□□□ Software Interfaces

Genisis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, **Genisis2_VIP_Build 86_API_Manual_0118201812222017.pdf**.

6.2.1.1.4.□□□□□□ Communications Interfaces

There are no communication interfaces involved with Genisis. This section is N/A.

6.2.1.1.5.□□□□□□ Memory Constraints

There are no memory constraints involved with Genisis. This section is N/A.

6.2.1.1.6.□□□□□□ Special Operations

There are no special operations involved with Genisis. This section is N/A.

6.2.1.1.7.□□□□□□ Product Features

Overall planned product features and those implemented in Builds 1 through 5 are covered in Sections 1.1.1 and 1.1.2 of this document.

6.2.1.1.8.□□□□□□ User Characteristics

New requirements for Genisis2 include infrastructure expansion to support MVP, a program to enroll up to one million users of the VHA into an observational cohort. The goal is to support robust genomic and bioinformatics-related data analysis, including over 50 concurrent studies, and translating to over 100 research users and 500 computational jobs. Genisis shall serve the VA genomic research community, whose members have working experience with academic computational platforms. Specialized user support, for both genomic Researcher and Genisis System Administrator groups, is implemented as the Unified Genisis System Administrator—Researcher Dashboard requirement.

6.2.1.1.9.□□□□□□ Dependencies and Constraints

There are no dependencies or constraints involved with Genisis. This section is N/A.

6.2.1.1.10.□□□□□□ Database Repository

The Database Repositories and individual tables for Genisis2 are outlined in Section 5.3.

6.2.1.1.11.□□□□□□ System Features

System Features are extensively described in sections 4 and 5 of this document – The overall goals, the workflow, status changes, user interface screens, etc.

6.2.1.1.12.□□□□□□ Design

Element Tables 6.2.1.1.13.

Routines (Entry Points)

Genisis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, **Genisis2_VIP_Build 86_API_Manual_0119201812222017.pdf**.

6.2.1.1.14.□□□□□ **Templates**

Genesis2 does not use any templates. This section is N/A.

6.2.1.1.15.□□□□□ **Bulletins**

Genesis2 does not use any bulletins. This section is N/A.

6.2.1.1.16.□□□□□ **Data Entries Affected by the Design**

Genesis2 does not use any variables or constants that affect design or its execution. This section is N/A.

6.2.1.1.17.□□□□□ **Unique Record(s)**

Genesis2 requirements do not use unique records in the design or execution. This section is N/A.

6.2.1.1.18.□□□□□ **File or Global Size Changes**

Genesis2 requirements do not use file or global sizes in the design or execution. This section is N/A.

6.2.1.1.19.□□□□□ **Mail Groups**

Genesis2 does not use file or global sizes in its design or its execution. It uses users VA email addresses for notifications of workflow status changes, but does not use mail groups per se. This section is N/A.

6.2.1.1.20.□□□□□ **Security Keys**

Genesis2 users their VA PIV Cards for authentication and the VA LDAP. There are no other security keys involved. This section is N/A.

6.2.1.1.21.□□□□□ **Options**

Genesis2 requirements do not use options in its design or its execution. There is an online help section that is available right next to the User Login/Profile details. That provides help regarding the User Interface. This section is N/A.

6.2.1.1.22.□□□□□ **Protocols**

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, **Genesis2_VIP_Build 86_API_Manual_0119201812222017.pdf**. Protocol involved is a **RESTful API** call. It is documented in detail in this document.

6.2.1.1.23.□□□□□ **Remote Procedure Call (RPC)**

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, **Genesis2_VIP_Build 6_API_Manual_12222017Genesis2_VIP_Build 8_API_Manual_01192018.pdf**. Protocol involved is a **RESTful API** call. It is documented in detail in this document

6.2.1.1.24. Variables Defined in Interface

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, [Genesis2 VIP Build 6 API Manual 12222017](#)[Genesis2 VIP Build 8 API Manual 01192018](#).pdf. Variables involved are documented in detail in this document

6.2.1.1.25. Types Defined in Interface

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document, [Genesis2 VIP Build 6 API Manual 12222017](#)[Genesis2 VIP Build 8 API Manual 01192018](#).pdf. Types involved are documented in detail in this document

6.2.1.1.26. GUI

The Graphical User Interface elements are described in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.27. GUI Classes

The Graphical User Interface classes are described in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.28. Current Form

Forms are not used in Genesis2. This section is N/A.

6.2.1.1.29. Modified Form

Forms are not used in Genesis2. This section is N/A.

6.2.1.1.30. Components on Form

Forms are not used in Genesis2. This section is N/A.

6.2.1.1.31. Events

Events are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.32. Methods

Methods are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.33. Special References

Special References are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.34. Class Events

Class Events are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.35.□□□□□ Class Methods

Class Methods are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.36.□□□□□ Class Properties

Class Properties are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.37.□□□□□ Uses Clause

Uses Clauses are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.38.□□□□□ Forms

Forms are not used in Genesis2. This section is N/A.

6.2.1.1.39.□□□□□ Functions

Functions are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.40.□□□□□ Dialog

Dialogs are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.41.□□□□□ Help Frame

Help frames are outlined in detail in the automated Java Docs generated by the environment automatically at any time. Not applicable in this section.

6.2.1.1.42.□□□□□ HL7 Application Parameter

HL7 is not used in Genesis2 anywhere. This section is N/A.

6.2.1.1.43.□□□□□ HL7 Logical Link

HL7 is not used in Genesis2 anywhere. This section is N/A.

6.2.1.1.44.□□□□□ COTS Interface

The only COTS interface involved is that the JAVA code makes JDBC calls to the database server, Microsoft Windows 2012 R2/Microsoft SQL server 2012. This follows standard JAVA and Microsoft SQL query formats.

6.3. Network Detailed Design

Section 4.1 and Figure 27 cover this topic.

6.4. Security and Privacy

6.4.1. Security

Genesis2 will adhere to all VA security requirements in accordance to VA Directive and Handbook 6500, Federal Information Processing Standard (FIPS) 199 and National Institute of

Standards and Technology (NIST) Special Publication (SP) 800-60, recommended Security Categorization. Genesis2 is housed at the Pittsburgh Information Technology Center (PITC), and information security controls are covered under the Pittsburgh facility General Support System (GSS). The implementation of a Continuous Readiness in Information Security Program (CRISP) is in place in order to review key security controls for regular and continuous monitoring.

The Genesis2 Security Categorization will drive the initial set of minimal security controls required for the information system. Minimum security control requirements are addressed in NIST SP 800-53, Revision 4, and VA Handbook 6500, March 2015, Appendix C: (References), Appendix E: (VA System Privacy Controls), and Appendix F: (VA System Security Controls).

For additional information, please reference Appendix C of the Genesis2 Requirements Specification Document (RSD).

6.4.2. Privacy

Genesis2 will adhere to all proposed VA Privacy requirements and controls, Identity Management and Security requirements, including VA Handbook 6500, March 2015, Appendix C: (References), Appendix E: (VA System Privacy Controls), NIST SP 800-53, Revision 4, Privacy Controls; any privacy or data security constraints that should be addressed in accordance with VA directives and HIPAA Privacy Act. Efforts that involve the collection and maintenance of Personal Identifiable Information (PII) must be covered by a Privacy Act system of records notice.

When a Privacy Impact Assessment (PIA) has been completed for Genesis2, a link to the PIA will be added.

6.5. Service Oriented Architecture / ESS Detailed Design

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document – [Genesis2_VIP_Build_6_API_Manual_12222017Genesis2_VIP_Build_8_API_Manual_01192018.docx](#).

7. External System Interface Design

Genesis2 has a Service Oriented Architecture and has an Application Programming Interface (API) that even its own Browser interface calls. Documentation for this is generated automatically. Details on individual calls and parameters for each are documented in a separate document – [Genesis2_VIP_Build_6_API_Manual_12222017Genesis2_VIP_Build_8_API_Manual_01192018.docx](#).

8. Attachment A – Approval Signatures

This section is used to document the approval of the System Design Document. The review should be conducted face to face where signatures can be obtained ‘live’ during the review. If unable to conduct a face-to-face meeting, then it should be held via LiveMeeting and concurrence captured during the meeting. The Scribe should add /es/name by each position cited. Example provided below.

The Business Sponsor and Project Manager are required to sign.

Sumitra Muralidhar
Business Sponsor

Date:

Katie Thomas
Project Manager

Date:

A. Additional Information

A.1. Identification of Technology and Standards

Details of this can be found in Section 4.5. It has a table that lists the components used in Genisis2 and their TRM approval statuses.

A.2. Constraining Policies, Directives and Procedures

Genisis2 has no additional constraining policies, directives, or procedures. This section is N/A.

A.3. Requirements Traceability Matrix

The Genisis2 RTM will be updated with the new requirements using IBM Rational Tools.

A.4. Packaging and Installation

The **Genisis2 Deployment Guide** contains extensive systematic instructions to set up any of the environments for this project – **DEV/TEST/SQA/PRE-PROD/PROD**.

A.5. Design Metrics

Design metrics and approaches are outlined in Section 3.1 in this document.

Template Revision History

Date	Revision	Description	Author
June 2015	2.10	Changed Heading 1 default setting to eliminate page break before	Process Management
May 2015	2.9	Edited for Section 508 conformance and remediated with Common Look Office tool	Process Management
February 2015	2.8	Incorporates revisions from PMAS Reform Lockdown; namely removing requirements for information that can be obtained from other PMAS authoritative sources.	Andrew Slawter, Office of Technology Strategies
September 2014	2.7	Adds Enterprise Shared Services terms and requires AERB Compliance Certificate attachment.	Process Management
August 2014	2.6	Signature block update authorized by AERB CR_018934	Process Management
March 2014	2.5	Section 508 repairs to new version approved by AERB Chair approved	Process Management
August 2013	2.3	Replaced the Service Architecture sub-section with new sub-sections for consumed and provided services. Also, applied miscellaneous feedback from VA team.	ASD Enterprise Shared Services (ESS) Work Group
June 2013	1.3	Upgraded to MS Office 2007-2010 format	Process Management
June 2013	1.2	Address inconsistencies in Section 3, Conceptual Design, Correct headings	Process Management
March 2013	1.1	Formatted to documentation standards and edited for Section 508 conformance	Process Management
January 2013	1.0	Initial Document	PMAS Business Office

See TOGAF® 9.1, Part III: ADM Guidelines & Techniques, Gap Analysis on TOGAF website at <http://pubs.opengroup.org/architecture/togaf9-doc/arch/chap27.html>