

## Bioinformatics Integration Support Contract (BISC)

# SYSTEM ARCHITECTURE AND SOFTWARE DESIGN SPECIFICATION



**Version 5.0**

**Period Of Performance: September 30, 2012 to September 29, 2017**

**Project Sponsor:**

National Institutes of Health (NIH)  
National Institute of Allergy and Infectious Diseases (NIAID)  
Division of Allergy, Immunology, and Transplantation (DAIT)

**Prepared by:**

***NORTHROP GRUMMAN***

Northrop Grumman Corporation  
Information Systems Sector  
Health IT  
2101 Gaither Road, Suite 600  
Rockville, MD 20850  
(301)527-6400 Main  
(301)527-6401 Fax  
[Jeff.Wiser@ngc.com](mailto:Jeff.Wiser@ngc.com)

## Contents

<b>1.0 Introduction .....</b>	<b>3</b>
1.1 Scope .....	3
1.2 Purpose .....	3
1.3 Background.....	3
<b>2.0 ImmPort Tools on Amazon Web Services.....</b>	<b>5</b>
2.1 Open ImmPort .....	5
2.2 Beta and Development Applications .....	13
2.3 Hosted Applications: ImmuneSpace.....	18
<b>3.0 ImmPort Tools at NIAID .....</b>	<b>19</b>
3.1 Features Table .....	19
3.2 Hardware Infrastructure.....	21
3.3 Database Architecture.....	23
3.4 Aspera Server .....	23
3.5 ImmPort 3.0.....	24
3.6 ImmPort 2.0.....	37

**BISC System Architecture and Software Design Specification Version History**

Version	Date	Description
1.0	4/29/2005	Initial release of the BISC Phase II System Architecture and Software Design Specification
1.1	6/1/2005	Completed Sections 3, 11, 14, 15, and 16, including addition of data models to Section 3. Updated sections 12 and 13. Added Section 17. Added acronym list to Appendix A. Added preliminary design and storyboards for future ImmPort versions to Appendix B.
1.2	12/28/2005	<p>Revised the Purpose section to note that most requirements for the ImmPort system version 1.0 were derived primarily from working with the Population Genetics Program (Section 1.2).</p> <p>Updated assumptions (Section 1.3).</p> <p>Updated references (Section 1.6).</p> <p>Updated descriptions and diagrams of hardware and software architecture (Section 2).</p> <p>Updated the database design to reflect the current database instances and servers across the development environment, test environment, and production environment (Section 3).</p> <p>Updated the previous Experimental and Reference Data Models Description: Supplement to System Architecture and Software Design Specification Version 1.1 and incorporated it into the sections on experimental data and reference data models (Sections 4 and 5, respectively).</p> <p>Added an ImmPort Ontology Browser schema (Section 5).</p> <p>Made minor changes to correct design diagrams in the following sections: 7, 8, 9, 10, 14, 16, and 17.</p> <p>Updated prototypical mock-ups to ImmPort system screenshots in the following sections: 7, 8, 9, 10, 11, 13, 14, 16, and 19.</p> <p>Expanded reference data query use cases and corresponding screenshots (Section 16).</p> <p>Removed preliminary design artifacts related to an upload wizard approach to experimental data (SASDS v.1.1, Appendix B).</p>
2.0	6/1/2007	<p>Updated section 1.2 Purpose - includes updated information regarding new sections and update details.</p> <p>Updated section 1.4 Key Objectives - updated relevant version information and clarification of objectives.</p> <p>Updated section 2.1 System Architecture Overview.</p>
2.2	2/6/2008	<p>Updated section 5.0 Reference Data Models.</p> <p>Updated section 3.2.1 Overall Database Design - provided more detail.</p> <p>Updated Table 3-1 ImmPort Database Schemas.</p>
3.0	10/29/2010	Re-factored the majority of the document to reflect updated data model and definitions, as well as updating the software design section to reduce redundancy and add new functionality.
4.0	11/01/2011	Updates to reflect changes in admin, flow cytometry, SFVT (Section 19) areas
5.0	3/24/2016	Rewritten for AWS instance, removal of old code, etc.

## 1.0 INTRODUCTION

### 1.1 SCOPE

The scope of the Bioinformatics Integration Support Contract (BISC) is to provide advanced information technology support in the production, analysis, archiving, and exchange of scientific data for a diverse community of life science researchers.

### 1.2 PURPOSE

The Immunology Database and Analysis Portal (ImmPort) applications have been developed by a team led by Northrop Grumman Information Systems (NGIS) with academic partners from the University of Texas-Southwestern in Phase I, and the University of California San Francisco in the current contract. The ImmPort system is intended to serve as a long-term, sustainable archive of data generated by investigators funded through the Division of Allergy, Immunology and Transplantation (DAIT) of the National Institute of Allergy and Infectious Disease (NIAID), National Institutes of Health (NIH). The ImmPort system consists of an extensive data warehouse containing an integration of experimental and clinical data supplied by NIAID/DAIT-funded investigators. The ImmPort system is freely accessible as a resource to all scientists in the research community.

This System Architecture and Software Design Specification (SASDS) defines the overall ImmPort architecture and software design specification identified by the NGIS BISC Team (hereinafter referred to as the BISC Team) for the ImmPort system that has been developed for NIAID/DAIT. The architecture and design described in this document focuses on the capabilities that are implemented in the ImmPort family of applications as of March 31, 2016.

### 1.3 BACKGROUND

The key objective of the SASDS version 5.0 is to provide an update to the hardware and software specifications of the system. The ImmPort project has evolved, such that systems are now hosted in a production or near production mode at both the NIAID hosting facility as well as Amazon Web Services (AWS) cloud environment. In general, the long term goal, which is in progress, is to provide hosting of data while it is private and being QC'ed and curated in ImmPort related applications housed at the NIAID hosting facility. When data is shared to the general scientific community, data would be transferred to AWS for easier re-use of these data for analysis, or in short, to bring the data to the analysis tool. This bifurcation of systems allows BISC and NIAID staff to maintain maximum control over data while it is sensitive and private, and more flexibility for re-use and distribution when the data is shared in AWS.

(Note: Some exceptions may occur to this policy, such as the current Data Browser that is hosted at NIAID, which allows downloading of shared research and clinical data. This decision to also host shared data for download at NIAID is driven by the desire to allow an alternative to accruing data transfer costs from AWS.)

Since the last version of the SASDS document, there has been a change in focus in the ImmPort project to lessen the level of effort spent on the development, maintenance and sales/outreach for analysis tools and reference data capabilities. In 2014-2015, based on usage statistics a decision was made in tandem with NIAID program staff to gradually remove analysis tools such as the HLA analysis pipeline, TagSNP, GenePattern and PED data generator as well as the reference data searches (Gene, Protein, HLA, SNP, etc) due to a lack of usage. The FLOCK flow analysis tool suite continued to have constant usage and increased interest and publications related to tool usage and results, so the decision was made to continue support of that application. As a result of these decisions, the retired tools and queries will not appear in this design document.

For the remaining features of ImmPort, it was recognized that a general code refresh was necessary given the overall age of the software and supporting stack of frameworks. As a result, the BISC team has progressed incrementally through the upgrade of the features into a new software architecture that is detailed in this document. For the purposes of this document, the newer code architecture is referred to as "ImmPort 3.0", while the prior architecture being gradually replaced is referred to as "ImmPort 2.0". This document will detail the ImmPort 3.0 architecture for features that have been upgraded or will soon be upgraded, and will keep the existing documentation in place for ImmPort 2.0 features not yet upgraded in the production environment. As features are completed, this document will be accordingly updated during 2016 and 2017. The functional requirements documents for features that are in progress for upgrade will also be available and referenced in this document.

Compared to prior versions of this document that detailed the ImmPort 2.0 architecture, the overall system architecture for ImmPort 3.0 is being simplified to have a less dense middle-tier. As a result, the need to document detailed design packages in this SASDS is reduced, since the same middle-tier approach is utilized across the features in a given application and the EJB tier has been removed.

The database documentation has been moved online, so is no longer described in detail in this document. References to the freely available online materials will be provided in this document. Far more detail about the database fields, tables and ERD diagrams are available online than in prior versions of the SASDS, so the overall amount of information has increased markedly. Additionally, since the SASDS document was released the entire database in MySQL is available for anyone to download and re-use, making comprehension of the database architecture much simpler.

This document is to be considered a “work in progress” and will evolve during the life of the BISC effort as additional requirements are implemented, new requirements are identified, and others are modified or deleted.

## 2.0 IMMPORT TOOLS ON AMAZON WEB SERVICES

The ImmPort tools deployed on Amazon Web Services (AWS) are designed primarily to identify studies of interest for users to evaluate for future analysis. Currently, the application performing this feature is Open ImmPort. In addition to Open ImmPort, the AWS infrastructure is utilized to develop Alpha and Beta tools to obtain feedback on utility from the user community. Tools such as ImmuneXpresso and the Cell Ontology browser fit into this category. Finally, the AWS infrastructure is being used to host production applications developed by other research teams funded by DAIT without the funding to support a federal system. ImmuneSpace is the first example of this usage of BISC resources.

### 2.1 OPEN IMMPORT

Open ImmPort is the application utilized by researchers to identify studies of interest for further analysis. As of March 31, 2016, 211 studies have been shared and are cataloged in Open ImmPort. Currently, no row level results are viewable from within Open ImmPort, with the advantage that no user authorization is then required. Open ImmPort utilizes a standard suite of Linux-based AWS servers and components described further below and a Spring/Java and JavaScript web application architecture.

#### 2.1.1 Feature Summary

**Table 2.1-1** below summarizes the major functionality of Open ImmPort.

**Table 2.1-1: Summary of ImmPort Capabilities and Features**

#	Capabilities/Features	Capability/Feature Description
1	Search for Studies	Allows users to perform a “Google like” search to identify shared studies of interest. The text entered by the user is searched against an index of the entire set of metadata stored in the ImmPort database.
2	View/Filter Query Results	After initial search, users are presented with a list of studies, summary information about each study, and the search hits. From here, users may utilize facets on the left hand panel to further filter the studies returned by categories of data such as assay method, species, sample type, etc.
3	View Study Details	Once a study of interest is identified, users are able to view extensive metadata and summary data about the study.
4	Visualize Summary Data	For selected aspects of the study data such as demographics, users are able to filter and visualize data in standard plots such as bar charts, line charts by factors such as gender, ethnicity, and arm
5	Link to download data	For a given study, users can click to download the raw study data, and are redirected to login to the Data Browser application at NIAID.
6	View Reference and Static content	Static content such as Tutorials, curated cytokine lists, system documentation, user documentation and ImmPort and BISC project information is hosted on the Open ImmPort site.

### 2.1.1.1 Home Page

The Open ImmPort home page provides the entry-point into Open ImmPort. Beyond being an information page containing announcements and ImmPort background information, users can right away begin searching and filtering for studies of interest. In addition, studies that may be of interest to the user community are highlighted in the slider bar based on factors such as recent publications, popularity of the study based on views or downloads, and recent additions to the shared ImmPort content.

The screenshot shows the Open ImmPort home page. At the top left is the logo and the text "OPEN IMMPORT". To the right is a sidebar titled "Announcements" with a message about the new Beta web site. Below the sidebar is a "Data Summary" section. A main content area features a study profile for "Human T Cell Profile". The profile includes a brief description of T cell differentiation and maintenance, a heatmap showing CD28 vs CD127 expression across various tissues, and a PubMed ID (25417158 Study: SDY702).

What is ImmPort?

OpenImmport is your resource for searching and downloading shared biomedical research data funded from NIAID DAIT and DMID, and other NIH and non-government funding agencies. Additional resources available at OpenImmport include step-by-step data reuse tutorials including example R and Python analysis code, the Cell Ontology Visualizer, the Cytokine Registry, and ImmuneXpresso - the cytokine and cell interaction literature mining tool.

ImmPort, is the data submission portal where researchers upload, QC, and curate their data prior to sharing in OpenImmport.

Download Study Data for Analysis

Featured Studies   Study Filter   Study Explorer

e.g. influenza

Study: Human T Cell Profile

Understanding of T cell differentiation and maintenance originates primarily from observations of the small percentage that reside in peripheral blood. In this study, viable T cells were characterized in immune-system tissues from organ donors. The lineage and migration of memory CD4+ and CD8+ T cells were investigated in donors aged three to 73. Distinct populations in blood, lymphoid, and mucosal tissues illuminate T cell function and changes throughout the human lifespan.

CD28

Blood Spleen ILN Lung MLN Ileum Colon

CD127

	CD4+ T cells	CD8+ T cells
Blood	23	19
Spleen	44 12	14 2
ILN	84 16	2 1
Lung	83 17	0 20
MLN	42 22	21 1
Ileum	74 4	3 0
Colon	82 14	14 1
	80	5

CD28

CD127

CD4+ T cells

CD8+ T cells

PubMed ID: 25417158 Study: SDY702

### 2.1.1.2 Search Bar and Results

Once the user has selected to view either all or a subset of studies, a list of studies are presented as shown below. The layout of the page is a familiar design with faceted search capabilities illustrated in the left-hand panel, and a “Google-like” simple text search bar at the top. From this page, users may either select a study to view more details, filter the study list further based on facets, click to view a larger version of the study schematic graphic, or click to download the study data. Clicking on the Download button directs a user to the Data Browser application in the directory for that study after authentication.

e.g. influenza

Filter Option

Clinical Trial

- N (117)
- Y (46)

Study Type

- Intervention Longitudinal (7)
- Interventional (80)
- Longitudinal (13)
- Observational (63)

Research Focus

- Atopy/Allergy (16)
- Autoimmune (12)
- Immune Response (44)
- Infection Response (23)
- Transplantation (16)
- Vaccine Response (58)

Species

- Anas platyrhynchos (1)
- Gallus gallus (1)
- Homo sapiens (114)
- Macaca fascicularis (1)
- Macaca mulatta (1)
- Mus musculus (44)
- Mustela putorius furo (1)
- Sus scrofa domesticus (2)

Biosample Type

- Bodily fluid (73)
- Cell (59)
- DNA (10)
- Not\_Specified (4)
- Organ (8)

Found 163 Studies in 1859 ms

« 1 2 3 4 5 »

**SDY1**  [Efficacy and Safety Evaluation of Allergen Immunotherapy Co-Administered with Omalizumab \(an anti-IgE Monoclonal Antibody\)](#)

[Download](#)

A series of allergy shots may reduce symptoms of seasonal ragweed allergies. This study will determine whether taking a drug called omalizumab (also known as Xolair) before getting the allergy shots is more effective than allergy shots alone or other treatments, such as prescription antihistamines.

Thomas Casale, Creighton University School of Medicine

**SDY10**  [Role of Antimicrobial Peptides in Host Defense Against Vaccinia Virus](#)

[Download](#)

Atopic dermatitis (AD) is a chronic inflammatory skin disorder characterized by recurrent viral skin infections. Recent studies have demonstrated that the skin of people with AD may have decreased antimicrobial peptide (AMP) expression. The purpose of this study is to compare small pox virus replication and the number of AMPs and other antiviral molecules in people with AD, as compared to those ...

Donald Leung, National Jewish Health

**SDY100**  [PPARG and microRNA-146 in mucosal immune responses to C. difficile](#)

[Download](#)

Clostridium difficile is typically a harmless anaerobic bacterium but recently it has re-emerged as a facultative pathogen that can cause nosocomial diarrhea, colitis and even death. Peroxisome proliferator-activated receptor (PPAR) gamma has been implicated in the prevention of inflammation in autoimmune and infectious diseases; however, its role in the immunoregulatory mechanisms modulating ...

Monica Viladomiu, VBI



### 2.1.1.3 Study Detail Page

When a study is selected, the user is presented with a study detail page. Within this page, there are multiple tabs presented with different aspects of study data based on what has been provided. The full set of tabs that may be displayed include:

- Summary: title, description, PI, type, arms/cohorts, study schematic
- Study Design: study timeline, inclusion and exclusion criteria, schedule of events
- Adverse Event: summary of adverse event data by severity, name, arm
- Assessment: summary of assessments taken
- Interventions: summary of interventions performed
- Medications: concomitant medications taken
- Demographics: summary information about gender, age, ethnicity
- Lab Tests: summary of laboratory panels and tests
- Mechanistic Assays: summary of assays performed, protocols, platforms, reagents, treatments
- Study Files: catalog of study data files provided

The screenshot shows the ImmPort interface for Study SDY1. The top navigation bar includes links for ImmPort, Resources, Studies, Submission, Documentation, Publications, Help, and About. The main content area displays study details under the heading "Study SDY1". A banner at the top states "Study data available for download for Registered Users". Below this, a tab navigation bar includes Summary (selected), Design, Adverse Event, Assessment, Interventions, Medications, Demographics, Lab Tests, Mechanistic Assays, and Study Files. The "Summary" section contains fields for Accession (SDY1), Title (Efficacy and Safety Evaluation of Allergen Immunotherapy Co-Administered with Omalizumab (an anti-IgE Monoclonal Antibody)), PI (Thomas Casale - Creighton University School of Medicine), Type (Interventional), Condition Studied (Seasonal allergy to ragweed), Brief Description (A series of allergy shots may reduce symptoms of seasonal ragweed allergies. This study will determine whether taking a drug called omalizumab (also known as Xolair) before getting the allergy shots is more effective than allergy shots alone or other treatments, such as prescription antihistamines.), Start Date (2003-04-01), and Schematic (button labeled "Hide"). The "Schematic" section contains a detailed study design diagram. The diagram illustrates two groups of subjects: "IT + anti-IgE" and "Placebo IT + anti-IgE". These groups undergo a "Pretreatment" phase from Week -12 to Week -9, followed by a "Ragweed Season" from Week 1 to Week 12. During the Ragweed Season, subjects receive "RIT Maintenance IT + study drug". At the end of the season, there is a "Follow-up" period and a "Specimen Collection". Arrows indicate the flow between the Pretreatment and Ragweed Season phases, and between the Ragweed Season and Follow-up.

## 2.1.2 Hardware and Software Architecture Components

**Table 2.1.2-1** below describes the hardware components and software components that constitute the architecture of Open ImmPort.

**Table 2.1.2-1: Hardware and Software Architecture Components**

Component Name	
Model View Controller (MVC)	A software design pattern for implementing user interfaces, often web based applications. The architecture divides the software into three interconnected domains in order to separate data representation from the presentation of the data to the user.
Business Logic/Services	The controller layer that converts data from the data access tier to the presentation/view tier and executes any necessary business logic.
Persistence/Data Access	The model tier that manages the connectivity to the data layer.
Presentation	The view layer that organizes data for presentation to the user.
RESTful web services	Based on Representational state transfer (REST) architecture and are scalable, light weight and maintainable components often used as web application API's.
Private subnet	In AWS design, instances in the private subnet cannot receive inbound traffic from the Internet and cannot send outbound traffic to the Internet unless provided through a network address translation (NAT) gateway that is housed in the public subnet. Data servers in the private subnet also access the public subnet through the NAT.
Public subnet	In the AWS design, instances in the public subnet can receive inbound traffic from the Internet and can send outbound traffic to the Internet. This would be the location of web servers and any NAT servers directing traffic to a private subnet.
JavaScript Object Notation (JSON)	A lightweight, human-readable data-interchange format widely used in web applications and web-based API's.
Spring Boot	Provides a ready to start Spring-based application deployment that utilizes best practices for the configuration ( <a href="http://projects.spring.io/spring-boot/">http://projects.spring.io/spring-boot/</a> )
Spring Web MVC	Provides model-view-controller architecture (MVC) and components to develop flexible and loosely coupled web applications in the Spring family of components
Aspera Connect Server	Server product developed by Aspera Software that optimizes connectivity speed and reliability over the internet between a client and the server using the proprietary FASP® transport technology. It is widely used for transporting large files, such as sequencing files, over the internet. <a href="http://www.asperasoft.com">www.asperasoft.com</a>
Aspera Connect Client	A free client software package that is a web browser plugin to connect to the Aspera Connect server to optimize connection speed across the internet.
MySQL	A widely used open source relational database system now owned by Oracle corporation. <a href="http://www.mysql.com">www.mysql.com</a>
Angular JS	A popular environment with interactive components for developing dynamic web-based applications using AJAX/JavaScript developed by Google but open source. <a href="https://angularjs.org">https://angularjs.org</a>
AJAX	Asynchronous JavaScript and XML (AJAX) is the use of XMLHttpRequest object to communicate with server-side scripts, widely used as the primary method for interactive web applications
Tiles	Software framework that emanated from Struts allowing developers to create page fragments that are assembled into pages at execution time and also allows for re-use of fragments to reduce redundancies ( <a href="http://tiles.apache.org">tiles.apache.org</a> )
Java Server Pages (JSP)	Framework to assist developers in making dynamically generated web pages based with Java as the backbone language.
Apache Tomcat	The most widely used open source software implementation of the Java Servlet, JavaServer Pages, Java Expression Language and Java WebSocket technologies. ( <a href="http://tomcat.apache.org">tomcat.apache.org</a> )

Component Name	
Apache SOLR	Popular open source enterprise search platform built on Apache Lucene™ utilizing Apache Tomcat as the servlet container (lucene.apache.org)
Google Analytics	Web monitoring tool framework provided by Google.
Amazon Web Services (AWS)	Amazon Web Services (AWS) provides cloud infrastructure and a wide array of server capabilities for developers to build production applications. (aws.amazon.com)
Elastic Compute Cloud (EC2)	Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud, designed to make development easier through ease of server creation, duplication, and shutdown along with providing a suite of configuration options for hardware and software/OS specifications.
Virtual Private Cloud (VPC)	Allows administrators to provision a logically isolated section of the AWS cloud where AWS resources can be created and launched in a customized virtual network
Simple Storage Service (S3)	Provides secure, durable, highly-scalable cloud storage.
Relational Database Service (RDS)	Provides pre-configured database servers for Oracle, SQL Server, PostgreSQL, MySQL and MariaDB where AWS performs the database administration allowing the development team to focus energies on application specific details and development.

### 2.1.3 Data Architecture

The data for Open ImmPort is stored in two technologies: MySQL relational database and Apache SOLR/Lucene text index. The MySQL relational database maps closely to the Oracle database used to collect the research and clinical data that is housed at the NIAID hosting facility, but de-normalized to optimize query performance. The MySQL database is used when retrieving data for the Study Detail pages shown above. The SOLR server provides the technology for the text search for both the Google bar as well as the faceted search. Indices are generated from the majority of the MySQL database schema plus documents provided by the researchers, such as protocols and Case Report Forms. The SOLR server also provides the capability of “type-ahead” in the search box.

An AWS RDS MySQL server is utilized for Open ImmPort to take advantage of Amazon’s management of the database server to reduce maintenance costs. The SOLR server exists on its own EC2 Ubuntu virtual machine. The MySQL database and the SOLR indices are refreshed completely with each data release, which occurs every 2-3 months.

A full data dictionary and ERD diagrams for the MySQL schema are available online as described below.

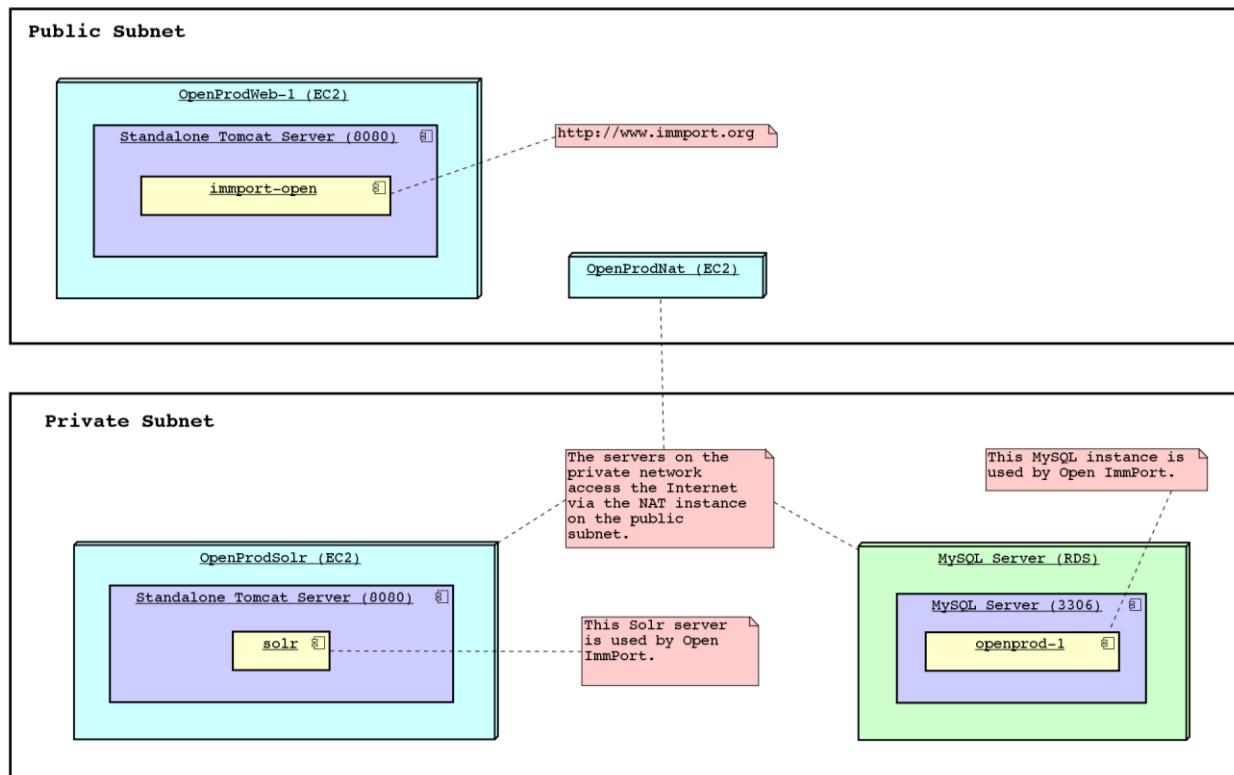
The table and column documentation for the Public ImmPort Schema are available [here](#).

Data Model Viewer is available [here](#).

Entity-Relationship diagrams are available [here](#).

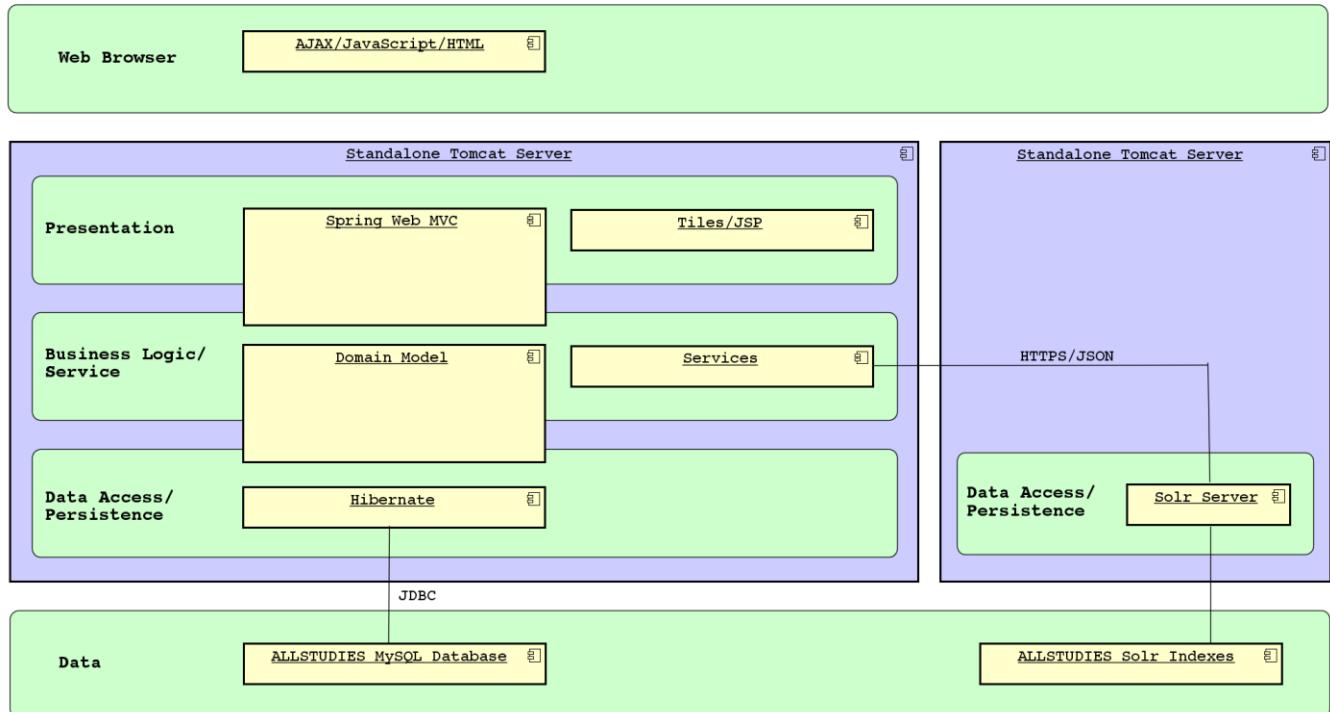
## 2.1.4 Open ImmPort Server Architecture

On AWS, Open ImmPort is separated into its own private network, or VPC. The outward facing Tomcat web server is contained in a public subnet where the site [www.immport.org](http://www.immport.org) is directed, while the data servers are contained in a private subnet only available through the NAT EC2 server within the public subnet. This separation limits exposure to security threats in the data tier of Open ImmPort. The configuration follows the AWS recommendations described at [http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC\\_Scenario2.html](http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/VPC_Scenario2.html)



## 2.1.5 Open ImmPort Software Architecture

Open ImmPort is a standard Java Spring based web application. The client layer utilizes JavaScript AJAX frameworks such as JQuery and Angular to provide interactive graphical user interfaces. The Model, View and Controller uses Spring Web MVC with Tiles/JSP to layout the web pages. Hibernate provides the data and persistence layer to the MySQL relational database via JDBC. Queries against SOLR are run through the web service connecting to the SOLR Tomcat server.



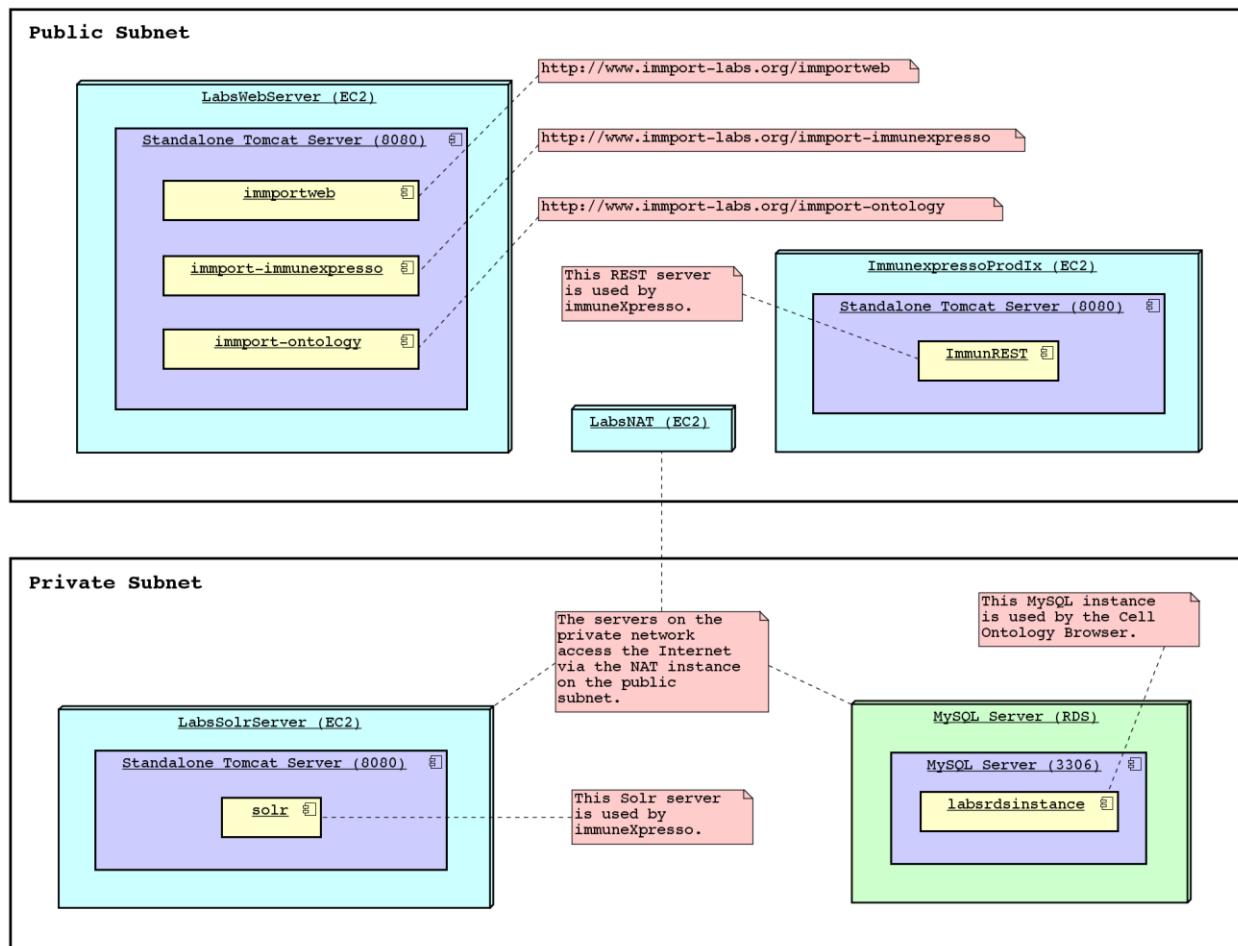
## 2.2 BETA AND DEVELOPMENT APPLICATIONS

### 2.2.1

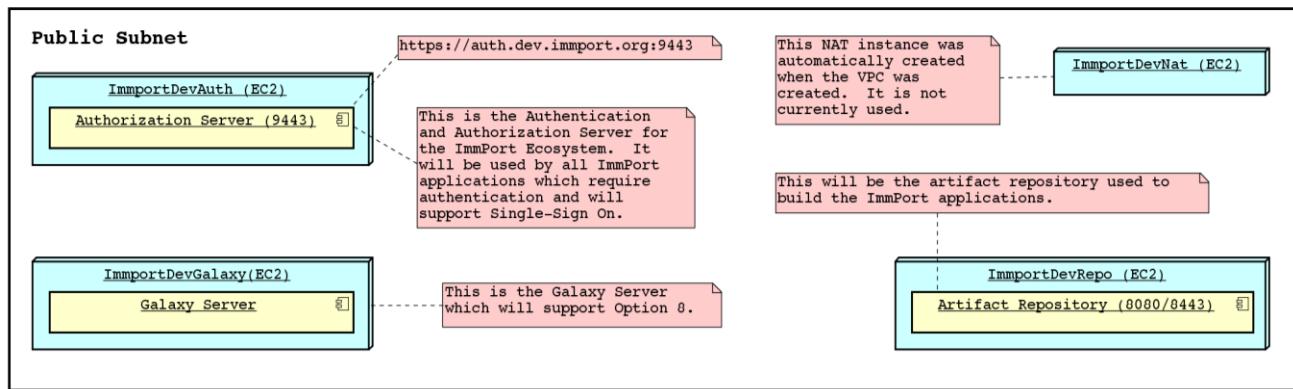
In addition to the production application environment for Open ImmPort, the BISC project also provides environments for applications that are in “Beta”, defined here as being reviewed by the user community, or in active development to be released in the future for review. These applications in Beta or development have been loosely grouped in what we call “ImmPort Labs”.

### 2.2.2 Server Architecture

The server architecture for applications in ImmPort Labs is described below. The VPC for the Beta applications is described below, and follows the same design pattern as noted for Open ImmPort.



The server architecture for applications that are in development is described below. These applications do not have a data component, and as a result only have a public subnet. Where required, access is limited by IP address filtering.



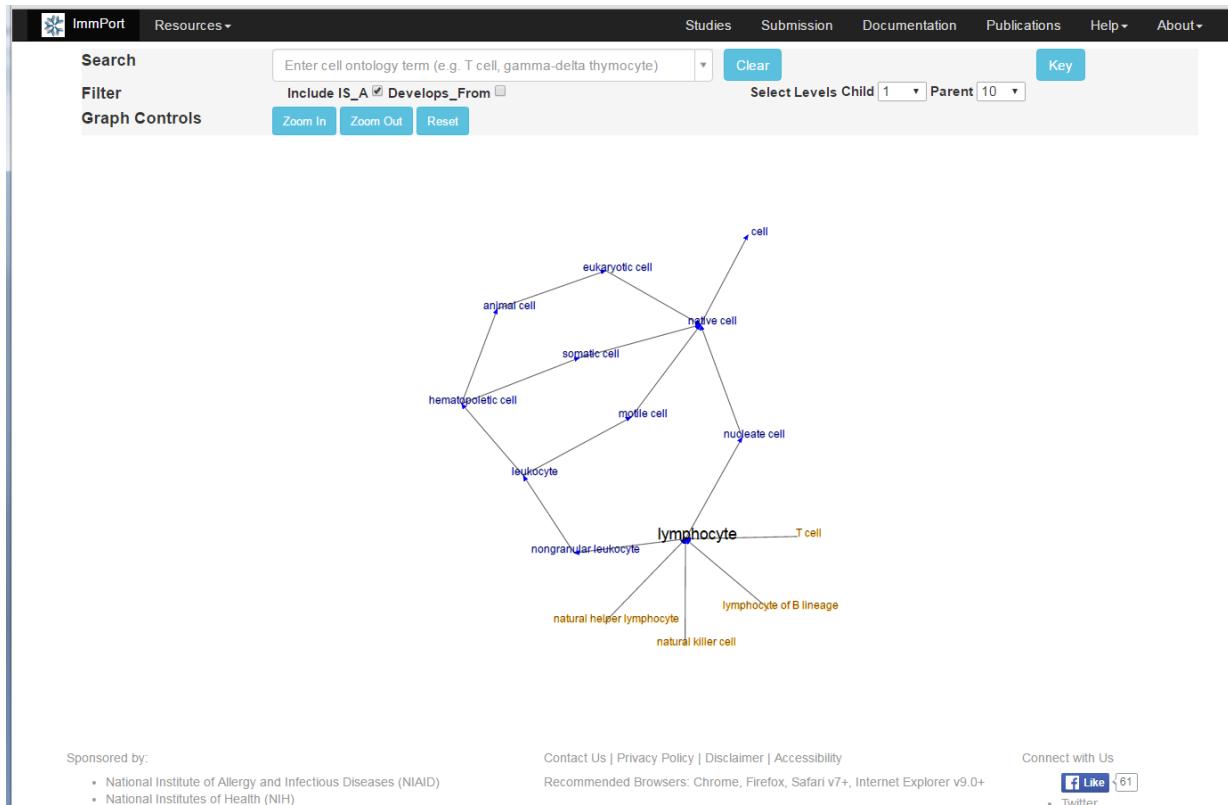
## 2.2.3 ImmuneXpresso

The screenshot shows the ImmuneXpresso application interface. At the top, there is a search bar labeled "Search immuneXpresso" and a "Got Feedback?" button. Below the search bar, there are sections for "Specify your filter term" (with a "Search Term" input field containing "CD4, T cell") and "Filters" (with options for Cell, Cytokine, Disease, Drug, Tissue, Interaction Type, and Article). The main content area displays a table of interactions, with the first few rows shown below:

Cell	Interaction	Cytokine	E-score	Paper #	Articles (Function / % Total)
CD4-positive, alpha-beta T cell	+ ➡	IFNG	4.68	194	<a href="#">See Articles</a>
CD4-positive, alpha-beta T cell	+ ➡	IL2	3.11	100	<a href="#">See Articles</a>
T-helper 1 cell	+ ➡	IFNG	3.99	75	<a href="#">See Articles</a>
T-helper 2 cell	+ ➡	IL4	12.27	52	<a href="#">See Articles</a>
CD4-positive, alpha-beta T cell	+ ➡	IL4	3.52	51	<a href="#">See Articles</a>
T-helper 1 cell	+ ➡	IL12	10.54	51	Differentiation / 17.31% Commitment / 1.92% Stimulation / 1.92% Polarization / 1.92% Other / 75.92% All / 100%
CD4-positive, CD25-positive, alpha-beta regulatory T cell	+ ➡	IL2	3.62	45	<a href="#">See Articles</a>

ImmuneXpresso is a Beta application built under the BISC Contract Option 7 and was built in collaboration with the team at the Shen-Orr lab at Technion (<http://shenorrlab.technion.ac.il/>). ImmuneXpresso continues the work of the lab in mining PubMed abstracts to determine relationships between cells and cytokines. The index is stored on a standalone Tomcat EC2 and accessible via a RESTful API. The front-end technology follows the same design pattern as Open ImmPort utilizing SOLR to query for matches to the Cell Ontology and Cytokine Registry available in ImmPort to build query terms to send to the ImmuneXpresso API. This project is in review for utility to the user community.

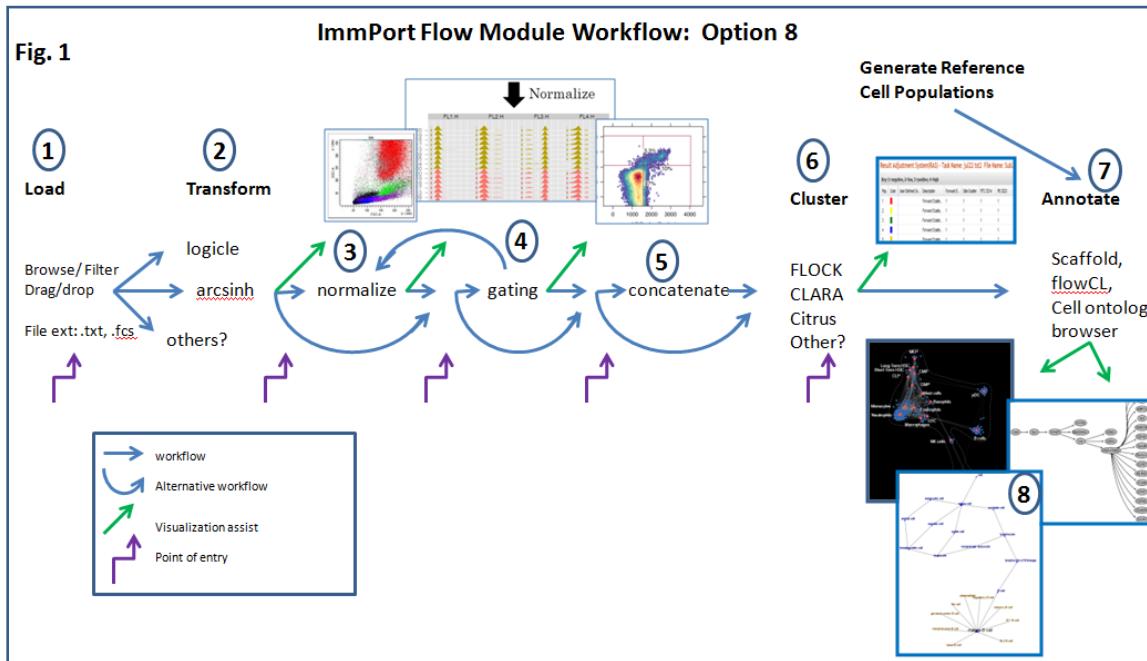
## 2.2.4 Ontology Browser



The ontology browser is another Beta application to support the visualization of the Cell Ontology (<http://obofoundry.org/ontology/cl.html>), which is of great utility in ImmPort data for standardization of cell populations. The browser provides a force directed graph visualization of the ontology, and utilizes the same software stack (JSP, AJAX, SOLR, MySQL) as Open ImmPort to allow for search for Cell Ontology terms with the addition of D3 for the visualization component. We are in the process of integrating the browser into the Open ImmPort application for visualization cell populations identified in ImmPort results as well as putting other ontologies into the browser to evaluate for utility.

## 2.2.5 Flow Analysis Tools (in development)

For BISC Option 8, our team was funded to perform a refresh on the Flow Cytometry Analysis tools in ImmPort largely based on the FLOCK algorithm. In addition, the science team from UCSF and Stanford proposed an expansion and modularization of the flow analysis tools, as reflected in the diagram below:



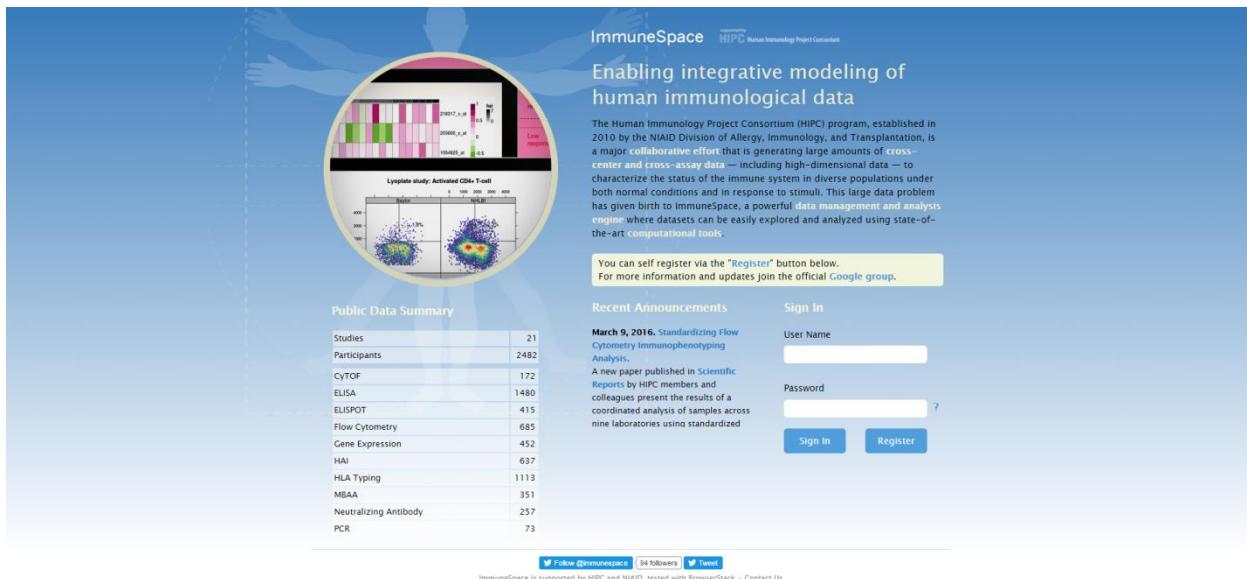
Further details about the methods being integrated and the steps of the process will be documented as the tool is developed. For the initial development work, we have chosen to utilize the popular analysis workflow engine Galaxy (<https://galaxyproject.org/>) to optimize modular method and component development and eventual sharing. The use of a publicly available web analysis framework was chosen over direct replacement of existing code in ImmPort because of the existence of open source tools that largely perform the workflow capabilities of queuing, bursting, and chaining methods in a generic way and the large cost to the ImmPort development team of development and maintaining code to perform those workflow capabilities.

Galaxy is an open web based platform that allows development of workflows and pipelines and facilitates reproducible research. We are implementing a version currently on AWS to provide existing users of FLOCK and the flow analysis tools the ability to review the new flow analysis tools for ease of use and utility. The individual methods and pipelines will be shared to the larger community through the appropriate venue, be it Bioconductor for R methods, the public Galaxy server, and/or in public servers of workflow engines such as GenePattern.

In terms of methods, the biggest upgrade in the pipeline will be the integration of the Single-Cell Analysis by Fixed Force- and Landmark-Directed (Scaffold) visualization method developed by the Nolan lab (<http://web.stanford.edu/group/nolan/>) partially under BISC funding and published in Science in 2015 (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4537647/>).

## 2.3 HOSTED APPLICATIONS: IMMUNESPACE

### 2.3.1



ImmuneSpace, available at [www.immunespace.org](http://www.immunespace.org) was developed by the team at the Gottardo lab (<http://www.rglab.org>) at Fred Hutchinson Cancer Center with the team at Labkey Software ([www.labkey.com](http://www.labkey.com)) under funding of the Human Immunology Project Consortium ([www.immuneprofiling.org](http://www.immuneprofiling.org)). Details about the project are available at the ImmuneSpace site, and the architecture in the Labkey product pages. The BISC team provides hosting and basic IT services on AWS for ImmuneSpace in the AWS instances funding by NIAID/DAIT.

It is in discussion that other applications funded by NIAID/DAIT may also be hosted in AWS instances managed by the BISC team in order to make use of the existing team's experience with AWS and managing federal systems in the cloud.

## 3.0 IMMPORT TOOLS AT NIAID

The NIAID hosting facility has been the home of the production instances of all parts of the ImmPort application since 2011 prior to the AWS applications development. Currently we are in a transition phase where a fraction of the applications are still based on ImmPort 2.0 technologies while the remainder has been transitioned to ImmPort 3.0. It is anticipated that at completion of the software refresh, the majority of the functionality hosted at the NIAID facility will be centered around private research and clinical data submission, QC, and curation before sharing that content to AWS for search and distribution on the cloud. It is also anticipated that the flow cytometry analysis tools will be deployed to AWS to ultimately make use of the burstable cloud capabilities for computation.

### 3.1 FEATURES TABLE

Extensive functionality has been built into the ImmPort web application during the duration of the contract period. **Table 3-1-1** summarizes the capabilities and features that have been developed to date.

**Table 3.1-1: Summary of ImmPort Capabilities and Features**

Capabilities/Features	Capability/Feature Description
Manage ImmPort User	System administration capabilities, which include allowing users to request system access (register), approve registration requests, create user accounts, update user information, query user information, and deactivate users.
Log In/Off	Authentication and authorization capabilities, which include allowing users to login, logoff, and retrieve login information when the account or password is forgotten.
Manage NIAID/DAIT Programs	Program management capabilities, which include creating, deleting, modifying, searching, and viewing programs and associating NIAID/DAIT-funded contracts and grants with DAIT programs.
Manage NIAID/DAIT Contracts and Grants	Contract/grant management capabilities, which include creating, searching, deleting, modifying, and viewing contracts and grants and assigning a PI for a contract or grant
Manage Research Project (RP) /Private Project Workspace (PPW)	Manage Research Projects (RP), which include allowing a user designated as a PI or PM on a contract or grant to create a project and its associated RP, and update the information associated with the project, manage user access to the RP.
Manage Collaborative Project (CP)	Manage the CP capabilities, which include allowing a user designated as a PI or PM on a contract or grant to create a CP, update the information associated with the CP, control user access to the CP, and share datasets in a CP.
Load Experimental Data	Batch loading of experimental data into a RP for multiple types of relevant metadata to provide the minimum information for multiple experimental assay types.
Browse, Simple Search, Advanced Search, Download Experimental and Clinical Research Data	Experimental and Clinical research data querying, includes allowing users to set up highly advanced multiple attribute search criteria and retrieve details of the experimental data from either a RP or from the semi-public part of the ImmPort database. Query results can be saved in the users RP as lists of entities or results downloaded in multiple files formats. Result files themselves can also be downloaded real-time online, through a queued mechanism, and as a request to receive a physical data delivery.
Browse, Simple Search, Advanced Search, Export Reference Data	Reference data querying, includes allowing users to set up highly advanced multiple attribute search criteria. Query results can be saved in the users RP as lists of entities or results downloaded in multiple files formats.
Analysis Dataset Generation and Analysis Tools	Analysis or visualization of reference data and experimental research data from the RP, CP and SPW which includes visualization tools such as GBrowse (Genome Browser), JMOL and Analysis tools including the Flow Analysis Tools, GenePattern for gene expression analysis, TagSNP Analysis Tool, HLA Typing Ambiguity Reduction tools and IGAT. Also allows for the creation of analysis data sets by combining disparate research data including interactive interfaces for creating PED files for genotyping analysis and HLA typing results.

Capabilities/Features	Capability/Feature Description
Database Audit History and Archiving	Maintain a complete audit history of research data (including analysis toolset data created) that is both updated and deleted. The audit history is defined as the ability to capture "who", "what", and "when" of the data involved in a change or deletion to Research data contained in the ImmPort System. Additionally, audit and collect limited summary information with respect to auditing/tracking of user session activity on a limited number of database areas. The focus is to obtain summary information on system activity such as logins, information created and updated in the several areas of the Administration Module, and usage of baseline Analysis Tools. Additionally, provide the capability to audit and track user session log information.
Share Research Data	Provide the capability to subset clinical and experimental metadata from RP areas to be shared into CP through an interactive user interface.
Browse and Download Shared Research Data	Provides the ability to navigate through a familiar file explorer styled interface to identify study results of interest, to select those result files, and then to download the result files using the Aspera transfer technology.

Development of the administrative functions of ImmPort, including User Management and Authentication, Program administration, Contract administration, and Project administration, are in the process of being refreshed into the ImmPort 3.0 codebase. Requirements for the features are captured in the ImmPort Administration Functional Requirements document that will be available on Open ImmPort by March 31, 2016.

Development of the Research Data Management capabilities in ImmPort for private data is ongoing. These requirements will be available on the ImmPort website within the coming two months as well.

The Flow Cytometry analysis tools are also in the process of being developed, and will be deployed for evaluation to the user community on AWS using the Galaxy framework to manage pipelines. Next steps for the flow cytometry tools will be based on this user feedback.

These functional requirements documents will provide an outline of the capabilities and general user interface descriptions for the given area.

## 3.2 HARDWARE INFRASTRUCTURE

The NIAID hosting facility is operated and, the encompassing hardware and networking systems and applications, administered by staff from the Office of Computational Infrastructure and Computational Biology (OCICB). The NIH network is operated by the CIT organization at the NIH level. Staff from NIAID or larger NIH operates all standard hosting activities, including OS administration of servers and network management,. Our team coordinates with NIAID staff in performing deployments, upgrading the OS on systems, upgrading database software, responding to security issues and alerts, scanning the ImmPort applications for security issues, storage planning, and overall issue resolution.

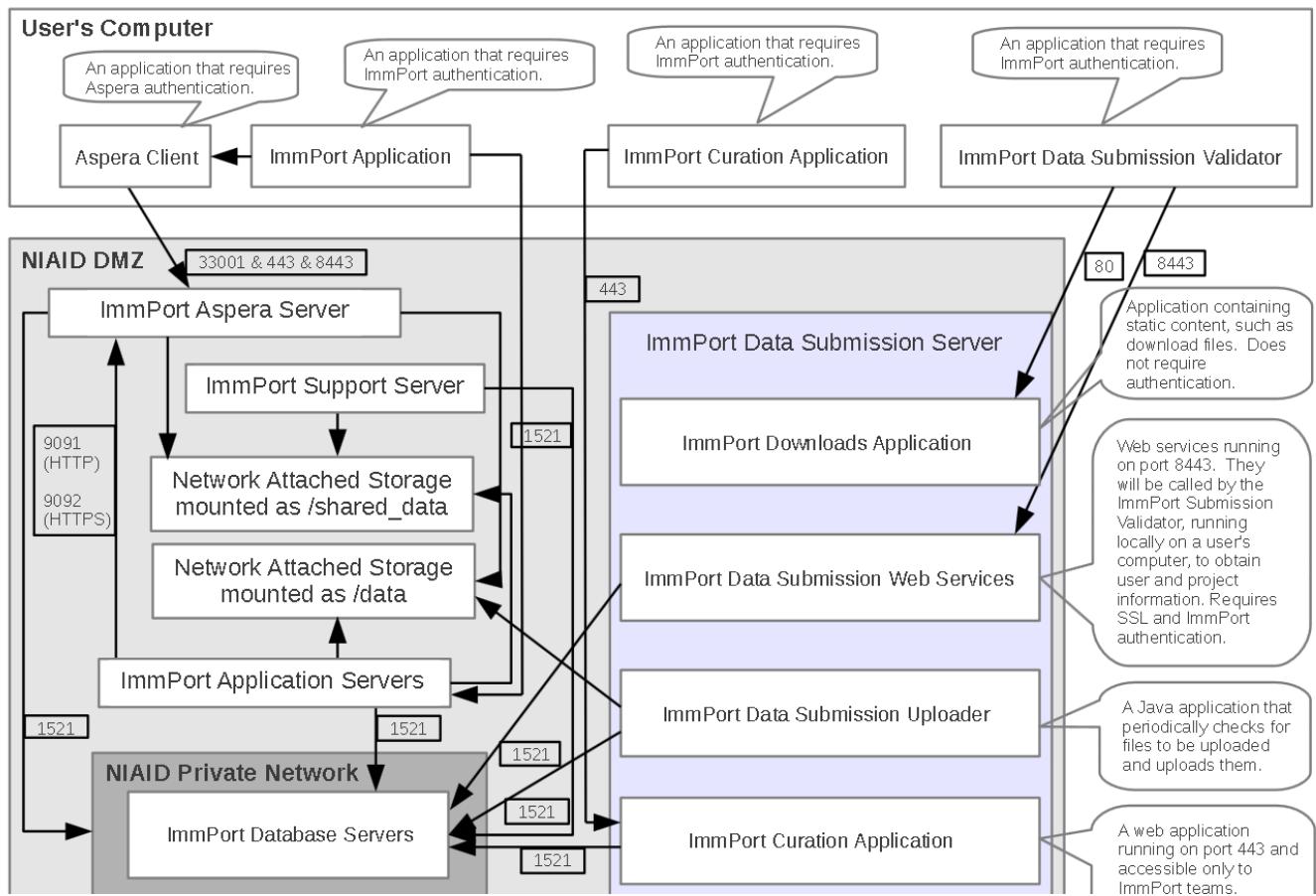
The overarching server architecture is illustrated in **Figure 3.2-1** below.

The pre-existing ImmPort 2.0 servers are identified as the ImmPort application servers (of which there are two) and the ImmPort Database servers (of which there are two). The Aspera Connect application is hosted on the Aspera physical server. The remainder of the ImmPort 3.0 developed code resides on the virtual ImmPort Data Submission Server.

All of the servers with the exception of the database servers reside in a NIAID DMZ to allow external access via defined ports. The database servers are further protected via limited access into a private network.

Specifications for the hardware are not available, but are based on the standard NIAID/OCICB family of servers. Firewall information is not available due to the proprietary nature of the NIH networks.

**Figure 3.2-1 Server Architecture at NIAID**



### 3.3 DATABASE ARCHITECTURE

The ImmPort system database architecture is stored and maintained in an Oracle 11g Enterprise Edition database utilizing Real Application Clusters (RAC), installed on a Linux EL5 operating system. Installed database options include Oracle Partitioning, On-Line Analytical Processing (OLAP), and Oracle Data Mining (ODM) options. The RAC environment provides necessary system load distribution and load balancing, while at the same time providing system redundancy and failover capabilities.

In addition to the BCPUB database instances deployed at the NIAID facility, an additional QA database instance is housed at that facility to perform installation testing. Specifications are identical to the production instance.

The data capture schema at NIAID for housing private data is being brought in line with naming schemes and relationships in the "shared" MySQL database available on Open ImmPort. A full data dictionary and ERD diagrams are available online as described below.

#### Schema Documentation

The table and column documentation for the Public ImmPort Schema are available [here](#).

Data Model Viewer is available [here](#).

Entity-Relationship diagrams are available [here](#).

### 3.4 ASPERA SERVER

We have integrated technology from Aspera, an IBM company, into the ImmPort system architecture. The Aspera Connect Server using the patented FASP® technology allows for optimized data transfer speeds across the Internet; we have utilized this technology for both data submission and data downloads for large files.

The Aspera security infrastructure provides user authentication and permissions on file systems. A free Aspera Connect Client provides a plug-in for users to install locally to take advantage of the FASP® UDP based transfer optimization. The Aspera SDK is utilized currently for accessing the Aspera Connect Server from the ImmPort application and data submission servers to queue up transfer tasks and return results to users.

More detailed documentation regarding the Aspera Connect Server, the Aspera SDK, FASP® technology, and the Aspera Connect Client can be found online at <http://asperasoft.com/software/transfer-servers/>

### 3.5 IMMPORT 3.0

The software architecture for ImmPort 3.0 applications largely mirrors that of Open ImmPort. Some aspects differ based on system requirements, but generally utilize Java, Spring, and AJAX components while making use of the Oracle 11 databases at NIAID.

#### 3.5.1 Hardware and Software Architecture components

The hardware and software components that differ or extend the Open ImmPort components previously described are listed below in **Table 3.5.1-1**.

**Table 3.5.1-1 Software and Hardware Architecture Components**

Component Name	
Network Attached Storage (NAS)	General networked storage providing access to diverse set of clients. Storage specifications are not available but are managed by NIAID OCICB.
Spring OXM	Spring's Object/XML Mapping support provides marshaling of objects to XML and un-marshalling in the reverse direction for object relational data mapping
Spring JDBC	Provided by the Spring framework as a value-added JDBC database connection that handles transactions, looping, connection opening and closing, and exception handling
Oracle RAC	Oracle Real Application Clusters (Oracle RAC) is a clustered version of Oracle Database based on a comprehensive high-availability stack that can be used as the foundation of a database cloud system as well as a shared infrastructure, ensuring high availability, scalability, and agility for any application. ( <a href="http://www.oracle.com">www.oracle.com</a> )
HyperSQL Database (HSQLDB)	Java based database that performs both in-memory and disk-based tables, and is widely used in embedded applications without a large server footprint. ( <a href="http://hsqldb.org">hsqldb.org</a> )

### 3.5.2 Data Browser

The Data Browser application provides a file explorer interface to users who have identified a study of interest either through Open ImmPort or through knowledge of the study accession. The user can navigate between studies or through directories of protocols, result files of various types, and study data files and select from those directories studies or directories to download. After selection and the click of the Download button, a request is sent to the Aspera server for download of specific files, and the Aspera Connect client coordinates download of the selected items to the user computer. The application requires installation of the Aspera Connect Client into the user's web browser to download content, but not to view files and directories.

The screenshot shows a file explorer interface for study SDY1. At the top right, it says "Logged in as: jwiser" and has "Logout" and a user icon. Below that, it says "Browse Shared Data > SDY1". The main area shows a table of files:

	Name	Size	Last Modified
▼	Protocols (5 files)	1.3 MB	Feb 9, 2016 12:17 PM
▼	ResultFiles (11,216 files)	3.9 GB	Jun 12, 2015 8:57 AM
▼	SDY1-DR17_SQL.zip	16.3 MB	Feb 9, 2016 7:55 AM
▼	SDY1-DR17_Subject_2_Flow_cytometry_result.txt	22.0 MB	Feb 9, 2016 7:55 AM
▼	SDY1-DR17_Tab.zip	14.7 MB	Feb 9, 2016 7:55 AM
▼	SDY1-DR17_manifest.txt	701.8 KB	Feb 9, 2016 7:55 AM
▼	SDY1-DR17_summary.txt	60.0 B	Feb 9, 2016 7:55 AM
▼	SDY1-DR17_table_count.txt	2.2 KB	Feb 9, 2016 7:55 AM
▼	StudyFiles (3 files)	630.9 KB	Jun 12, 2015 8:57 AM
▼	archive (44 files)	5.7 GB	Feb 9, 2016 8:45 AM

At the bottom left, there are navigation icons and a page size selector (100 items per page). At the bottom right, it says "1 - 10 of 10 items".

Sponsored by:

- National Institute of Allergy and Infectious Diseases (NIAID)
- National Institutes of Health (NIH)
- Health and Human Services (HHS)

[Contact Us](#) | [Privacy Policy](#) | [Disclaimer](#) | [Accessibility](#)

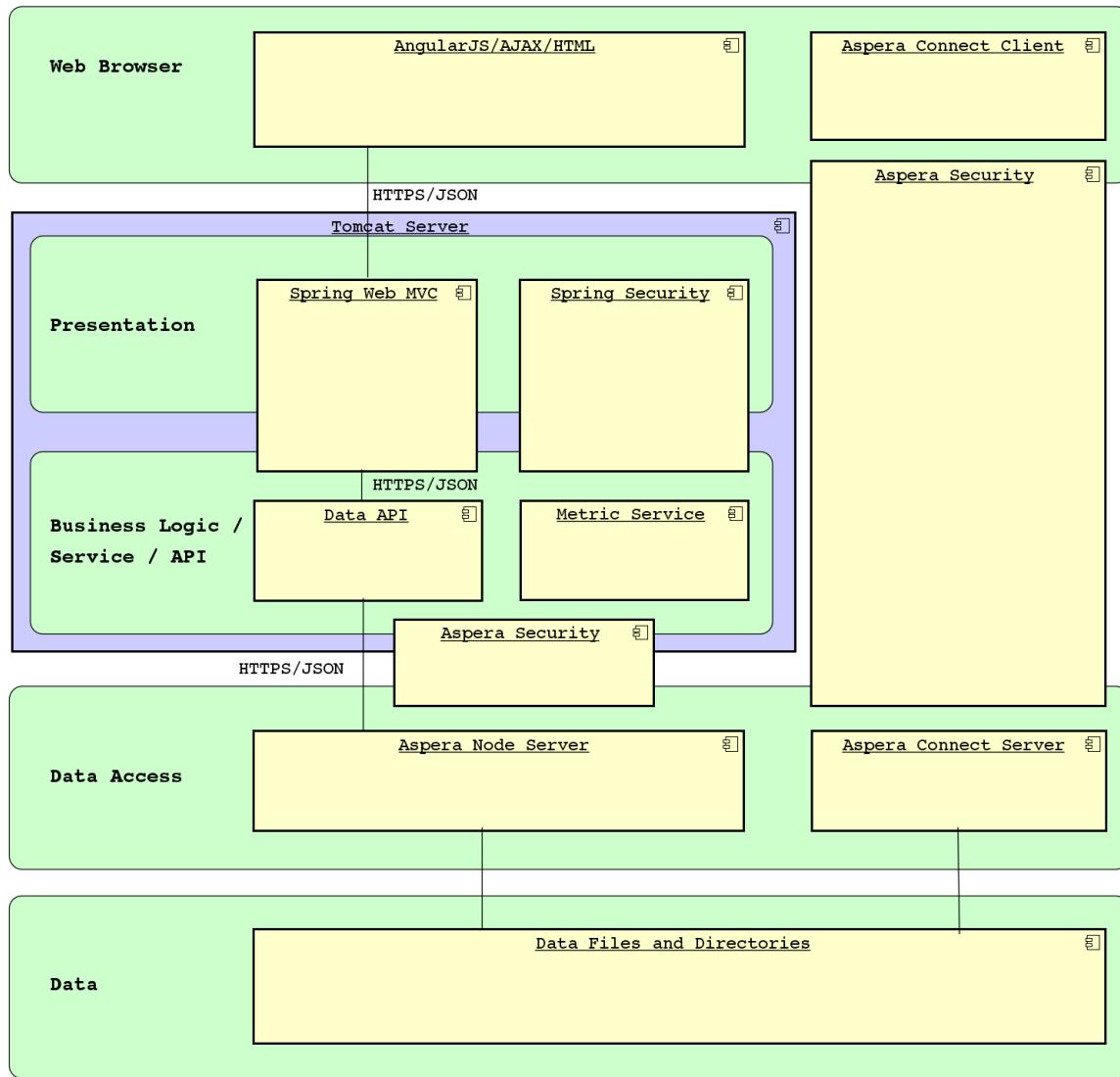
Recommended Browsers: Chrome, Firefox, Safari v7+, Internet Explorer v11+

Connect with Us

- [Like 61](#)
- [Twitter](#)
- [RSS Feed](#)
- [Google Group](#)

### 3.5.2.1 Data Browser Software Architecture

The Data Browser is a web-based application with use of similar web frameworks as Open ImmPort described previously. The interactive UI is handled by Angular JS components using AJAX and JavaScript, while the MVC architecture is implemented in Spring. Aspera Security is utilized when calls are made to download content on the data files and directories identified in the Aspera Node Server. The Aspera Connect Server governs the interactions with the data files and directories via the Aspera Connect Client.



### 3.5.3 Data Submission

Please check that you are using the [latest version](#) of the ImmPort data transfer templates.

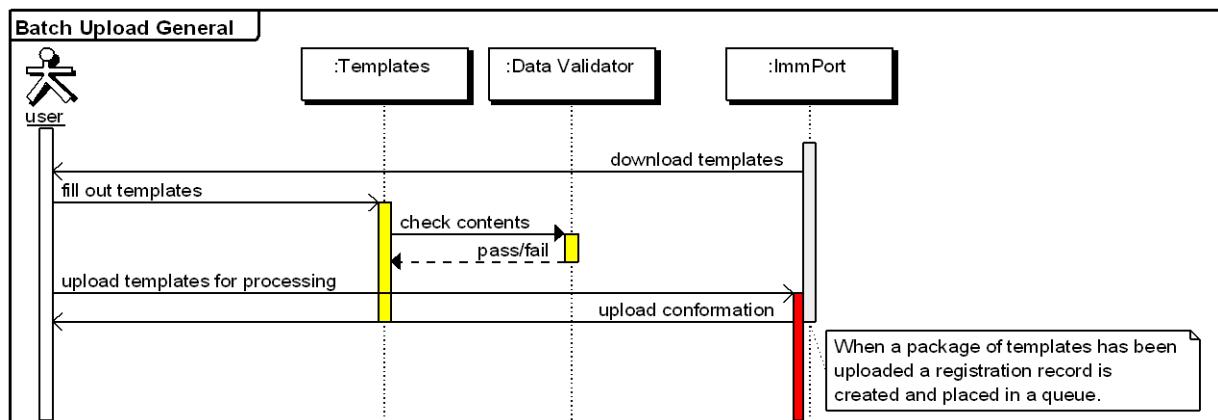
Research Project Title: TLR Innate Immunity - Obj 1/2

Upload Notes:

Upload Online     Upload By Aspera     Upload Offline

One of the primary goals of the BISC project is the sharing of research data at a level of detail that is beyond the current published literature dissemination mechanism. The ImmPort data upload module is designed to capture information regarding experimental results and the accompanying metadata, including subjects, biological samples, experiments details, etc. Uploaded information is access controlled by linking to a private research project and the user must be authorized to submit information for this project.

#### Data Submission Workflow Overview



The ImmPort system has developed a series of templates using Excel to define how information is to be submitted for each entity (<https://immport.org/immportWeb/display.do?content=DataTemplates>).

## Data Submission Templates

### Data Submission / Data Submission Home

Submit Data | Submission History | Resources ▾

Welcome to the ImmPort Experiment Data Submission Main Page!



The Experiment Data Submission module allows users to upload experiment data into the ImmPort DATABASE. Information about the experiment and its results are organized into several components including Subjects, Biological Samples, Experiment Samples, Protocols, Reagents, and Experiment Sample Results.

If this is your first time entering data, you will need the following:

- ▶ A set of protocol documents delineating your SOPs
- ▶ Reagent list
- ▶ A list of types of samples (blood vs lymphocytes vs serum, etc.)
- ▶ A system for uniquely defining each sample such that sample type, date of sample, type of analysis can be discerned

Please check that you are using the latest version of the ImmPort data transfer templates.

- ▶ [Data Submission Template Change History](#)
- ▶ [User Guide](#)
- ▶ [ImmPort Upload Templates Description](#)
- ▶ [Example Data Packages](#)
- ▶ [Data Package Validator](#)
  - ▶ This tool checks the format and content of the files in the data package and reports issues.
  - ▶ The tool is downloadable client side and executed from a graphical user interface.
- ▶ YouTube Video Tutorials:
  - ▶ [Introduction to ImmPort data loading, part 1 confirming access to your ImmPort project](#)
  - ▶ [Introduction to ImmPort data loading, part 2 basic study design and protocols templates](#)
  - ▶ [Introduction to ImmPort data loading, part 3 creating your zip package](#)
  - ▶ [Introduction to ImmPort data loading, part 4 installing the ImmPort validator on windows](#)
  - ▶ [Introduction to ImmPort data loading, part 5 validating your data submission package](#)
- ▶ [Minimum Clinical Information For Data Archives](#)
- ▶ [Subject De-identification Process](#)

Extensive documentation, including a user guide, example data packages, YouTube video links and full template specifications are available at the ImmPort site.

## Template Documentation

The data submission process is described in the User Guide ([https://immport-submission.niaid.nih.gov/downloads/documentation/ImmPort\\_Data\\_Submission\\_Guide.pdf](https://immport-submission.niaid.nih.gov/downloads/documentation/ImmPort_Data_Submission_Guide.pdf))

The ImmPort templates descriptions are available in interactive form [here](#).

The ImmPort templates are available to download and fill [here](#).

The ImmPort template's history is available [here](#).

The ImmPort templates descriptions in PDF form are available [here](#).

Data Model Viewer is available [here](#).

Excel templates are used to annotate columns and provide features such as lists of controlled vocabulary terms. The research data content is organized into generally understandable biological domains including

study, subjects, samples, protocols, reagents, and experiments. There is a template for each domain. The records in one domain often reference user defined records in other domains (e.g. subject records reference protocols for subject assessment and treatment). The data upload process supports a modular or incremental approach so that subsets of a study's content may be uploaded to ImmPort at the discretion of the data provider. Data uploaded in an earlier upload session may be referenced in a subsequent session without having to redefine the records. The data uploaded by users may be referenced by their own identifiers or by ImmPort accessions. User defined identifiers must be unique within a research project. ImmPort accessions are unique across the data repository. Each upload template has required fields that represent the minimum information that must be submitted (as defined by the NIAID DAIT minimum information guidelines- <https://immport.org/tutorials/MinimumInformationGuidelines.doc>), plus additional fields that are optional. In some cases, references from a record in one domain to a record in another domain are required and in other cases they are optional. This establishes links between uploaded data and ensures that the data model is effectively populated.

A commonly adopted practice is for the data provider to upload their data when their study is completed. The user fills out one or more templates. If the templates are in the Excel format, they saved from Excel into tab-delimited text files. This eliminates the need to support multiple versions of the Excel spreadsheets. These templates are bundled together into one ZIP archive with additional documents such as protocols and assay results. The ZIP archive is submitted to ImmPort via web based file upload technology. This generates an upload ticket that is used to track status and audit uploaded data. After the package is submitted, the system displays a confirmation screen and the package is queued for later processing.

Aspera transfer technology is again utilized as an option in data submission to optimize the speed of transfer. The user is required to have installed the Aspera Connect plug-in to perform a submission in this manner.

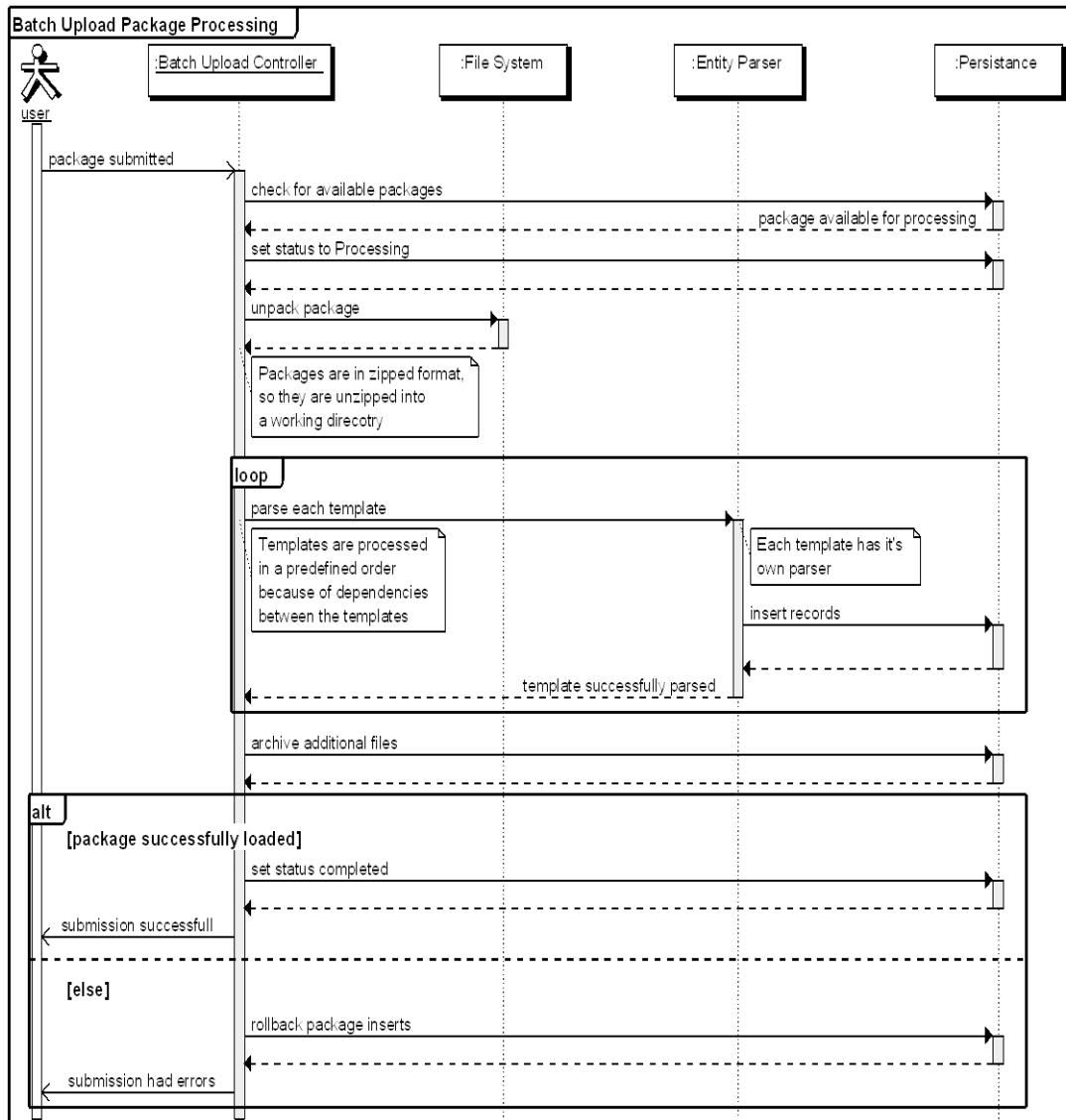
There are alternate procedures to deal with large data upload requirements that do not scale well with web-based technology. An upload ticket may be generated without uploading a ZIP archive. This ticket is used to link data transferred to the data upload staging area via non-web based means. For large data sets, an external drive may be sent to a data provider who uploads data to the drive and then sends it back to ImmPort. The data is transferred to the ImmPort repository and the standard data processing that parses and links the data to the content in the repository is completed. In some cases, a hybrid approach is taken whereby descriptive data is uploaded via the web interface and additional data is sent to ImmPort by offline methods.

#### **Data Submission Processing Workflow**

When a data package is uploaded, the system assigns a ticket number and records an entry in the upload registration table for this package. The batch upload daemon runs every 5 minutes and checks the database registration table looking for packages to process. When the upload daemon finds a package to process, it does the following:

1. Marks the package as processing
2. Unpacks the package in a working directory
3. Parses each template in predefined order, because there can be dependencies between the templates. For example, every biological sample record must link to a subject record, so the subject template, must be loaded before processing the biological sample template
4. As each template is parsed, the records are inserted into the ImmPort database.

5. Any file in the package that do not match the template names, is treated as a file to be archived and linked to the project
6. If the package is successfully loaded, the user is notified that the submission was completed successfully. If any error occurs during the load, the entire package is rolled back and the user is notified of the errors.



#### Data Submission History

The user may review the data submission history to their project(s) to get a status update and a log of their data submission sessions.

**Data Submission / Data Submission History**

Submit Data | Submission History | Resources +

Submit Data Main Page → Step 1: Download and Fill Templates → Step 2: Check Data in .zip file → Step 3: Send Data in .zip file → Step 4: Review Submission Status & Results

This page is a summary overview of the data submissions to a project. If you have access to more than one project, use the project filter to view the submission queue from other projects.

The data submitted to ImmPort can be queried and reviewed in [Research Data Search](#).

Filter By Project: Public Gene Expression Data ▾ Set Filter

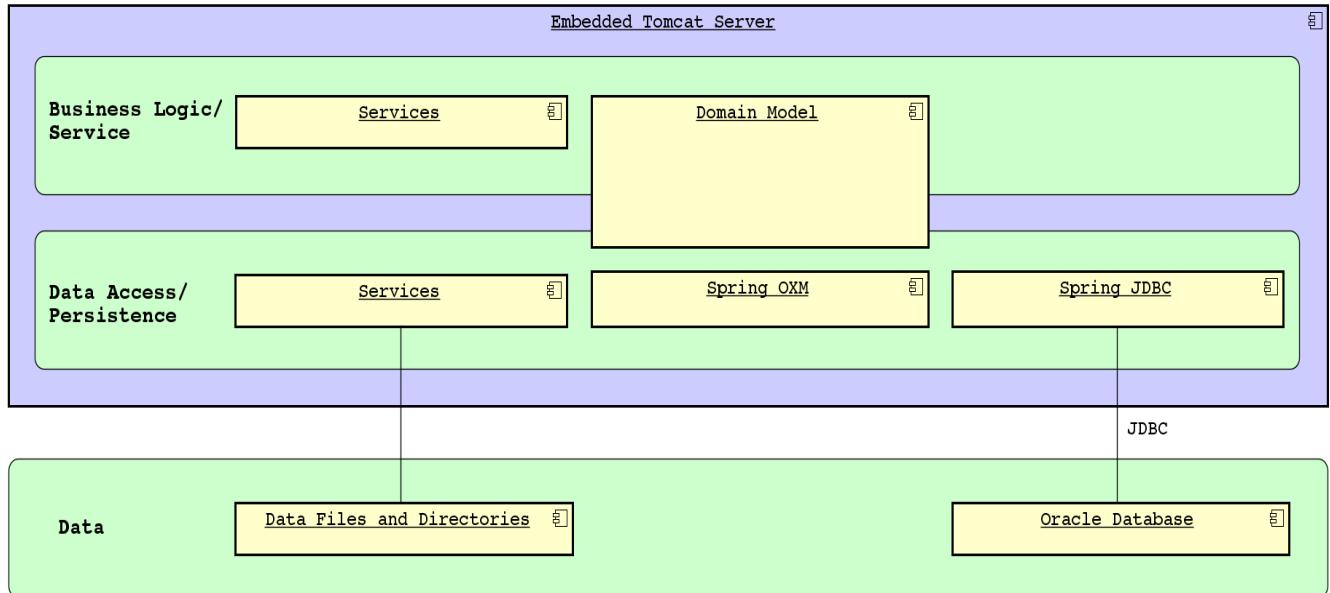
Click on a Ticket Number below to view a submission's details.

105 items found, displaying 1 to 25. [[First/Prev](#)] [1](#), [2](#), [3](#), [4](#), [5](#) [[Next/Last](#)]

Ticket Number	.ZIP File Name	Format	Status	Submitter	Submit Date
Bea_20090504_1166	HLA_1reagentperlocus_Example_Package.tab-delimited.zip	HLA_typing_summary	Completed	Bea	05/04/2009 10:51:28
Bea_20090504_1165	Illumina_BeadStudio_Final_Report_SubjectIDs.tab-delimited.zip	Affy_GCOS-GDAS-GTYPE_Genotyping_output	Completed	Bea	05/04/2009 10:50:46
Bea_20090504_1164	ELISPOT_example_Package.tab-delimited.zip	ELISPOT_report_file	Completed	Bea	05/04/2009 10:50:40
Bea_20081219_1084	Flow_Cytometry_example_Package.tab-delimited.zip	FCM_report_file	Completed	Bea	12/19/2008 17:04:26
Bea_20081125_1059	Genotyping_Small_Example_Package_111kb_XML.zip	Affy_GCOS-GDAS-GTYPE_Genotyping_output	Failed	Bea	11/25/2008 15:58:35

### 3.5.3.1 Data Uploader Software Architecture

The data submission “Uploader” does not consist of an interactive user interface since it relies on files loaded into a specified directory and a queue established in the database. The software architecture has been upgraded to the ImmPort 3.0 stack, and utilizes the Spring MVC, object relational mapping, and database connection frameworks.



### 3.5.4 Data Submission Validator

To assist the user with checking the validity of the templates they have filled out, ImmPort has developed a stand-alone Java application (DataValidator) that can be downloaded and run on the user's desktop before they upload the ZIP archive for processing. The DataValidator checks content for referential integrity between the templates, controlled vocabulary fields, previously assigned Project based identifiers, etc. Using the DataValidator is optional, but if used helps find most of the common data submission errors, before the package is submitted for final processing.

The screenshot shows the ImmPort Validator Application interface. At the top, there is a logo and links for 'Tools' and 'About'. Below the header, a welcome message says: 'Welcome to the ImmPort Validator Application. The ImmPort Data Submission Package Validator is used to check the format and content of the files in a data submission package ".zip" file before submitting to ImmPort.'

**Run Validation against ImmPort**

This section contains a green checkmark icon and instructions: 'Run Validation against ImmPort. This option requires you to be connected to the Internet and logged into ImmPort. Project data is automatically downloaded to your computer and used to validate a Data Submission Package File. A validation report will be displayed and can be downloaded. There is an option to save the Project Data to a local directory for use in Validation against local copy mode.'

1. Select the project you want to validate against.
2. Select the Data Submission Package File. (Please see the ImmPort Submission Data User Guide for more details) Check the "Do you want to save the Project Data File for later use?" if you want to save the Project Data file on your computer.
3. Click the Run Validation button.
4. Click on the Download Validation Report File link to save the report to your computer.

**Run Validation against local copy**

This section contains a computer monitor icon and instructions: 'Run Validation against local copy. Internet connection and logged into ImmPort is not required. A Project Data File is required and downloaded to your local directory. The validator application will validate the Data Submission Package File against the Project Data file. A validation report will be displayed and can be downloaded.'

1. Select the project data file that you want to validate against.
2. Select the Data Submission Package File. (Please see the ImmPort Submission Data User Guide for more details)
3. Click the Run Validation button.
4. Click on the Download Validation Report File link to save the report to your computer.

**View Report**

This section contains a document icon and instructions: 'View Report. This option will display the last report that was generated by the Validator Application.'

1. Click on the Download Validation Report File link to save the report to your computer.

**Download Project Data File**

This section contains a download icon and instructions: 'Download Project Data File. This option requires you to be connected to the Internet and logged into ImmPort. Project data will automatically be downloaded and a Project Data File link will be provided so you can save to your computer for use in the Run Validation against local copy mode.'

1. Select the Project to download data.
2. Click the Download Project Data File button.

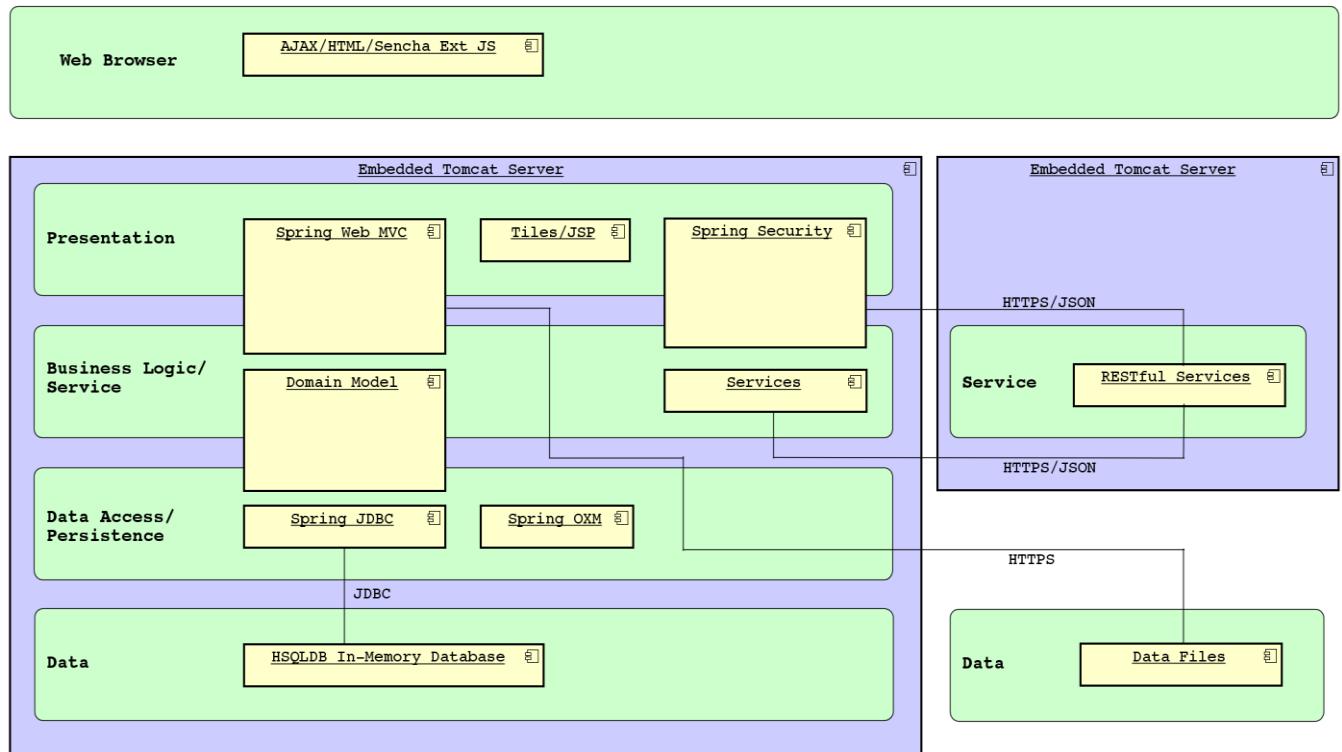
The Validator uses a HSQLDB database that is populated from content downloaded by the user via a RESTful web service connected to the ImmPort database. The HSQLDB database stores a subset of the content in the project necessary to validate referential integrity. The Validator synchronizes with the ImmPort repository automatically via the web service. After synchronization, the user selects the data upload archive to evaluate and the Validator provides a report on the success of the validation or issues it encountered.

### **3.5.4.1 Data Validator Software Architecture**

The Data Validator software has been upgraded to the ImmPort 3.0 stack. One overriding difference between the Data Validator and the rest of the ImmPort 3.0 applications is that the Data Validator is installed on a user's computer directly. This allows the user to be disconnected from the Internet and ImmPort to perform basic validation, and additionally, saves the effort of submitting large packages to ImmPort before knowing that there is an issue. The entire embedded Tomcat Server is installed in a package on the user machine via a multi-step installation wizard, which works for Windows, Mac, and Linux based computers. Having the entire web server installed on a remote site opens the possibility that users could install the Validator on a central location and having multiple users access the application.

Again, AJAX components are utilized for the interactive UI, and the Spring Web MVC handles the web tier model, view and controller. The Web Service is utilized for user authentication and retrieval of projects to which the user has access. Following authentication and project selection, the Web Service retrieves a subset of the submitted data in the ImmPort database (namely identifiers for each submitted object) that is used within a session of the Data Validator to validate local data packages. A copy of the data content retrieved from the web service is also stored locally, so a user may operate in off-line mode. The identifiers are stored within the session into the HSQLDB, which offers a high performance in memory database that opens at run-time and closes completely at session close.

Having the installed Data Validator in a similar web application framework as the rest of ImmPort 3.0 provides substantial benefits in code re-use and maintenance versus the prior iteration of the Data Validator that was implemented in Java Swing. Usage of HSQLDB for the validation steps, allows consistent use of SQL for running validation queries between the code executed in the Data Validator and the Data Submission codebase.



### 3.5.5 Curation Application

The Curation Application is a relatively simple web-based application utilized as an alpha tool by the internal scientists and data curators on the BISC team in Northrop Grumman. The MVC architecture is the same as the previously highlighted applications. The primary function of the Curation application as currently developed is to assist curators in executing and saving database SQSL queries that help with daily operational tasks in getting data collected, QC'ed, shared and curated. Based on usefulness to the internal curation team, it can be evaluated over time if the application would also be useful for external users.

### 3.5.6 Downloads Application

The Downloads application provides a means to host static content with no other purpose. No authentication is required. The goal is to move all static documentation and files used by the web applications in the NIAID facility into this application.

### 3.5.7 Support Server

The support server listed in the hardware diagram is a development server used for trying out computational and data generation tasks. There is no web application installed on this hardware and it is not available to the outside world other than through VPN access to the NIH network by badged employees.

## 3.6 IMMPORT 2.0

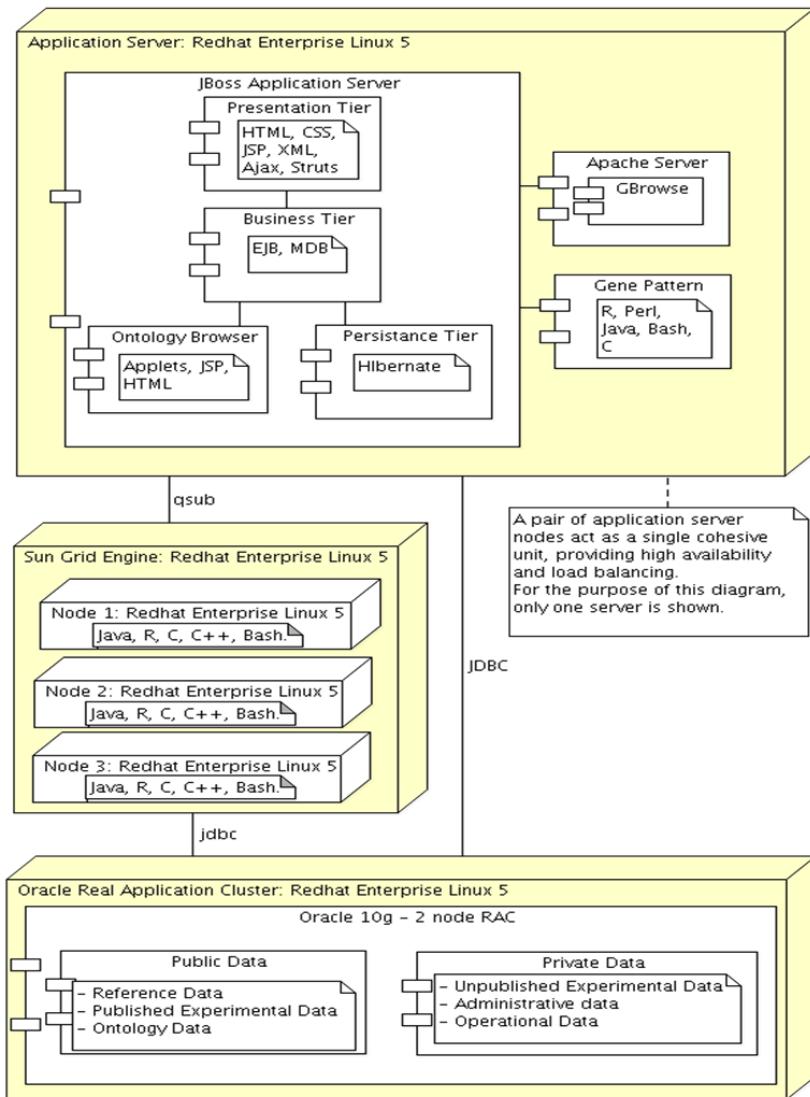
The ImmPort 2.0 system is implemented as an n-tier distributed architecture, providing a scalable, pluggable, and reusable componentized solution where additional layers or components can be added to accommodate current and future needs. The architecture is comprised of BigIP F5 hardware load balancers, a pair of application server nodes running the Redhat Enterprise Linux 5 OS, a 2-node Real Application Cluster (RAC) Oracle database, and a Network Attached Storage (NAS).

The software organization is implemented using a conventional 3-tier approach, which decouples software components into three major areas: the presentation, the business, and the persistence layers. Each layer specializes in specific orchestrated functions. The presentation layer separates the Graphical User Interface (GUI) from the complexity of the business rules. The persistence layer abstracts data retrieval and storage from yet another complex layer such as that of a Relational Database Management Systems (DBMS).

The presentation layer utilizes key web technologies, included but not limited to, the Hypertext Markup Language (HTML), Extensible Markup Language (XML), Cascading Style Sheets (CSS), Java Server Pages (JSP), Asynchronous JavaScript and XML (Ajax), Extjs, and the Struts Framework. The business tier leverages the Stateless Enterprise Java Beans (EJBs) and the Message Driven Beans (MDBs) stack of the Java2 Enterprise Edition (J2EE) framework. Using the Facade Object Oriented Design pattern, the EJBs abstracts the complexity of the underling complex business rules. Lastly, the persistence layer manages all database access and storage and it relies on the successful object-relational mapping (ORM) Hibernate library for the Java platform.

### 3.6.1.1 Software Architecture

The ImmPort system software architecture diagram shown below depicts all the components comprising the BISC software architecture that have not yet been upgraded and that are housed on the *ImmPort Application Servers* and the *Database Servers* previously described. Each software component is described in **Table 3.6.1-1**, which follows the figure.



**Figure 3.6.1-1: Software Architecture**

**Table 3.6.1-1. Software Architecture Components**

Component Name	Description
Application Server	The Application Server contains JBoss Application Server instance(s) and an Apache Web Server. The JBoss App Server is responsible for hosting the presentation tier, business logic tier, and the data access tier. For purposes of performance and scalability, multiple application server instances may be utilized to support running a cluster of JBoss application servers.
Business Tier	The business tier abstracts the complexity of the business rules and manages the communication between the presentation and database tiers.
Database Servers (2 nodes)	This software component contains two major database areas: Public Data Warehouse and Private Project Workspace. There is also a Raw File Repository for staging and backup/archive purposes. To ensure proper performance and scalability, there will be at least two of these servers running Oracle 10g in Real Application Cluster mode.
EJB Container	The application tier consists of the Enterprise Java Bean (EJB) container and the Java Messaging Service. The EJB container houses the majority of the application logic in terms of data loading/publishing, browsing/querying, journaling, and any administrative/management activities. The JMS component, on the other hand, is responsible for the execution of any application logic that is asynchronous in nature. This helps reduce the load and distribute the work across the network.
Hibernate	The data access tier utilizes the open source Hibernate library for data access. Hibernate provides fine-grained access/update, caching, and transactional persistence to the Private Project Workspace database where transactional support is required.
Java Messaging Service	Java Messaging Service (JMS) is the standard API for sending and receiving messages.
Java Server Pages	Java Server Pages (JSP) is server-side technology for building presentation tier components.
JBoss Application Server	JBoss Application Server is an open-source J2EE application server.
Oracle 11g Database	Oracle 11g is a relational database management system (RDBMS). This software component contains all subcomponents that make-up the server-side RDBMS software. This includes, but is not limited to: Cluster Ready Services (CRS) software, Real Application Clusters (RAC) software, and RAC database server software. The server configuration is a 2-node RAC environment with shared Automatic Storage Management (ASM) utilizing ASMLib software.
Oracle 11g Client	This software component contains all subcomponents and utilities that are installed as part of the Oracle 10g Client software, currently residing on the ImmPort Application Server. This SQL*Plus Client subcomponent is used by the GBrowse application software for establishing a remote client database connection for retrieval of data for the GBrowse Visualization Tool.
Persistence Tier	The persistence tier manages all database access and storage using the open source object relational mapping Hibernate library.
Presentation Tier	The presentation tier manages and renders Graphical User Interface (GUI) components using web technologies such as HTML, CSS, JSP, XML, Ajax, and the Struts framework.
Struts	Struts is an open-source Model View Controller (MVC) framework for developing J2EE web applications.

### 3.6.2 User Management and Authentication

The Manage User design artifacts model system administration capabilities, which include allowing users to request system access (register), approve registration requests, creating user accounts, updating user information, querying users, and deactivating users from the system.

#### **Create/View/Update/Search User**

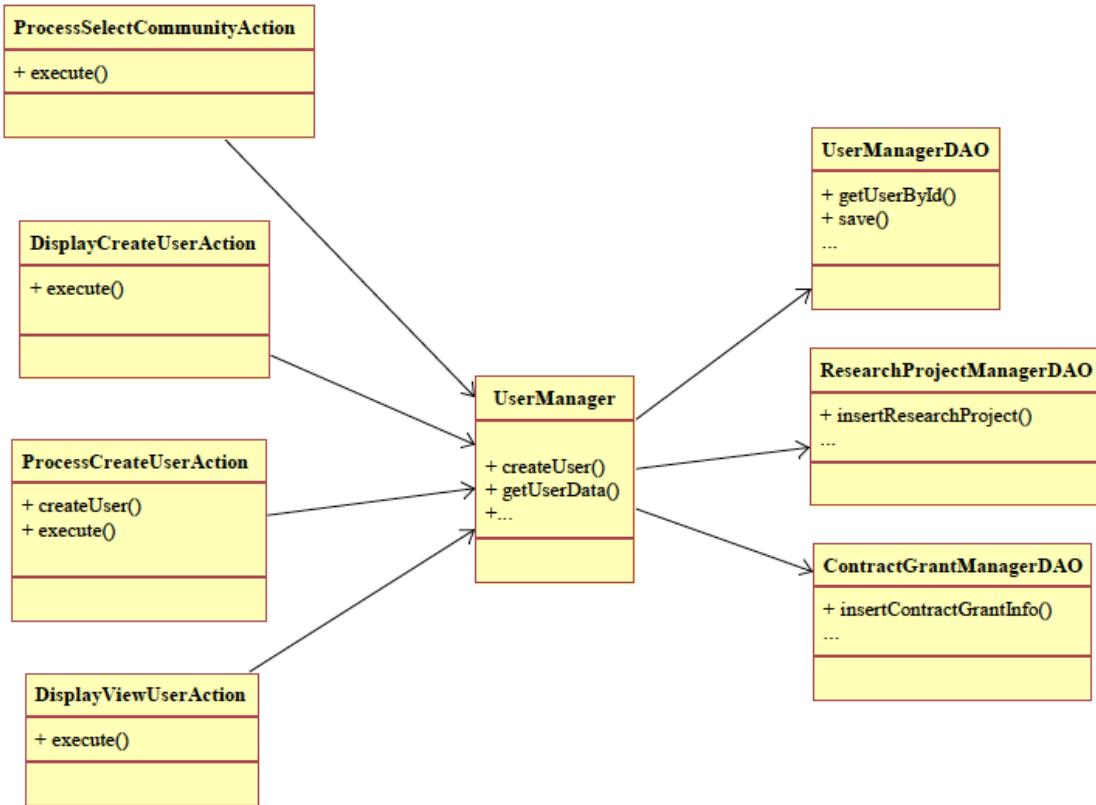
The Create/View/Update/Search User depicts the events for creating/viewing/updating/searching a user account. The system collects the user information and creates a user account, provided the submitted data are valid. If the system encounters errors, then it will display the appropriate message. Once the user information is collected the system will display the data while giving the user an option to update the information.

The screenshot shows a web-based application interface titled "Admin / Create User". At the top, there is a navigation bar with links: My Profile, Programs, Grants - Contracts, Users, and Reports. Below the navigation bar, the title "Create User: Life Science Research Community" is displayed, followed by a note that fields marked with an asterisk (\*) are required.

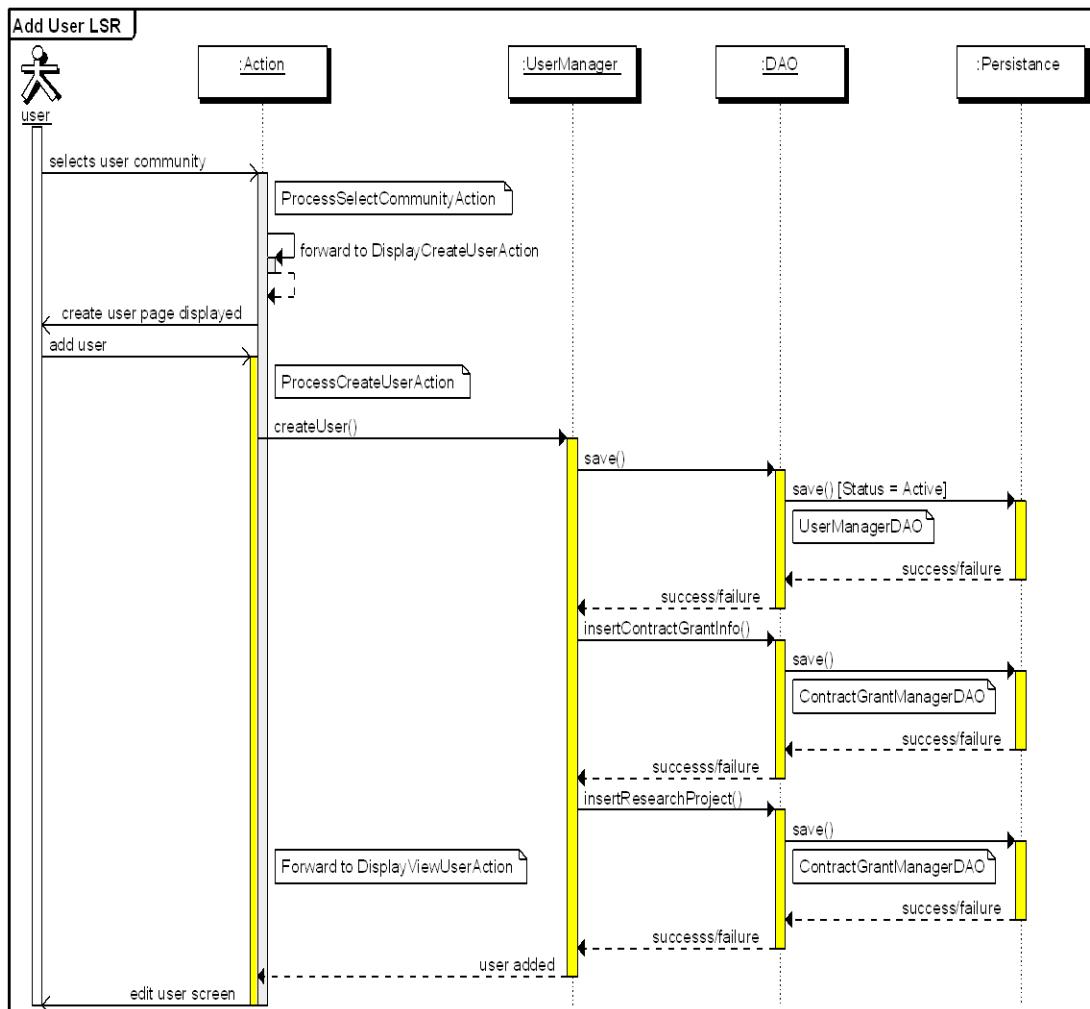
The form contains the following fields:

- ImmPort Login Name \* (up to 12 characters allowed): A text input field with a validation message: "Please enter a minimum of 8 characters for your password."
- Password \*: A text input field.
- Re-Enter Password \*: A text input field.
- Last Name \*: A text input field.
- First Name \*: A text input field.
- Middle Initial: A text input field.
- Organization \*: A text input field.
- Department: A text input field.
- Principal Investigator Name: A text input field.
- Phone Number (US/Canada xxx-xxx-xxxx) \*: A text input field next to radio buttons for "US and Canada" (selected) and "International".
- Email Address \*: A text input field.
- How can ImmPort assist your research efforts?: A text input field.
- How did you learn of ImmPort \*: A dropdown menu with options including "Colleague" (selected).

At the bottom of the form are two buttons: "Create" and "Cancel".

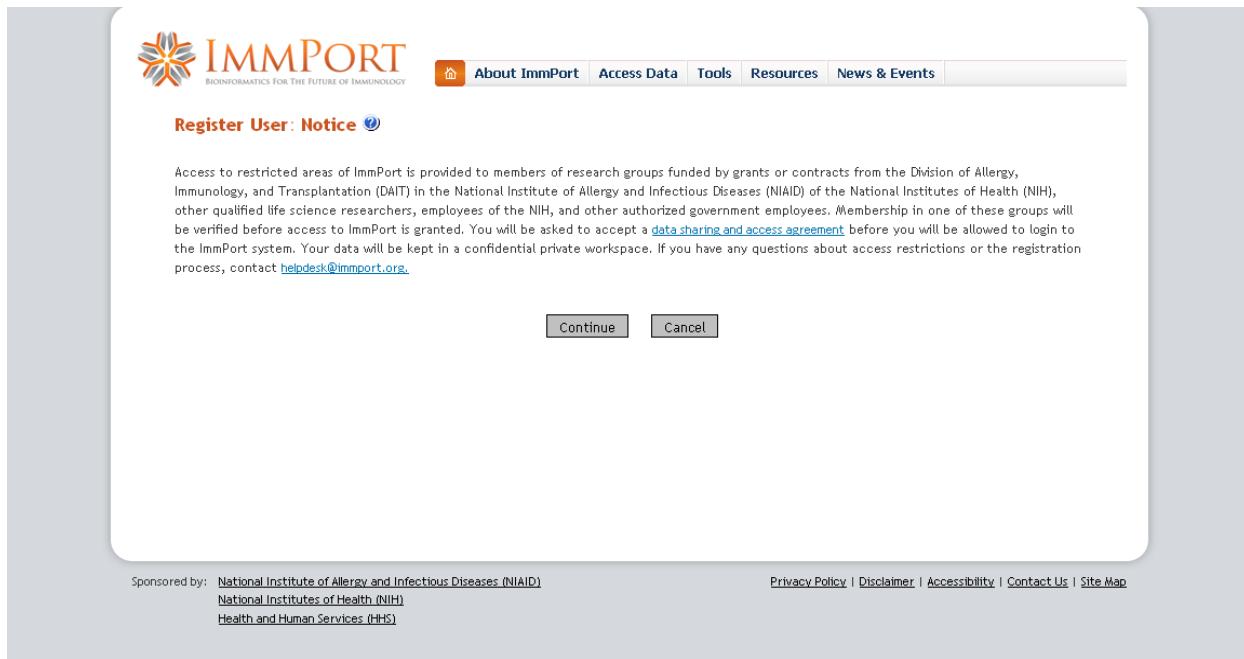
**Create New User Class Diagram**

**Create New User Sequence Diagram**

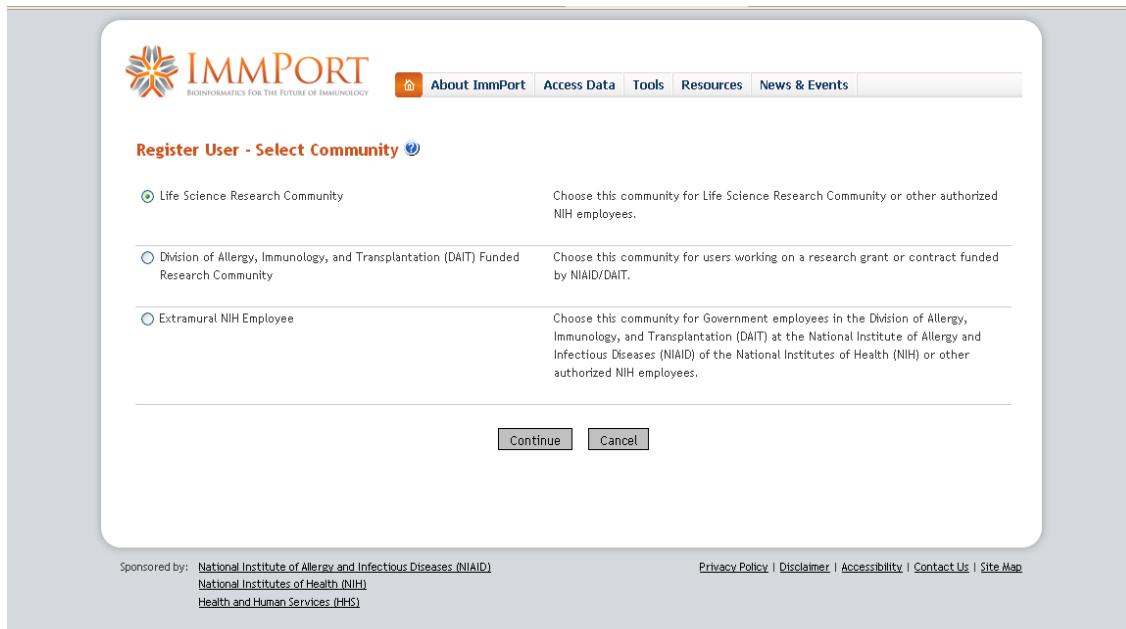


### 3.6.2.1 Users Registration

The User Registration depicts the events for requesting access to ImmPort. First, the individual initiates a registration request.



Depending upon the type of access requested, the system prompts the individual for additional information. The user may select to be a general member of the life sciences community, a member of the DAIT-funded research community, or an extramural NIH employee.



The system validates the user information and saves the registration request, provided the data submitted are valid.

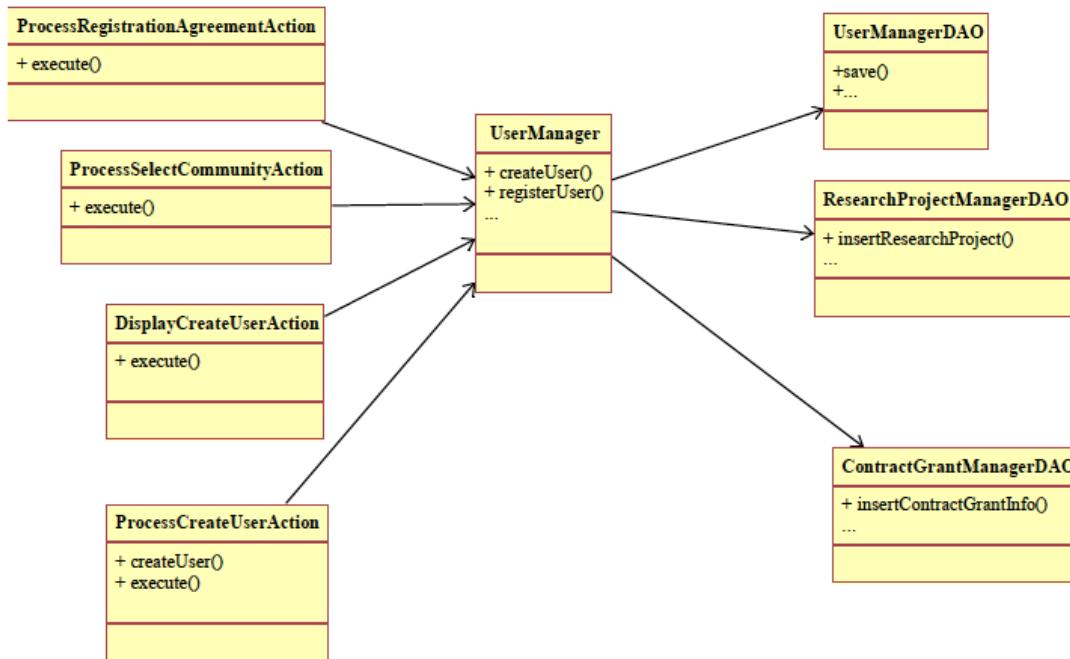
**Register User: Life Science Research Community** ⓘ

Fields marked with an asterisk \* are required.

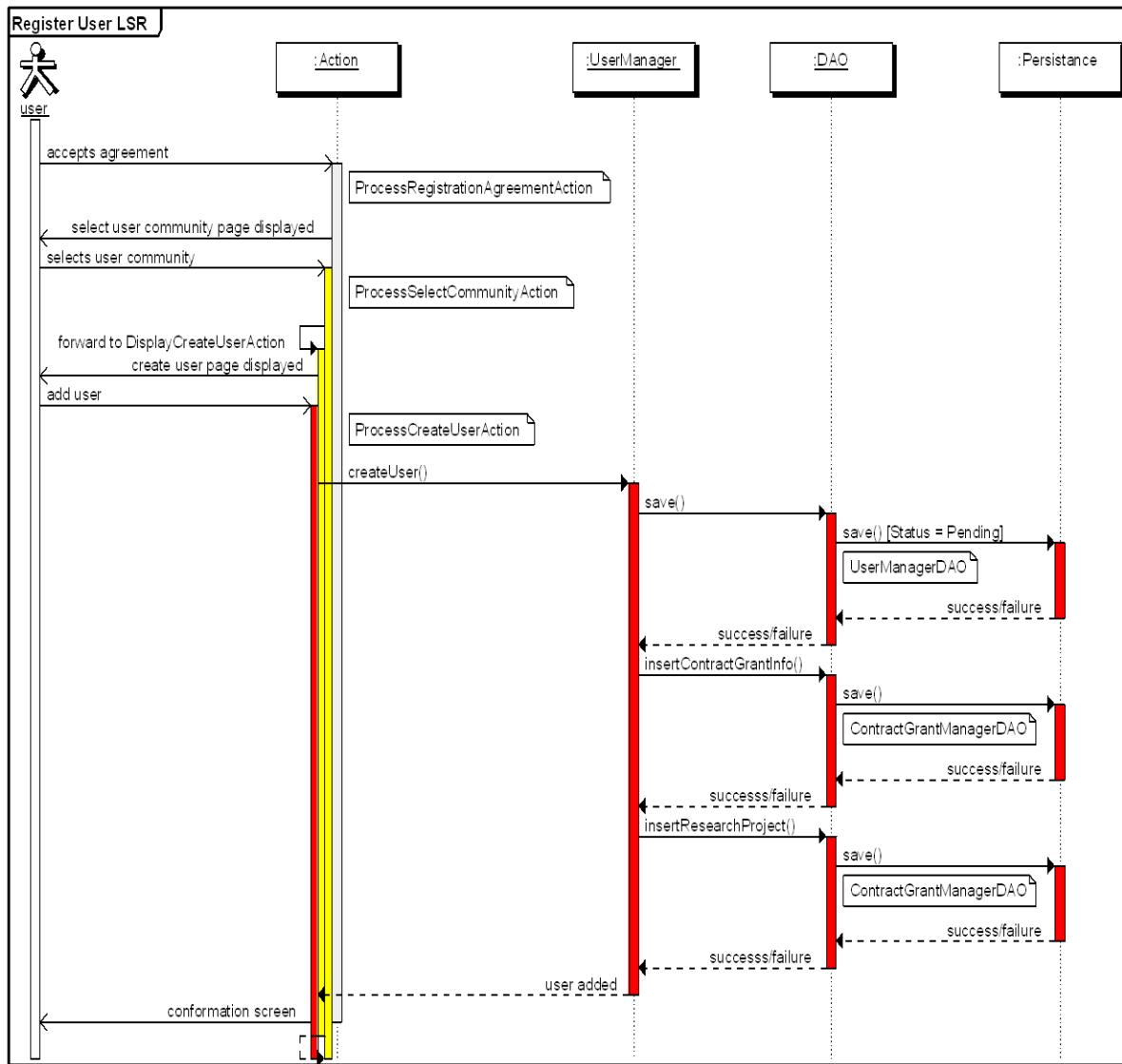
ImmPort Login Name *	<input type="text" value="johnd"/>
(up to 12 characters allowed)	
Password *	<input type="password" value="*****"/>
Please enter a minimum of 8 characters for your password.	
Re-Enter Password *	<input type="password" value="*****"/>
Last Name *	<input type="text" value="Doe"/>
First Name *	<input type="text" value="John"/>
Middle Initial	<input type="text"/>
Organization *	<input type="text" value="BISC"/>
Department	<input type="text"/>
Principal Investigator Name	<input type="text"/>
Phone Number (US/Canada xxx-xxx-xxxx) *	<input type="text" value="301-527-0000"/> <small>ImmPort recommends using an email address from an education or research institution to expedite registration approval.</small>
Email Address *	<input type="text" value="johndoe@research.org"/>
Password Retrieval Question *	<input type="text" value="What is your favorite color?"/>
Password Retrieval Answer *	<input type="text" value="blue"/>
How did you learn of ImmPort *	<input type="text" value="Colleague"/>
How can ImmPort assist your research efforts?	<input type="text" value="Research"/>

Appropriate error messages are displayed if errors are encountered. At the end of the process, the system notifies those users responsible for reviewing registration requests.

**User Registration Class Diagram**



**User Registration Sequence Diagram**



### 3.6.2.2 Approve/Reject User

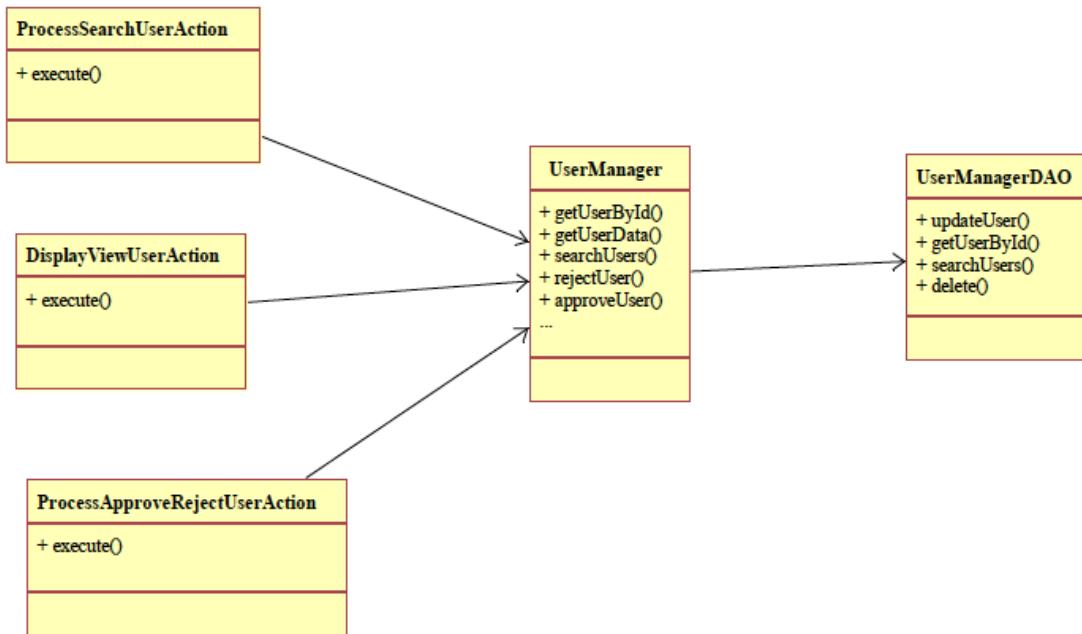
The Approve/Reject User depicts the events for approving/rejecting a user account. Appropriate error messages are displayed if errors are encountered.

The screenshot shows a web-based administrative interface titled "Admin / Review Registration Request". The page includes a navigation bar with links for "My Profile", "Programs", "Grants - Contracts", "Users", and "Reports". A note at the top states: "Fields marked with an asterisk \* are required." The form contains the following fields:

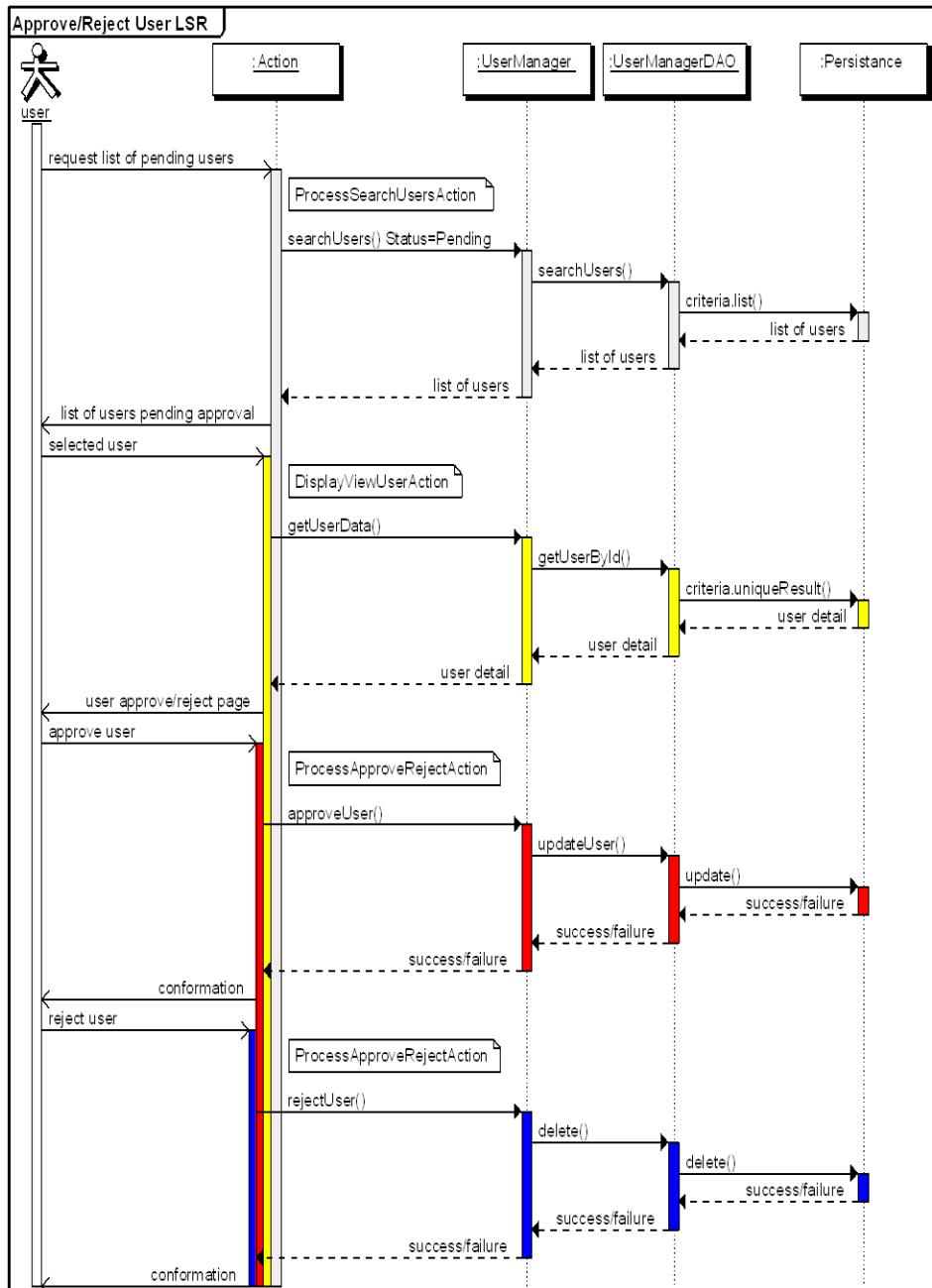
User ID	john
Last Name *	Doe
First Name *	John
Middle Initial	
Organization *	BISC
Department	
Principal Investigator Name	
Phone Number (US/Canada xxx-xxx-xxxx) *	<input checked="" type="radio"/> US and Canada <input type="radio"/> International 301-527-0000
ImmPort recommends using an email address from an education or research institution to expedite registration approval.	
Email Address *	john.doe@research.org
How did you learn of ImmPort?	Colleague
How can ImmPort assist your research efforts	Research
User Community*	Life Science Research Community

At the bottom of the form are three buttons: "Approve", "Reject", and "Cancel".

**Approve/Reject User Class Diagram**

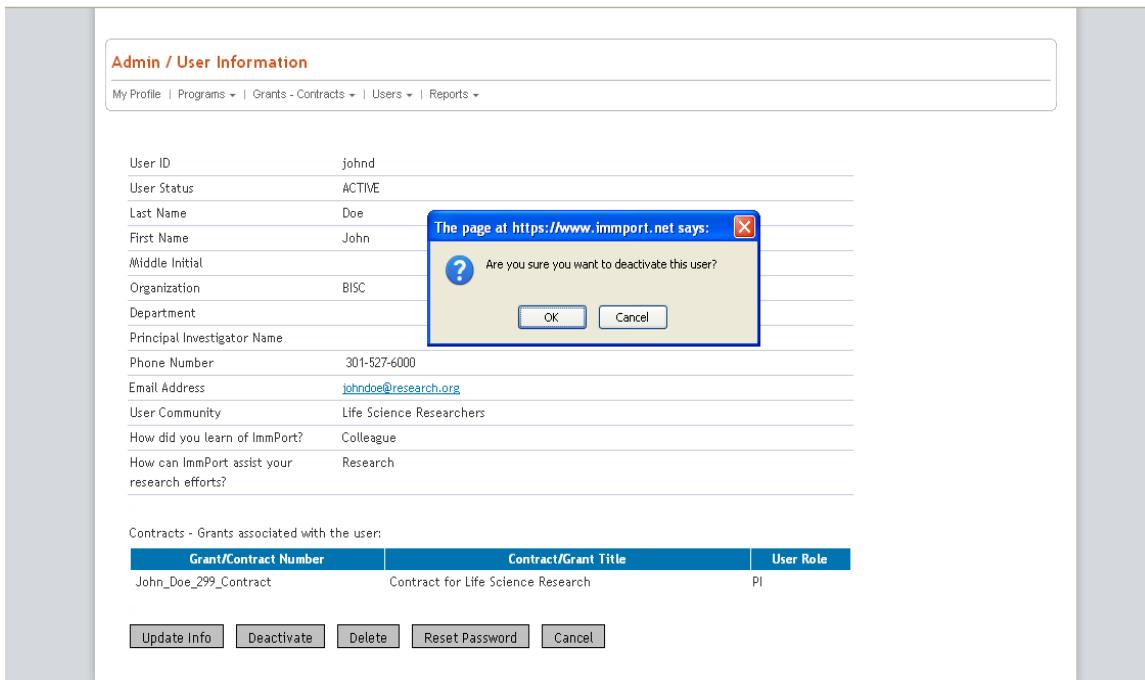


**Approve/Reject User Sequence Diagram**

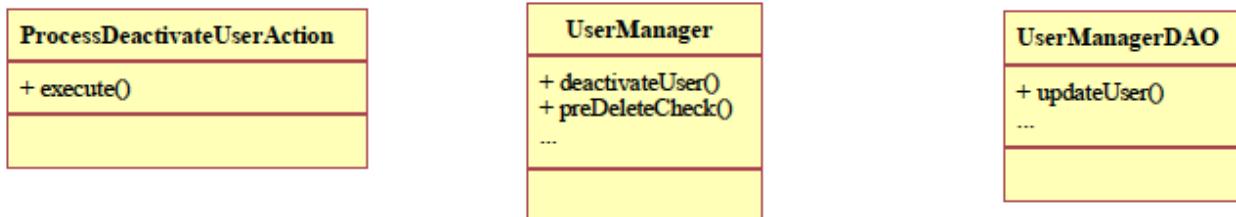


### 3.6.2.3 Deactivate User

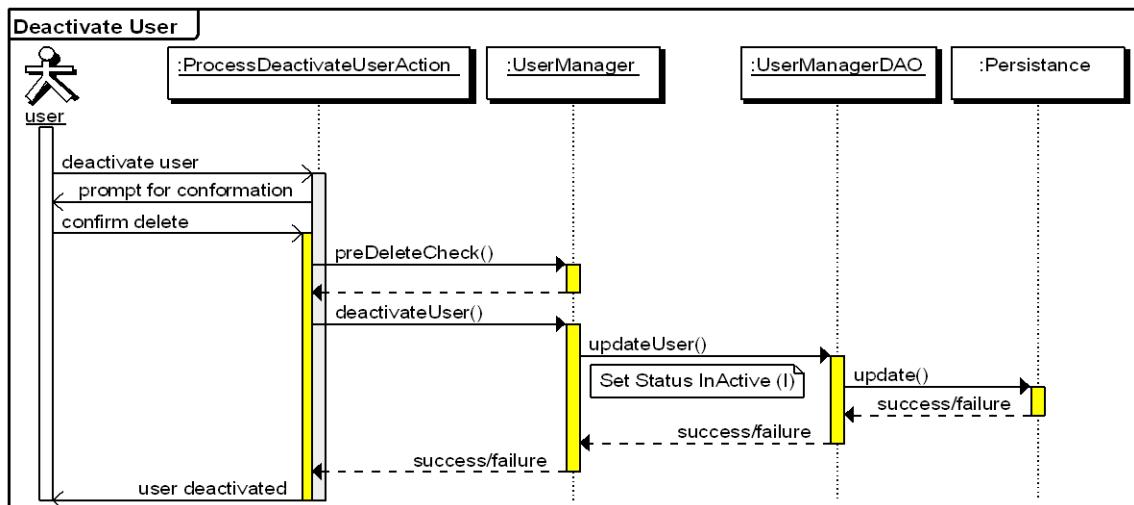
The Deactivate depicts the events for deactivating a user account. Appropriate error messages are displayed if errors are encountered.



### Deactivate User Class Diagram



### Deactivate User Sequence Diagram



#### 3.6.2.4 Activate User

The activate user depicts the events for activating a user account that has been deactivated in the past. Appropriate error messages are displayed if errors are encountered.

**Admin / User Information**

User was deactivated by patty-po on 2010-10-26 16:04:28.0

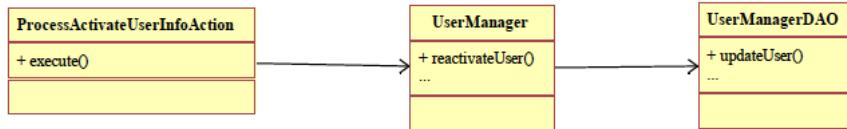
User ID	johnd
User Status	INACTIVE
Last Name	Doe
First Name	John
Middle Initial	
Organization	BISC
Department	
Principal Investigator Name	
Phone Number	301-527-6000
Email Address	<a href="mailto:johndoe@research.org">johndoe@research.org</a>
User Community	Life Science Researchers
How did you learn of ImmPort?	Colleague
How can ImmPort assist your research efforts?	Research

Contracts - Grants associated with the user:

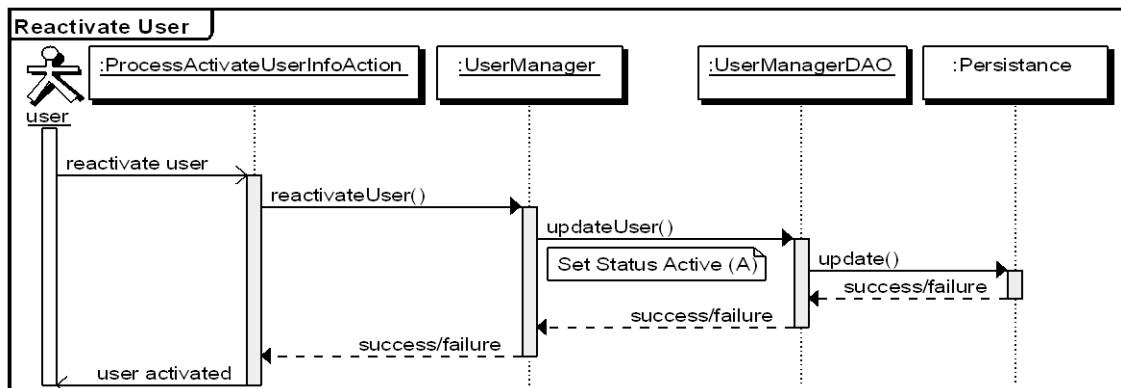
Grant/Contract Number	Contract/Grant Title	User Role
John_Doe_299_Contract	Contract for Life Science Research	PI

**Activate User**

**Activate User Class Diagram**



**Activate User Sequence Diagram**



### 3.6.2.5 Search Users

The search user depicts the events for searching a user account. The user is first presented with a search screen.

ImmPort Bioinformatics FOR THE FUTURE OF IMMUNOLOGY

[Edit Profile](#) | [Sign Out](#)

**Admin / Search Users**

My Profile | Programs | Grants - Contracts | Users | Reports

Enter criteria in the form below to search for a specific ImmPort user. If you submit the form without entering search criteria, you will retrieve records for all users.

Last Name: Doe

First Name:

Grant/Contract Number:

Status:  Active  Inactive  Active & Inactive  Pending

[Submit](#) [Reset](#)

After the search, results are presented and the user may click to view more details about the selected user.

ImmPort Bioinformatics FOR THE FUTURE OF IMMUNOLOGY

[Edit Profile](#) | [Sign Out](#)

**Admin / User Search Results**

My Profile | Programs | Grants - Contracts | Users | Reports

[Modify Search](#)

Users matching your search criteria are listed below. Click on a user's name to view user details.

2 items found, displaying all items.

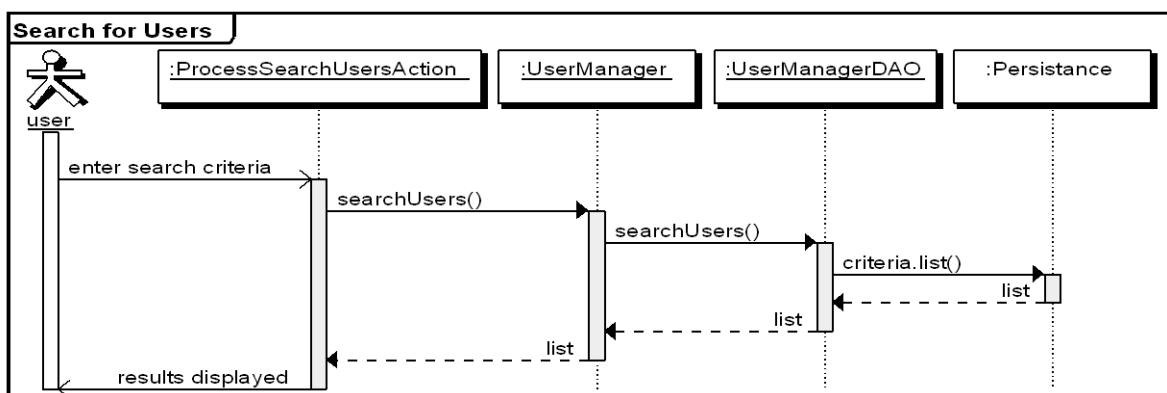
Last Name	First Name	MI	Status	Organization	Agreement Acceptance	Data Agreement Date	Community
Doe	Jane		ACTIVE	University			Life Science Researchers
Doe	John		ACTIVE	BISC			Life Science Researchers

2 items found, displaying all items.

**Search Users Class Diagram**



**Search Users Sequence Diagram**



### 3.6.2.6 Forgot password

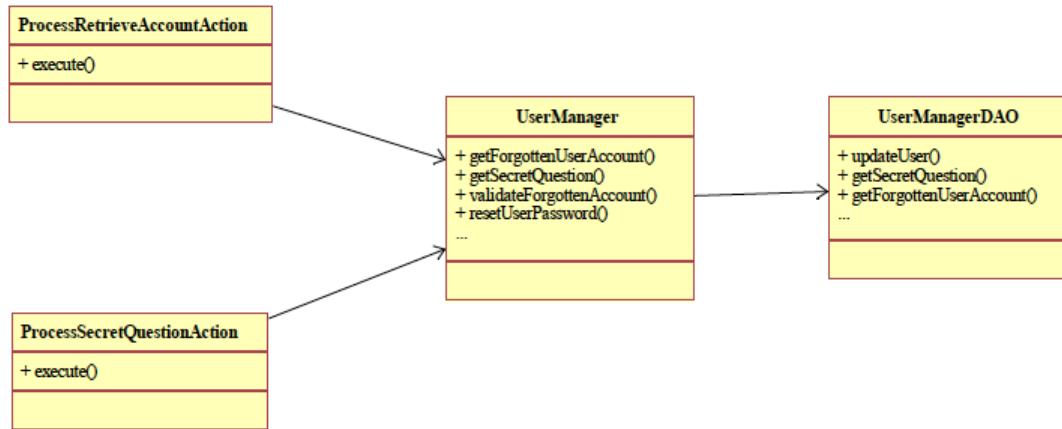
The Forgot Password Sequence Diagram depicts the sequence of events for a user that has forgotten their username/password.

The screenshot shows the ImmPort website's login page. At the top, there is a logo with a stylized flower icon and the word "IMMPORT" in orange, with the subtitle "BIOINFORMATICS FOR THE FUTURE OF IMMUNOLOGY". Below the logo is a navigation bar with links: Home, About ImmPort, Access Data, Tools, Resources, and News & Events. The main content area has a title "Retrieve User Name and Password" in orange. A sub-instruction says: "Please enter the following information to confirm your identity. The security question you provided during registration will be displayed." Below this, a note states: "Fields marked with an asterisk \* are required." There are four input fields: "First Name \*" with value "John", "Last Name \*" with value "Doe", "Middle Initial" (empty), and "Registered Email Address" with value "johndoe@research.org". At the bottom are two buttons: "Submit" and "Cancel". At the very bottom of the page, there is a footer with links: "Sponsored by: National Institute of Allergy and Infectious Diseases (NIAID)", "National Institutes of Health (NIH)", "Health and Human Services (HHS)", "Privacy Policy", "Disclaimer", "Accessibility", "Contact Us", and "Site Map".

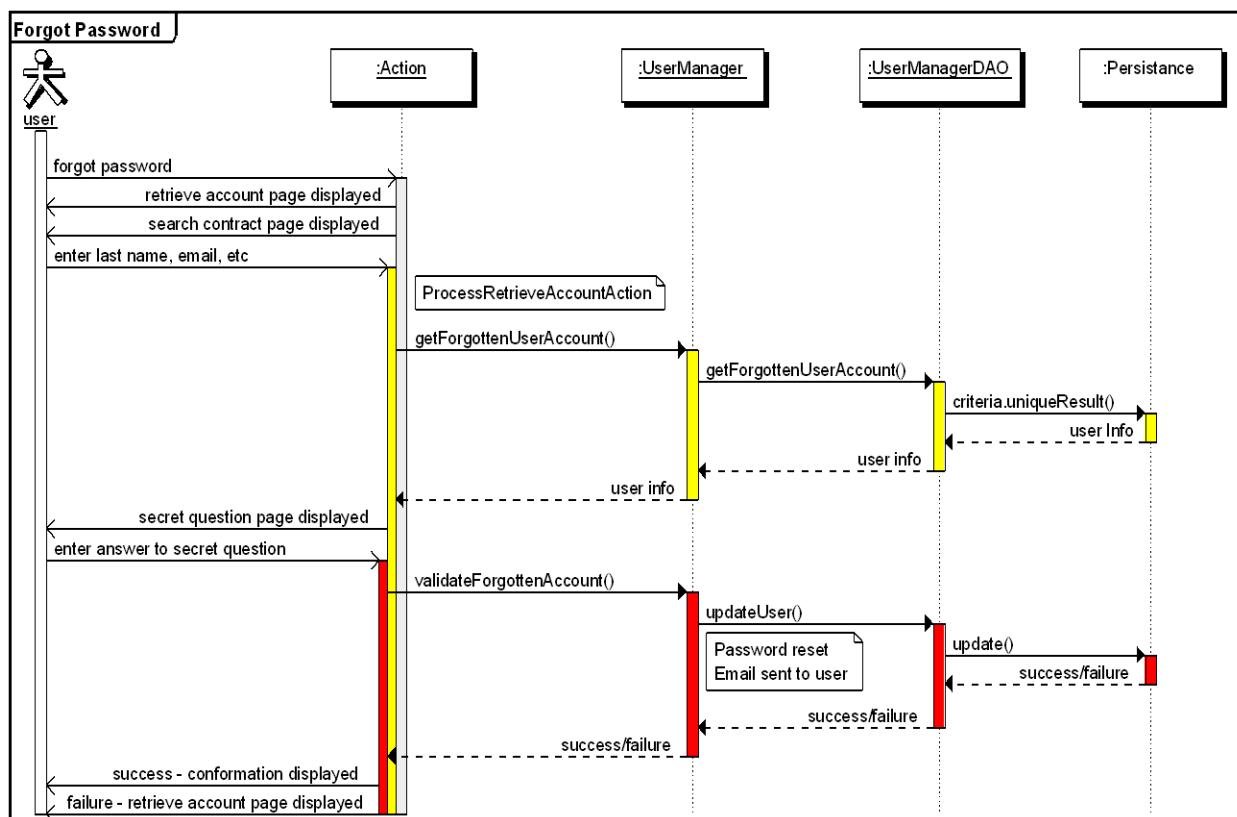
A security question is utilized to verify the user's identity, and the password is emailed upon success.

This screenshot shows the continuation of the password retrieval process. The top part is identical to the previous screenshot, featuring the ImmPort logo, navigation bar, and "Retrieve User Name and Password" form. The "Security Question" field contains "What is your favorite color?", and the "Answer" field contains "blue". Below these fields are "Submit" and "CANCEL" buttons. The footer at the bottom of the page includes the same "Sponsored by" links and "Privacy Policy" through "Site Map" links as the first screenshot.

**Forgot Password Class Diagram**



**Forgot Password Sequence Diagram**



### 3.6.2.7 Log in/Off

In order to log into the ImmPort system, a unique user-id is required with a password. These are created during the registration process or creation of an account. The password is restricted to 8-12 characters with one being a numeric value. Appropriate error messages are displayed, if applicable.

The screenshot shows the ImmPort sign-in interface. At the top, it says "Sign In". Below that is a "User Name:" field with a placeholder "User Name". Underneath is a "Password:" field with a placeholder "Password". To the right of the password field are two buttons: "SIGN IN" and "Forgot Password?".

Below the sign-in form is a section titled "Register for ImmPort" with a question mark icon. It lists "Benefits of Registration:" followed by three bullet points: "» Access to data visualization tools", "» Compare data sets to other researchers", and "» Access to data analysis tools".

The screenshot shows the ImmPort homepage. At the top, there is a navigation bar with links: "Edit Profile | Sign Out", "About ImmPort", "Admin", "Access Data", "Tools", "Resources", and "News & Events".

Below the navigation bar are four main tabs: "Browse Data", "Search Data", "Visualize Data", and "Analyze Data".

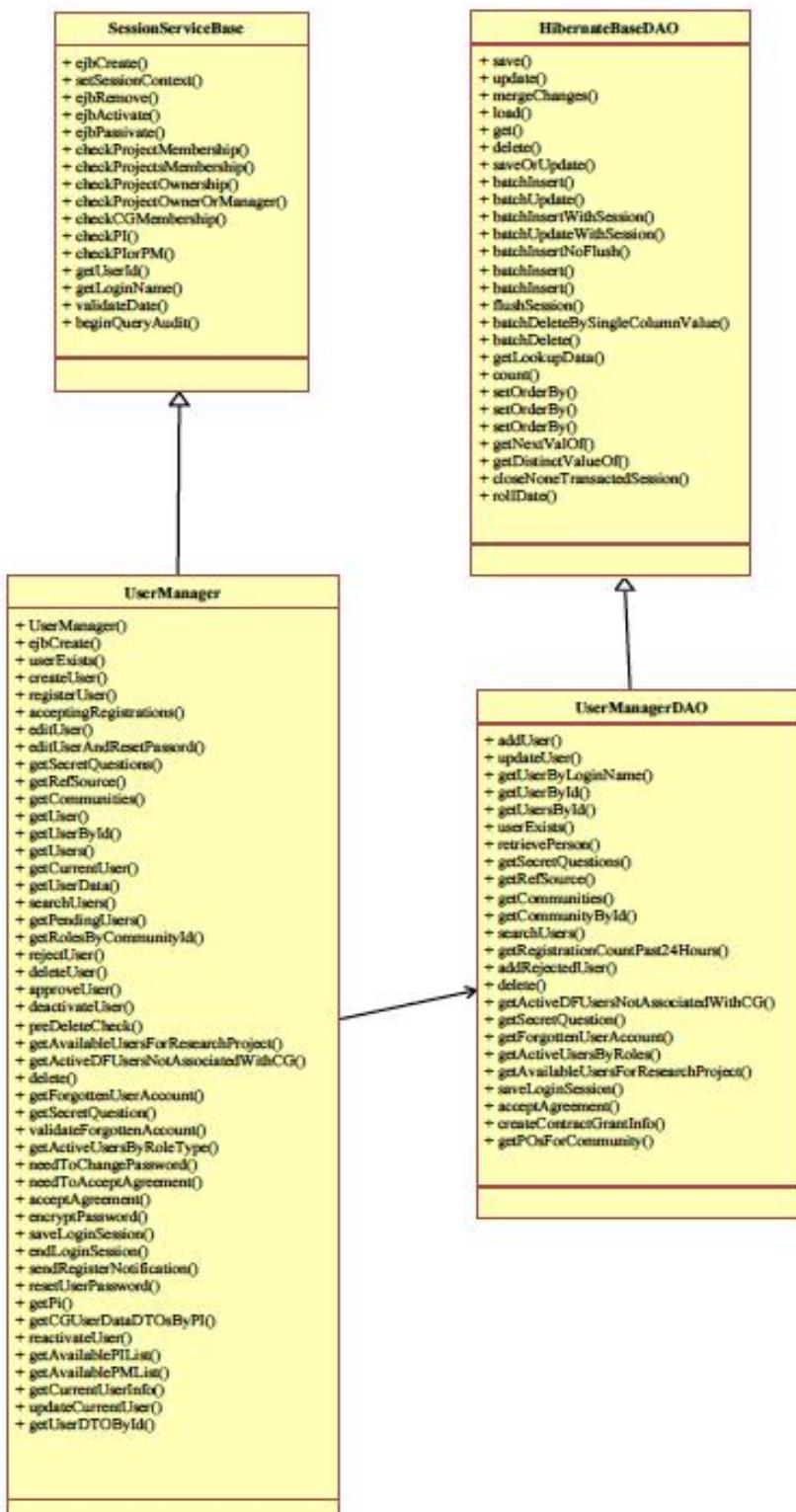
On the left side, there are two columns of links:

- ImmPort Research Data**: [NIAID Programs in ImmPort](#)
- Shared Research Data**: [Saved Lists](#), [Analysis Results](#)
- My Workbench**: [Private Research Data](#)
- News**: [Recent Publications in Immunology](#)

On the right side, there is a "Notifications" sidebar with a list of recent updates:

- Release of research data from the study by Dr. Thomas Casale: 'Allergen immunotherapy Co-administered with Omalizumab' is available [now](#).
- Doing flow cytometry? [FLOCK](#) has been upgraded for improved automated population identification.
- Analyze HLA typing data with ImmPort's [new ambiguity reduction tool](#) for typing data.
- Download summary tables, pathways, protein-protein interaction networks and GO terms for immunologically relevant genes from the [ImmPort Gene Lists](#).
- Submitting data to ImmPort? We want to help! We're available to assist you in all aspects of data submission, including formatting, validating and uploading your data. Send us an [email](#), and we'll work with you to ensure everything runs smoothly.

**User Administration Class Diagram**



### 3.6.3 Program, Contract, Project Management

ImmPort 2.0 vs. 3.0, etc.

#### 3.6.3.1 Create Program

The create program diagrams depict the events for creating a program. Appropriate error messages are displayed if errors are encountered.

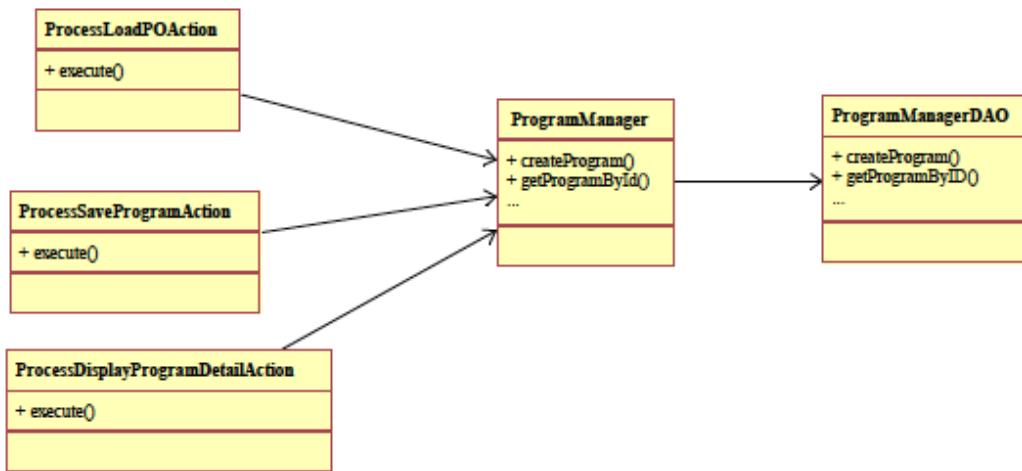
The screenshot shows a web-based application interface for creating a new program. At the top, there is a navigation bar with links: My Profile, Programs, Contracts-Grants, Users, and Reports. Below the navigation bar, the main title is 'Create Program'. A note states: 'Fields marked with an asterisk (\*) are required.' The form fields include:

- Title:** The Program
- Short Title (or Acronym):** Short Title for Program
- Description:** Program Description (text area)
- Category:** NIAID Program that will share data: \*\*\*will make Program metadata viewable on Home page (dropdown menu)
- Start Date (mm/dd/yyyy):** 09/30/2004 (date input field)
- End Date (mm/dd/yyyy):** 09/30/2011 (date input field)
- ImmPort Access End Date (mm/dd/yyyy):** (date input field)
- DAIT Information:**
  - Program/Project Officer:** Cheryl Kraft (text input field)
  - Deputy Program/Project Officer:** (text input field)

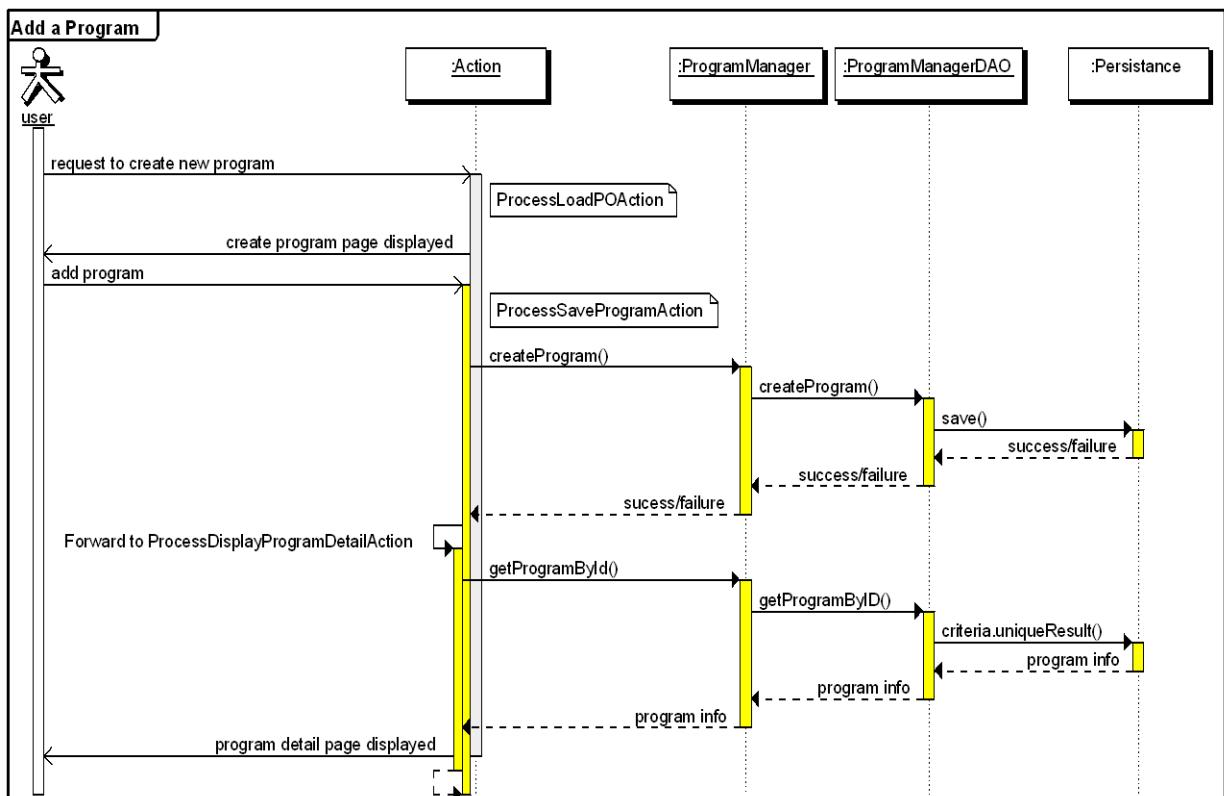
At the bottom of the form are two buttons: 'Create' and 'Reset'.

Once the program is created, contracts are associated.

**Create Program Class Diagram**



**Create Program Sequence Diagram**



### 3.6.3.2 Update Program

The Update Program diagrams depict the events for updating a program. Appropriate error messages are displayed if errors are encountered.

My Profile | Programs | Contracts-Grants | Users | Reports

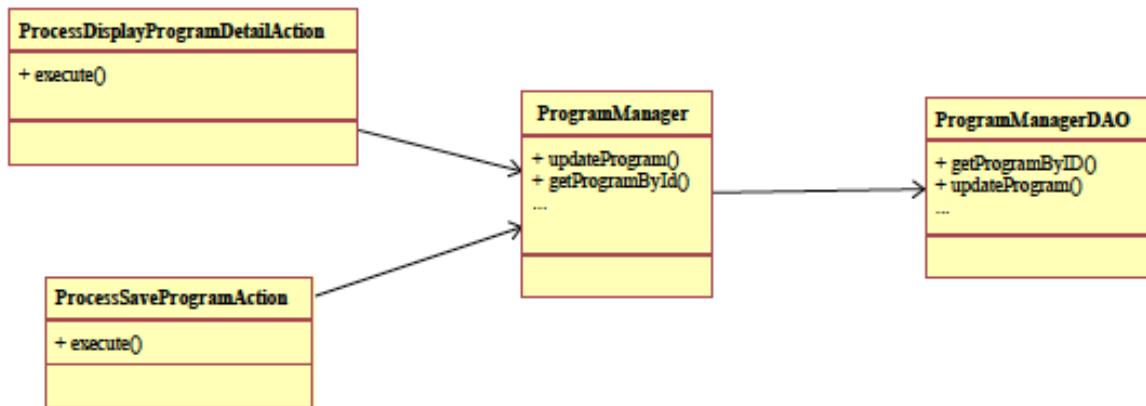
### Update Program

Fields marked with an asterisk \* are required.

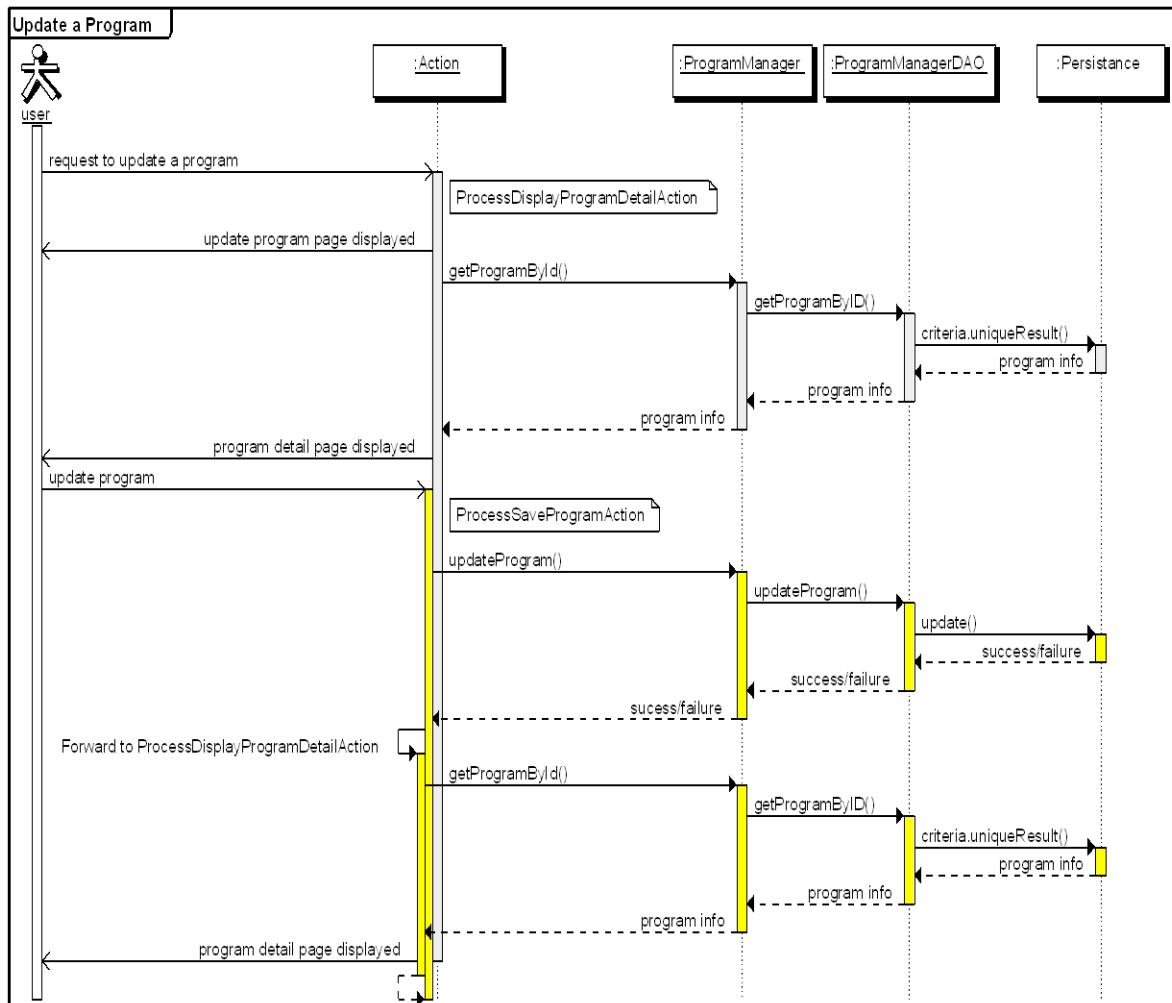
Title*	The Program
Short Title (or Acronym)*	Short Title for Program
Description*	Program Description
Category*	NIAD Program that will share data: ***will make Program metadata viewable on Home page
Start Date (mm/dd/yyyy)*	09/30/2004
End Date (mm/dd/yyyy)*	09/30/2011
Import Access End Date (mm/dd/yyyy)	09/30/2011
<b>DAIT Information</b>	
Program/Project Officer *	Cheryl Kraft
Deputy Program/Project Officer	

Save | Reset

#### *Update a Program Class Diagram*



**Update a Program Sequence Diagram**





### Search Program

The search program diagrams depict the events for searching a program. Appropriate error messages are displayed if errors are encountered.

Admin / Programs / Search Programs

My Profile | Programs | Contracts-Grants | Users | Reports

Enter criteria in the form below to search for a specific program. If you submit the form without entering search criteria, you will retrieve records for all programs.

Title

Short Title (or Acronym)

Category

Start Date

End Date

ImmPort Access End Date

Program/Project Officer (Last Name)

Status  Active  Inactive  All

The user is able to search on Title, Short Title, Start Date, End Date, ImmPort Access End Date, or Program/Project Officer.

Admin / Programs / Program Search Results

My Profile | Programs | Contracts-Grants | Users | Reports

[Modify Search](#)

Programs matching your search criteria are listed below. Click on a program title to view program details.

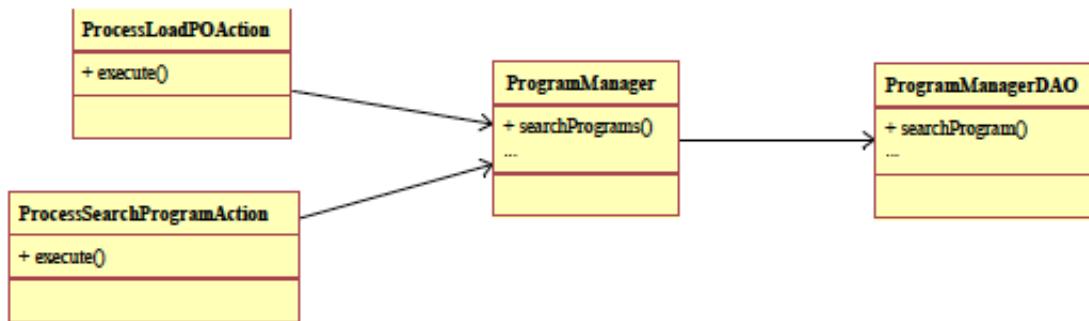
One item found.

Title	Short Title	Category	Status	Program/Project Officer
The Program	Short Title for Program	DAIT	Active	Cheryl Kraft

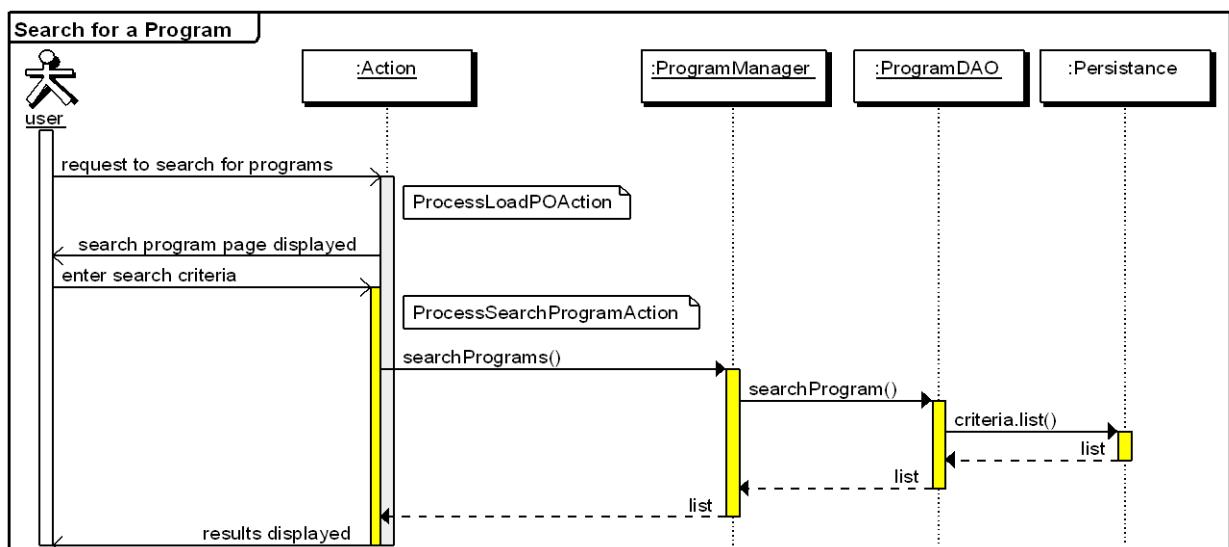
One item found.

A list of program search results is displayed. This allows the user to select a program to view or modify. In order to modify the search, click the “Modify Search” link.

**Search Program Class Diagram**



**Search Program Sequence Diagram**



**Program Administration Class Diagram**



### Add/Update Contract

The Add/Update Contract diagrams depict the events for adding and updating a contract. Appropriate error messages are displayed if errors are encountered.

The screenshot shows a Windows-style application window titled "Add/Update Contract". The form contains several groups of input fields:

- Contract Metadata:** Title\*, Short Title\*, Contract-Grant Number\*, Abstract\*
- Keywords:** Contract Keywords
- Category:** NIAID Contract-Grant that will share data: \*\*\*makes Contract metadata viewable on Home page
- Dates:** Start Date (mm/dd/yyyy)\*: 09/30/2004; End Date (mm/dd/yyyy)\*: 09/30/2011; ImmpOrt Access End Date (mm/dd/yyyy)
- DAIT Information:** Program\*: Other Programs dropdown (selected), Ken KTest; Program/Project Officer\*: Ken KTest; Deputy Program/Project Officer: Ken KTest

A "Submit" button is located at the bottom of the form.

Once the contract is added, a Principal Investigator, Program Manager and Other Staff is associated with the contract.

The screenshot shows the "Associated Users" and "Associated Projects" sections of the application:

- Associated Users:** A table listing users with checkboxes:

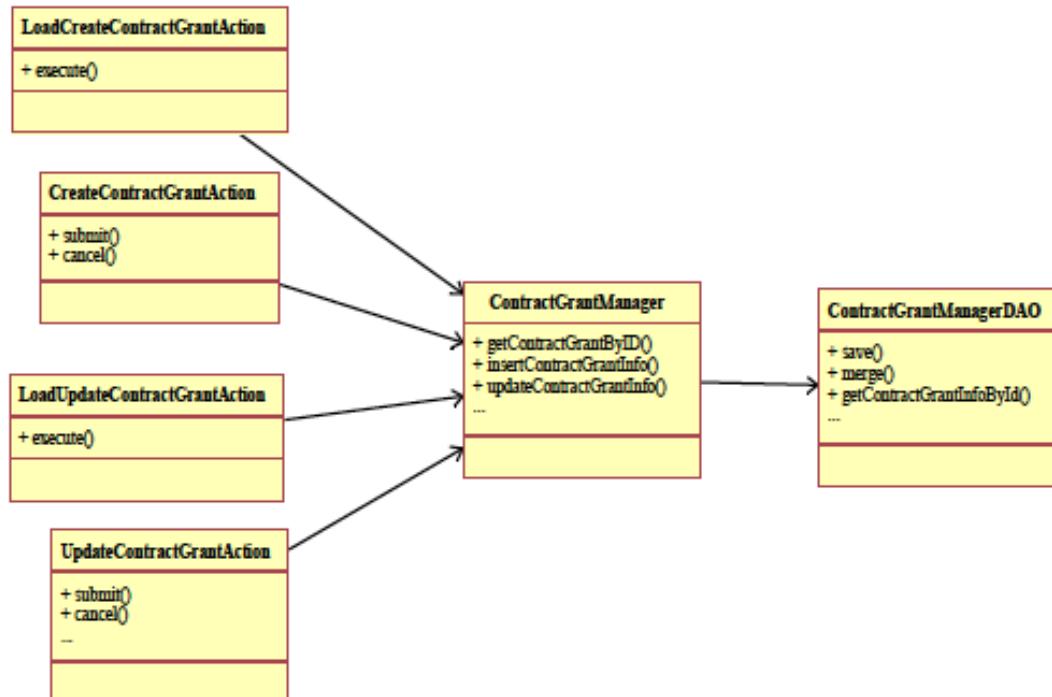
Name	Phone	Email	Role
Dunn, Patrick	301-527-6603	pdunn@immpOrt.org	OS
Wiser, Jeff	301-527-6673	jwiser@immpOrt.org	PM
Xia, Ashley	301-496-7551	axia@nIaID.nIh.gov	PI

Buttons for "Add..." and "Remove" are below the table.
- Associated Projects:** A section containing a "Create Default Project" button.

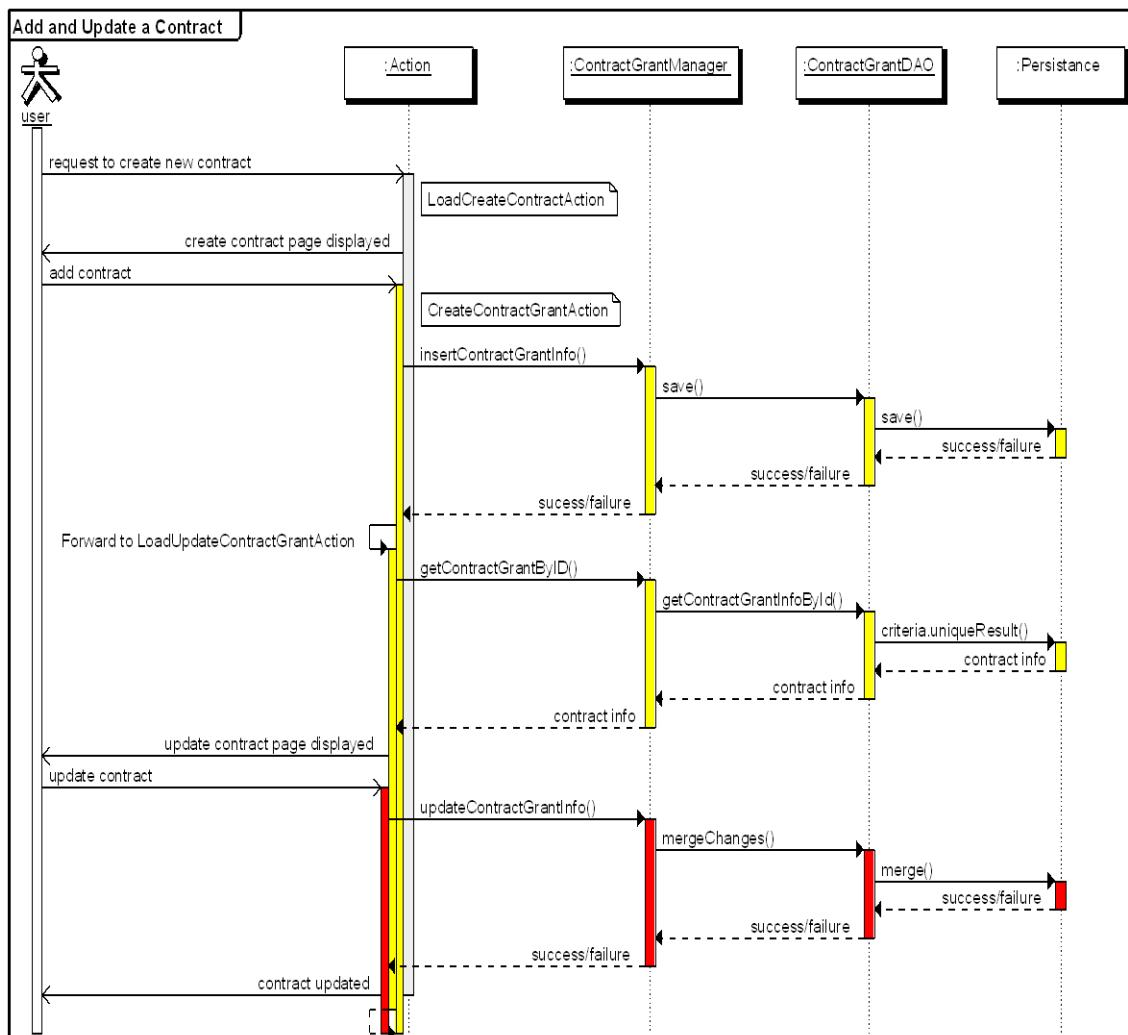
Once the associated users are added, a default project may be created. The associated users to the contract will obtain access to the default project.

Program	OTHER PROGRAMS																				
Program/Project Officer*	Ken KTest																				
Deputy Program/Project Officer																					
<input type="button" value="Submit"/> <input type="button" value="Cancel"/>																					
<hr/>																					
Associated Users																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th>Name</th> <th>Phone</th> <th>Email</th> <th>Role</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/></td> <td>Dunn, Patrick</td> <td>301-527-6603</td> <td>pdunn@immporg.org</td> <td>OS</td> </tr> <tr> <td><input type="checkbox"/></td> <td>Wiser, Jeff</td> <td>301-527-6673</td> <td>jwiser@immporg.org</td> <td>PM</td> </tr> <tr> <td><input type="checkbox"/></td> <td>Xia, Ashley</td> <td>301-496-7551</td> <td>axia@nifld.nih.gov</td> <td>PI</td> </tr> </tbody> </table>			Name	Phone	Email	Role	<input type="checkbox"/>	Dunn, Patrick	301-527-6603	pdunn@immporg.org	OS	<input type="checkbox"/>	Wiser, Jeff	301-527-6673	jwiser@immporg.org	PM	<input type="checkbox"/>	Xia, Ashley	301-496-7551	axia@nifld.nih.gov	PI
	Name	Phone	Email	Role																	
<input type="checkbox"/>	Dunn, Patrick	301-527-6603	pdunn@immporg.org	OS																	
<input type="checkbox"/>	Wiser, Jeff	301-527-6673	jwiser@immporg.org	PM																	
<input type="checkbox"/>	Xia, Ashley	301-496-7551	axia@nifld.nih.gov	PI																	
<input type="button" value="Add..."/> <input type="button" value="Remove"/>																					
<hr/>																					
Associated Projects																					
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Project Title</th> <th>Project Category</th> <th>Project Description</th> </tr> </thead> <tbody> <tr> <td>Contract Title</td> <td>DAIT</td> <td>Contract Abstract</td> </tr> </tbody> </table>		Project Title	Project Category	Project Description	Contract Title	DAIT	Contract Abstract														
Project Title	Project Category	Project Description																			
Contract Title	DAIT	Contract Abstract																			

### Add/Update Contract Class Diagram



**Add/Update Contract Sequence Diagram**



### Search Contract

The Search Contract diagrams depict the events for searching a contract. Appropriate error messages are displayed if errors are encountered.

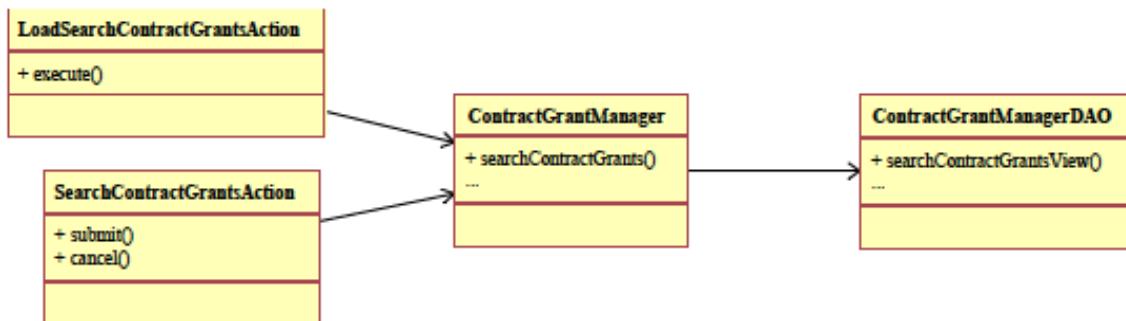
The screenshot shows a search form titled "Admin / Search Contracts-Grants". The form includes fields for Title, Contract-Grant Number, Contract-Grant Category (a dropdown menu listing categories like Immune Function and Biodefense in Children, Elderly, and Imm...), Start Date, End Date, ImmPort Access End Date, Principal Investigator (Last Name), Associated Program Name (a dropdown menu listing programs like Population Genetics II, Human Immunology Project Consortium, Modeling Immunity for Biodefense II, Atopic Dermatitis Research Network (ADRN)), and Status (radio buttons for Active, Inactive, or All). Below the form are "Submit" and "Reset" buttons.

A user is able to search on the fields Title, Grant – Contract Number, Category, Start Date, End Date, ImmPort Access End Date, Principal Investigator, Associated Program Name or Status.

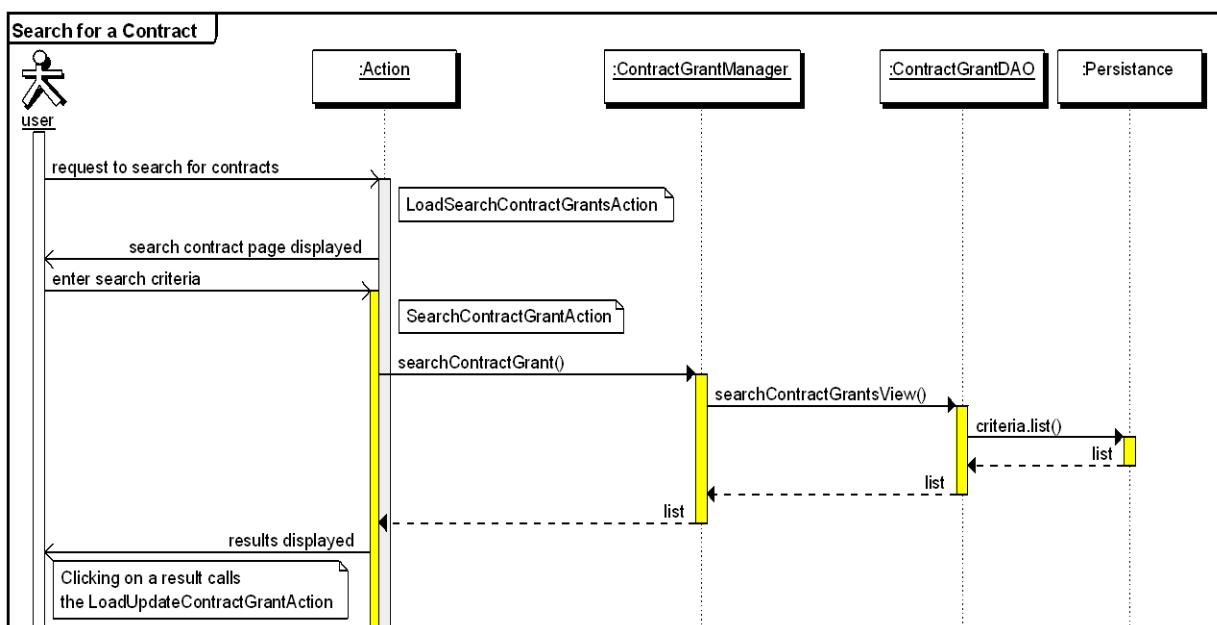
The screenshot shows a search results page titled "Admin / Contracts-Grants Search Results". It features a header with the IMMPORt logo and navigation links for About IMMPORt, Admin, Access Data, Tools, Resources, and News & Events. The main content area displays a table of search results with columns for Contract-Grant Number, Title, Status, Category, Program, Principal Investigator, and End Date. One row is visible in the table, showing Contract Number, Contract Title, Active status, DAIT category, Other Programs program, Xia Ashley principal investigator, and 2011-09-30 end date.

A list of Grants - Contracts search results is displayed. This allows the user to select a Grants - Contracts to view or modify. In order to modiy the search, click the “Modify Search” link.

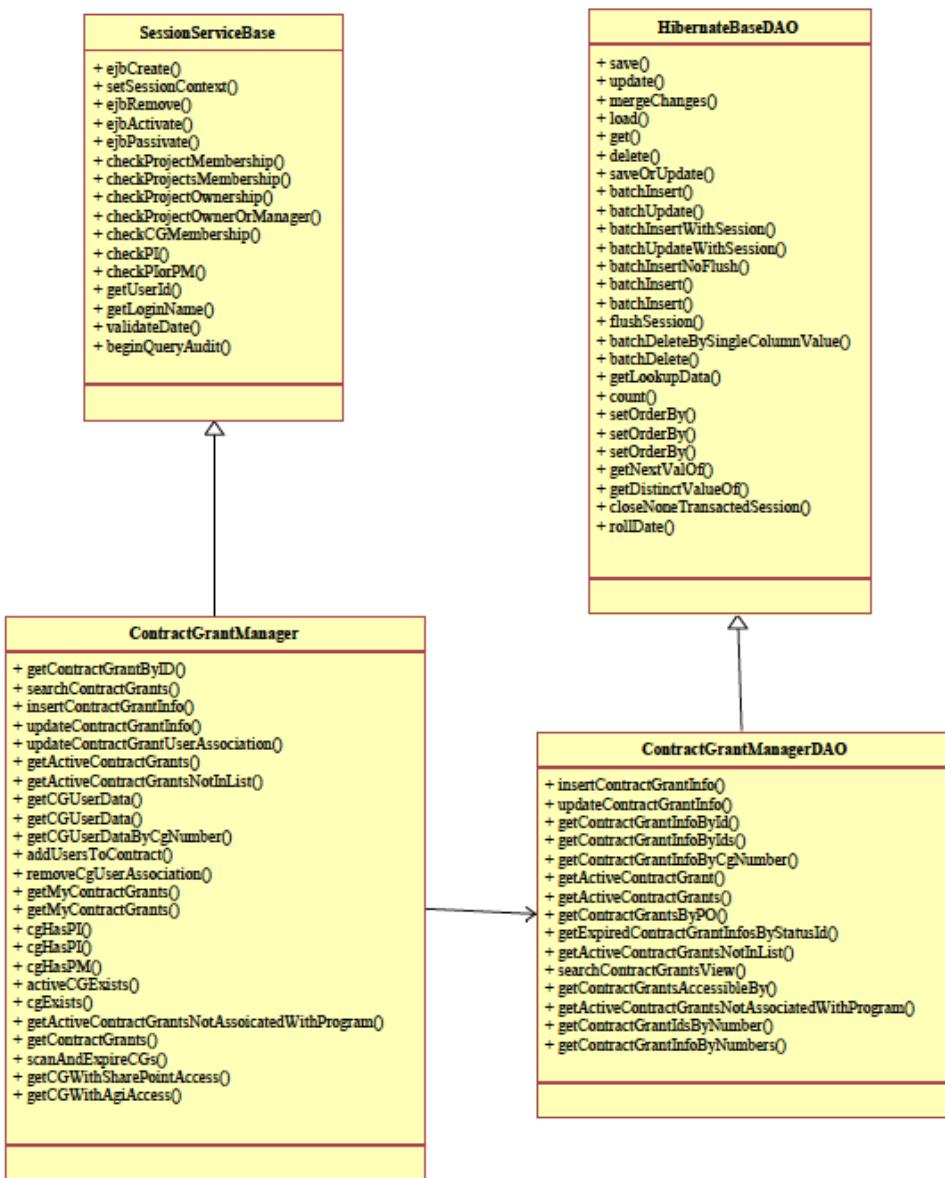
### Search Contract Class Diagram



### Search Contract Sequence Diagram



**Contract Administration Class Diagram**



### Add/Update Projects

The Add/Update project diagrams depict the events for adding and updating a project. Appropriate error messages are displayed if errors are encountered.

The screenshot shows a web application interface for 'IMMPORT' (Bioinformatics for the Future of Immunology). At the top, there's a navigation bar with links for 'Edit Profile', 'Sign Out', 'About ImmPort', 'Admin', 'Access Data', 'Tools', 'Resources', and 'News & Events'. The main content area has a header 'Admin / Create Research Project: Notice'. Below it, there are links for 'My Profile', 'Grants - Contracts', and 'Projects'. A section titled 'Terms and Conditions for Creating and Using Private Project Workspaces' contains text about the purpose of workspace creation and a list of rules. The rules include: creating more than one workspace, facilitating data analysis, ensuring confidentiality, reporting breaches, and maintaining system security. Below the rules, a note states that by continuing, the user agrees to these terms. At the bottom, there are 'Continue' and 'Cancel' buttons.

In order to create a project, a user must read the terms and conditions for creating a Private Project Workplace.

**Admin / Create Project**

My Profile | Contracts-Grants | Projects

**Research Project**

Fields marked with an asterisk \* are required.

Title *	<input type="text" value="Project Title"/>
Description *	<input type="text" value="Project Description"/> (max 4000 chars)
Keywords *	<input type="text" value="Project Keywords"/>
Category *	<input type="text" value="Test Project: can be used for any testing purpose"/>
Contract-Grant*	<input type="text" value="Sally-Contract (TEST)"/>
Project Owner *	Patty-PI Tester
Project Manager	<input type="button" value="..."/> <a href="#">Remove Project Manager</a>

The user enters the required fields, Title, Description, Keywords, Category, and Contract-Grant. The Project Owner is auto generated. An appropriate error message will display if an error has occurred.

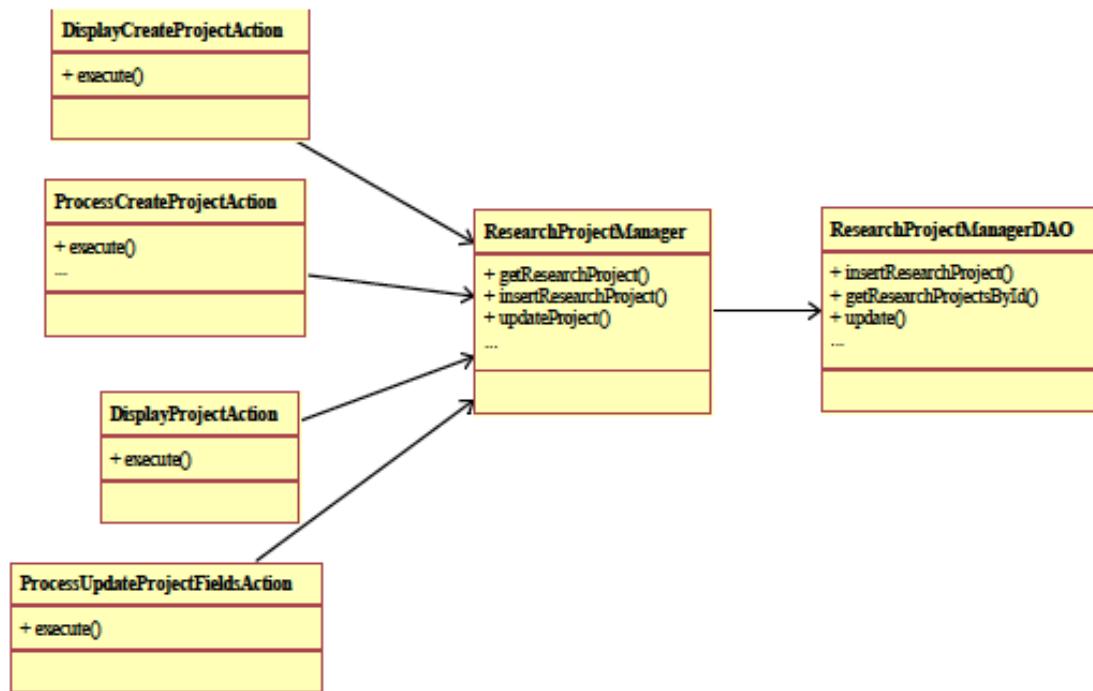
Your workspace has been created.

**Associated Users**

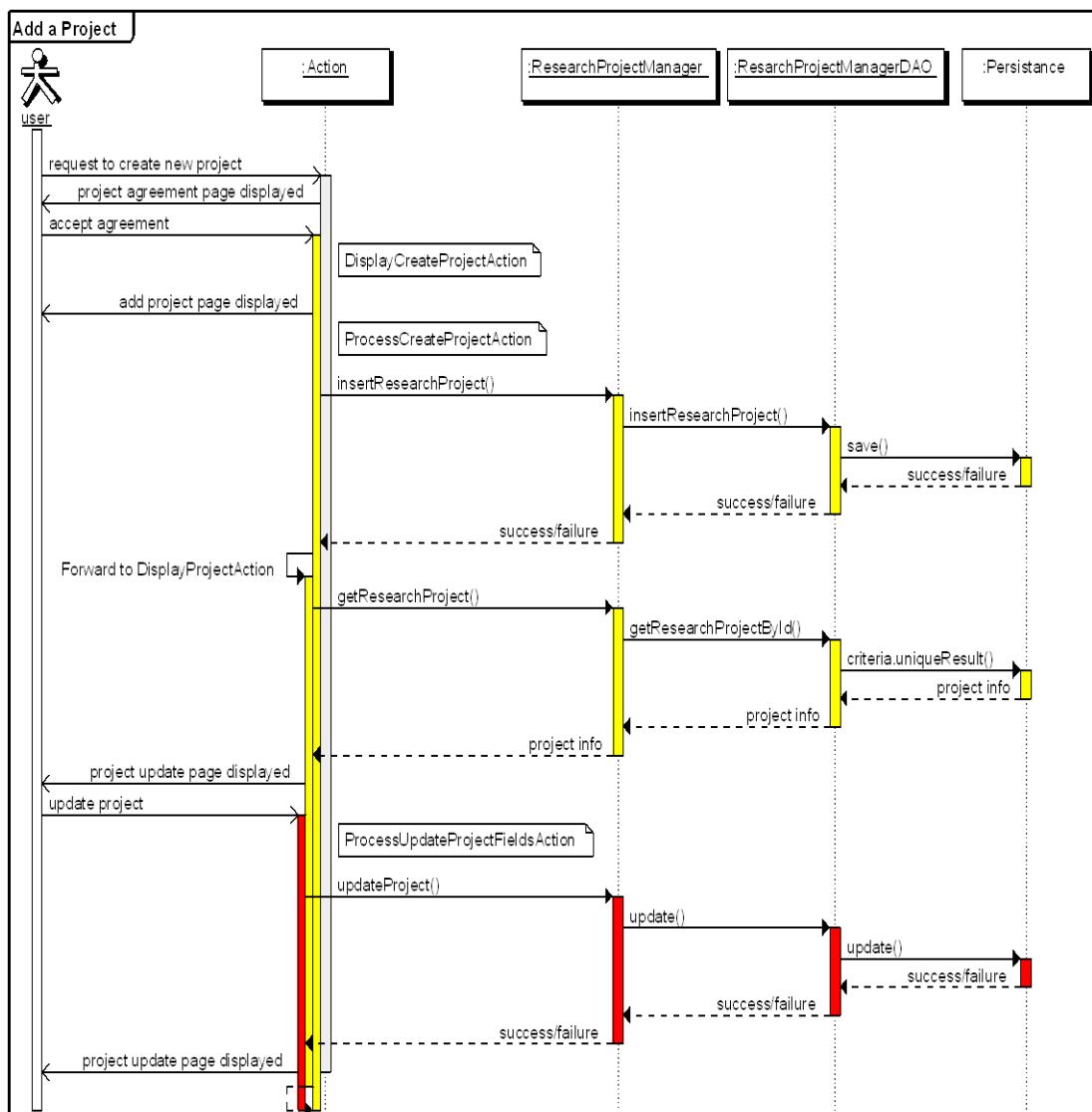
The following users are associated with this research project. Add users by clicking "Add..." Remove users by clicking on the corresponding checkboxes and clicking "Remove."

No results found.

Once the project is created, the user is able to update the project and/or associate users to their project.

**Adding/Updating A Project Class Diagram**

**Adding/Updating a Project Sequence Diagram**



### Display Projects

The display projects diagrams depict the events for viewing a project.

My Research Projects			
	Project Title	Project Category	Project Description
<input type="radio"/>	Atopic Dermatitis & Vaccinia Network (ADVN) Cli...	DAIT	Atopic Dermatitis & Vaccinia Network (ADVN) Clinical Studies Consortium
<input type="radio"/>	ITND019AD - Clinical Trial	DAIT	Phase II, Double-Blinded, Placebo-Controlled, Efficacy and Safety Evaluation...
<input type="radio"/>	McMaster Test Project	DAIT	McMaster
<input type="radio"/>	Mt. Sinai VIP004	DAIT	for testing Mt Sinai load files
<input type="radio"/>	Project Title	TEST	Project Description
<input type="radio"/>	Test Example Packages	TEST	testing
<input type="radio"/>	Test Project	DAIT	Test Project
<input type="radio"/>	UAB PopGen	DAIT	study design subs, bioSams, etc
<input type="radio"/>	UWash PoPGen test GT data	TEST	UWash PoPGen test GT data
<input type="radio"/>	p2	BISC	a
<input type="radio"/>	rene test	DAIT	rene test
<input type="radio"/>	t	DAIT	t

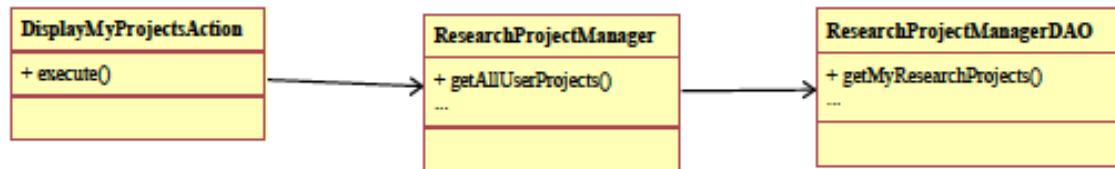
  

My Collaborative Projects			
	Project Title	Project Category	Project Description
<input type="radio"/>	Collaborating	DAIT	Collaborating with users
<input type="radio"/>	Collaborative Project 123	DAIT	Sandbox for me to play
<input type="radio"/>	Rene's cw	BISC	a cw for test
<input type="radio"/>	hey	DAIT	what.
<input type="radio"/>	rene cp 1	LSR	rene cp2 k

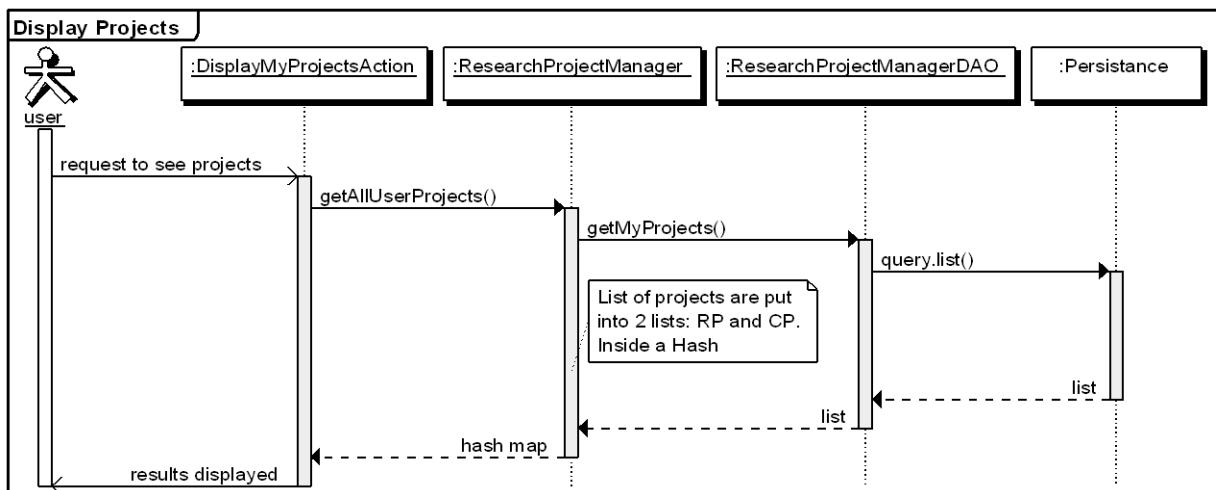
[View Project Details](#)   [Manage Project Data](#)

A user is able to select the project while viewing the project details or manage project data.

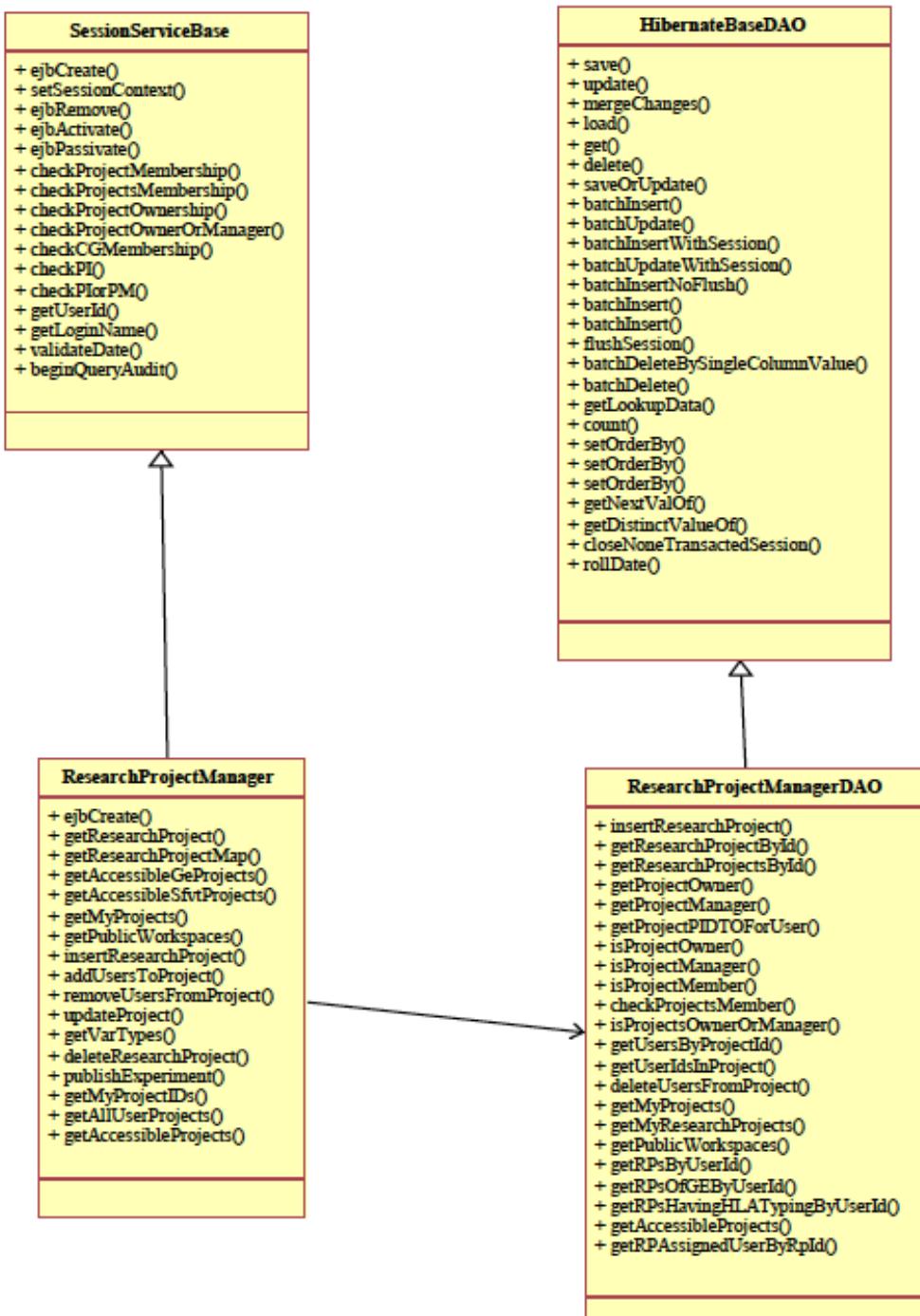
### Display Projects Class Diagram



**Display Projects Sequence Diagram**



**Project Administration Class Diagram**



### 3.6.4 Flow Cytometry Analysis

Flow cytometry analysis defines and quantifies cellular features making it an essential tool employed in immunology research. Challenges inherent in the management and interpretation of flow cytometry data led to the development of the ImmPort Flow Cytometry Analysis tool and its four supporting modules:

1. Data Management—single/multiple file upload, data set creation/editing
2. FCS file analysis using the FLOCK algorithm
3. Cross Sample analysis
4. Viewing/editing analysis results

The Data Management module allows the user to upload individual or multiple flow cytometry files, edit the descriptive content of the files and create/edit sets of uploaded files. To further facilitate data management and data set creation ImmPort employs components of the search module to find files of interest.

The FLOCK (FLOw Clustering without K) algorithm was developed as a novel method to identify unique populations within a single flow cytometry file. Analysis output includes two dimensional images of FLOCK-defined populations enhanced by the inclusion of the following tables for viewing or download: results summary, centroid, mean fluorescence intensity. (FLOCK publication: <http://www.ncbi.nlm.nih.gov/pubmed/20839340>).

Cross Sample analysis automatically maps populations of cells across multiple flow cytometry files utilizing centroid values saved from previous FLOCK results. Analysis results are displayed in two-dimensional format for each file in the comparison allowing the user to select which population/fluorescent marker to view across all files. Population proportions are available for all files in the comparison analysis as are results for individual files including centroid and mean fluorescence intensities.

Viewing and editing analysis results applies to FLOCK and individual results in Cross Sample Comparison. The Result Adjustment System (RAS) allows the user to edit two-dimensional images via moving, splitting or deleting a centroid for improved representation of the analyzed results. Supporting statistical output is recalculated to support the analysis edits.

The ImmPort Flow Cytometry analysis workflow is depicted below:

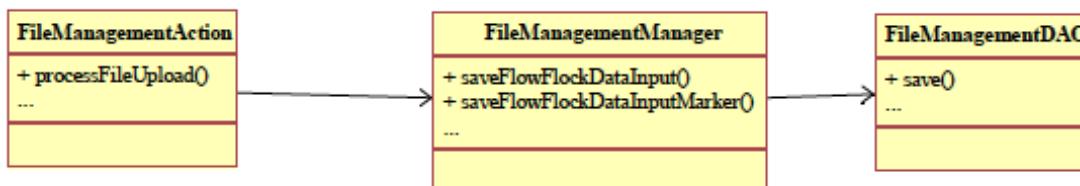


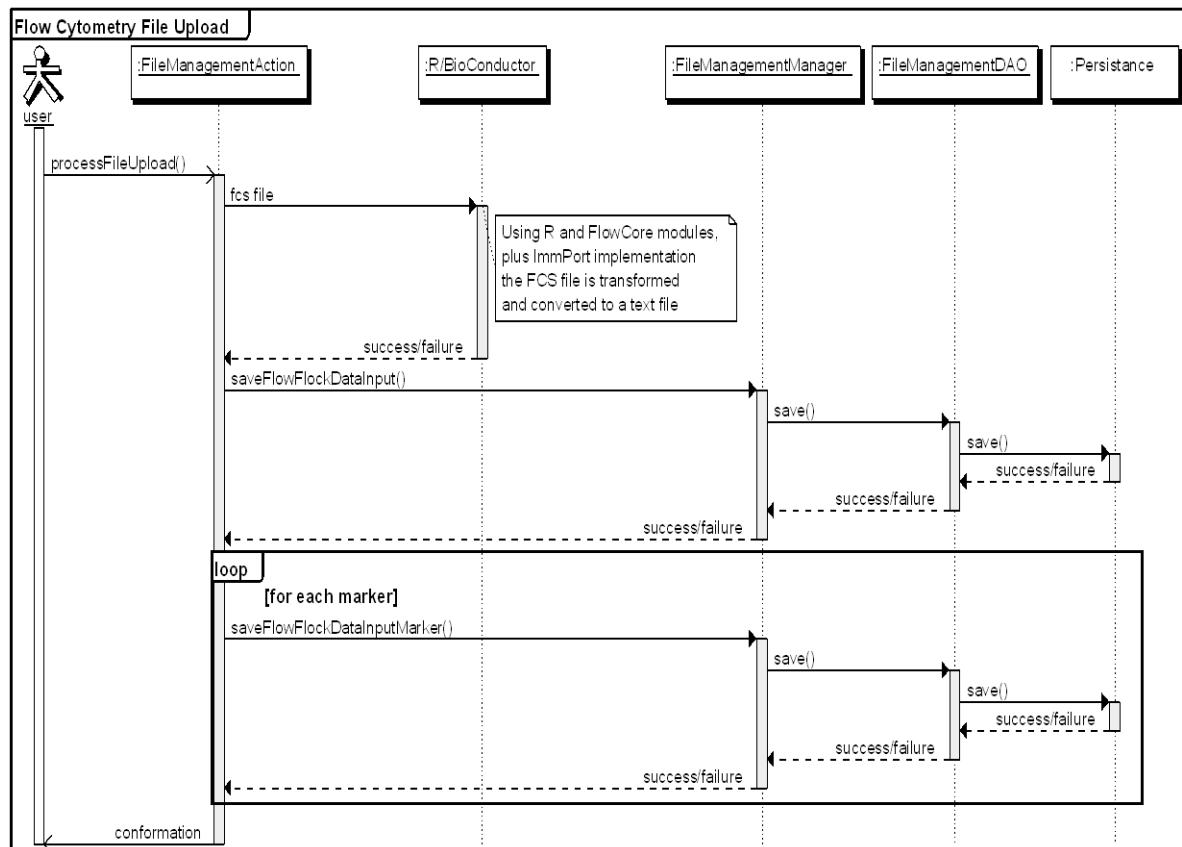
### Data Management: File Upload

The output from a flow cytometer is typically a binary file that conforms to the Flow Cytometry Standard (FCS). Most analysis algorithms including FLOCK require the FCS binary file be converted and possibly transformed into a text file representation of the fluorescence expression values. The ImmPort Data Management File Upload module support both single and multiple FCS file upload. Uploaded files are automatically converted, transformed and stored in the ImmPort database. The conversion algorithm, FCSTrans, was written using R/BioConductor modules by the BISC team as described in <https://www.immport.net/help/ImmPort.FCS.conversion.pdf>. The ImmPort system supports linking of FCS files to metadata such as subject, study and experimental information, previously loaded into the ImmPort system.

The screenshot shows the ImmPort web application interface. At the top, there is a navigation bar with links for 'Edit Profile' and 'Sign Out'. Below the navigation bar, the main header reads 'Flow Cytometry Analysis (Beta) / Data Management / File Upload'. A sub-header below it lists navigation items: 'Flow Cytometry', 'Data Management', 'FLOCK', 'View/Edit Results', 'Cross Sample', and 'Help'. A message at the top states: 'A project is required to upload data. Please see the [User Guide](#) or contact the [help desk](#) for more information.' A dropdown menu labeled 'Select project to store uploaded file' contains the option 'Test research project for v2.8'. Below this, a note specifies: 'The uploaded data file can be in either .fcs or .txt format. The filename can't contain the following characters: " " \ and /. ImmPort automatically converts .fcs files to .txt files when the upload includes only .fcs files. The .txt files can be created using third party tools including Tree Star FlowJo™ on MacOS. For conversion details, please see the [help section](#).' There are input fields for 'Name \*' (with a red asterisk), 'Description', 'Data File (.fcs)', 'Data File (.txt)', and 'Marker File (.info)'. Each input field has a 'Browse...' button to the right. To the right of each file input, there are three hyperlinks: 'example Data File (.fcs)', 'example Data File (.txt)', and 'example Marker File (.info)'. A note at the bottom left says: 'NOTE: Clicking "Upload" will upload the above files into the private project workspace of your chosen project and create the data set.' A large blue 'Upload' button is located at the bottom left.

### Upload Class Diagram



**Upload Sequence Diagram**

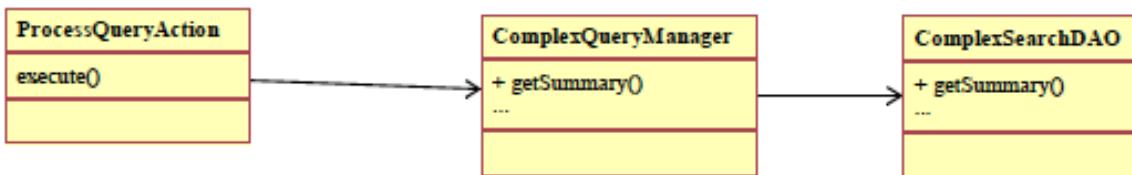
**Data Management: Create Analysis Set**

The Data Management module facilitates the creation of analysis sets by employing existing search module code and technologies outlined in the search module portion of this document. Created analysis sets may be used as input to either FLOCK or Cross Sample Analysis while existing data sets may be edited to add or remove FCS files. The UI consists of 3 panels:

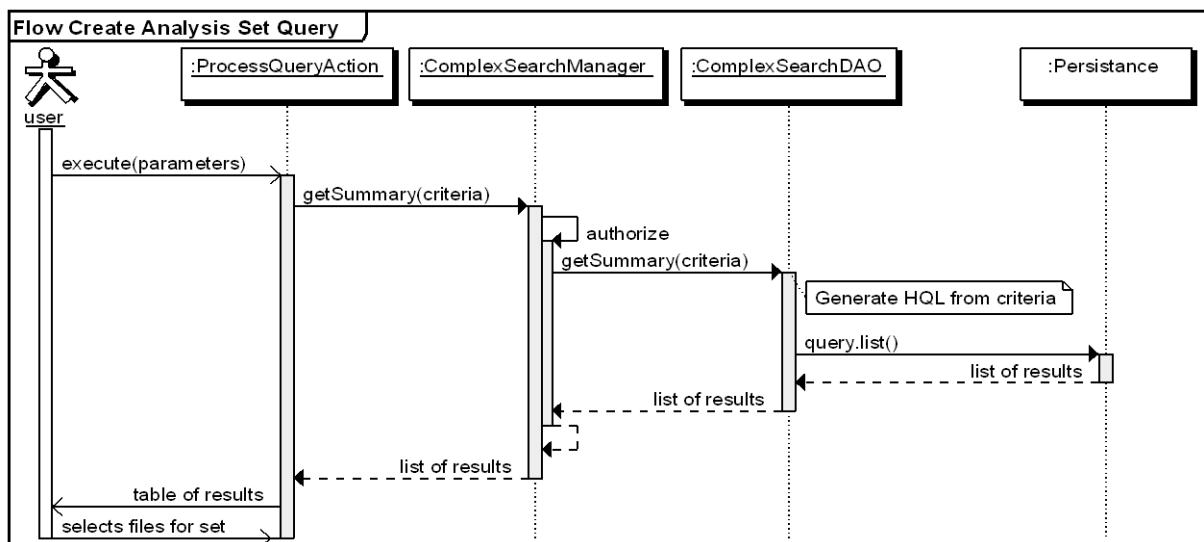
- Search—used to query for available FCS files
- Results—displays search results
- Create Set—files of interest are moved to this panel to save the set to the database.

The meta information for the analysis sets is stored in the ImmPort data base.

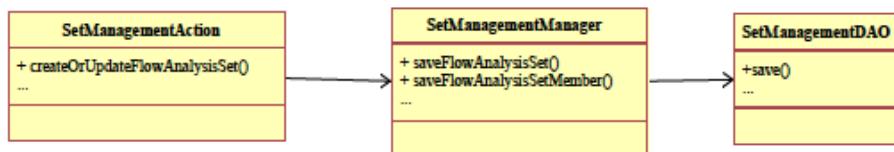
The screenshot shows the 'Flow Cytometry Analysis (Beta) / Data Management / Create Analysis Set' interface. At the top, there is a navigation bar with links: Flow Cytometry, Data Management, FLOCK, View/Edit Results, Cross Sample, and Help. Below the navigation bar, a message says 'Use the "Uploaded Flow Files Search" section to find data files for dataset creation. The number of items in a dataset is limited to 300.' A dropdown menu 'Select project:' shows 'Test research project for v2.8' and an unchecked checkbox 'Include Semi-Public files in the Search'. The main area has three sections: 'Uploaded Flow Files Search' (with a search bar for 'Flow: Name' and a dropdown for 'Results per page: 25'), a table of loaded files (Displaying 1 - 25 of 49), and a 'Create Set' panel (listing 'test fcs upload' as an FCS Text File). A note at the bottom says 'Use arrow button to move files in or out of analysis set. Mouse-over column titles, click down arrow to hide or show columns.'

**Create Analysis Set Query Class Diagram**

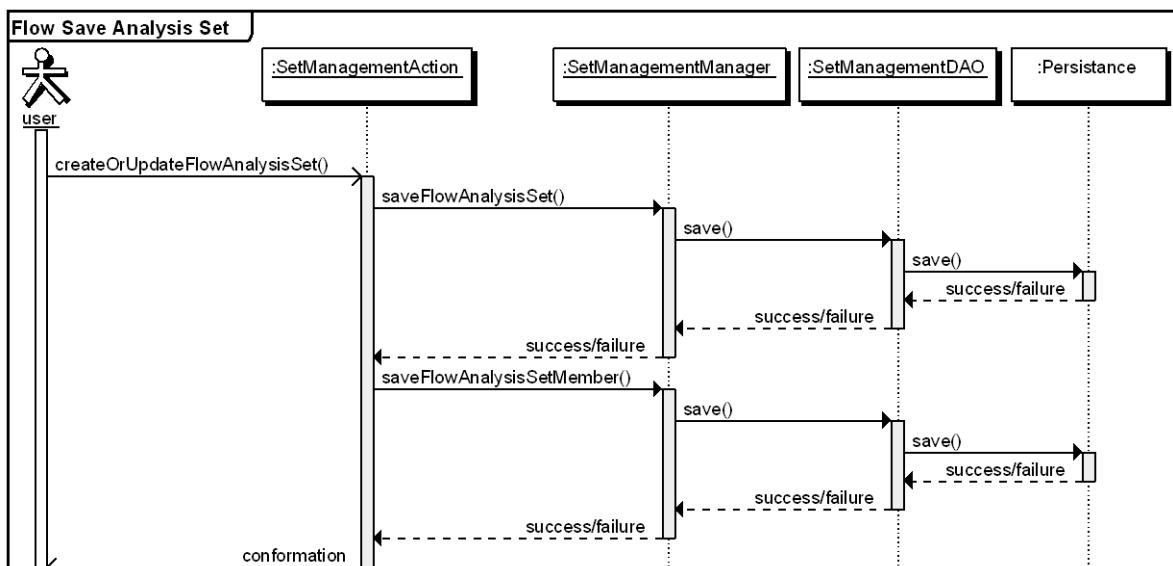
**Create Analysis Set Query Sequence Diagram**



**Create Analysis Set Save Class Diagram**



**Create Analysis Set Save Sequence Diagram**



## FLOCK Analysis

FLOCK analysis uses a rapid binning approach to determine the number of unique populations in high dimensional flow cytometry data. The user interface provides two means to initiate a FLOCK analysis:

- Individual file selection—single or multiple files can be selected for FLOCK analysis from the Uploaded Flow Files Search screen accessed via FLOCK/Analyze Individual Files dropdown menu
- Data set selection—previously created data sets can be selected for FLOCK analysis via FLOCK/Analyze Data Sets dropdown menu

From either analysis initiation point the process for running FLOCK requires the same steps: create an analysis task with the data set or selected individual files, select analysis settings, place the task in a queue, the FlockWorker retrieves the task and a Java program is called for the following:

1. Extracting the text version of the FCS file from the database and storing it in a task folder.
2. The FLOCK program written in ‘C’ reads the extracted FCS file and generates detailed and summary results representing the cell populations.
3. The result files are processed by the Java program generating overview images.

The FlockWorker process marks the task as completed making the results available for user review and subsequent editing.

### ***FLOCK analysis: individual file selection***

Name	Description	Exp Sample Acc	Status	FCS Text File	Date Created	Created By
Sub2Sample1			loaded	Sub2Sample1.txt	2010-09-21	Thomson
Sub2Sample1			loaded	Sub2Sample1.txt	2010-09-21	Thomson
Treg pure			loaded	DD082009_017.txt	2010-09-21	Thomson
Teff pure			loaded	DD082009_015.txt	2010-09-21	Thomson
CD4 purified			loaded	DD082009_012.txt	2010-09-21	Thomson
Spleen-pLN			loaded	DD082009_009.txt	2010-09-21	Thomson
Single file upload test IE8			loaded	Sub1Sample1.txt	2010-09-21	Thomson

**FLOCK analysis: data set selection**

**Flow Cytometry Analysis (Beta) / FLOCK / Run FLOCK Using Data Set**

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

Select project: Test research project for v2.9.1

Double click on a row to view detail screen. Mouse-over column titles, click down arrow to hide or show columns

Name	Description	Status	Project	Date Created
20100923 save all from spw 2	xxxxxx	loaded	Test research project for v2.9.1	2010-09-23
Dataset4	4files 7col	loaded	Test research project for v2.9.1	2010-09-21
Dataset3	save_all_7items	loaded	Test research project for v2.9.1	2010-09-21
Dataset2	4 items 7 col	loaded	Test research project for v2.9.1	2010-09-21
Dataset1	2 items 6 col	loaded	Test research project for v2.9.1	2010-09-21
20100921 SPW set	xxxxx	loaded	Test research project for v2.9.1	2010-09-21

**FLOCK Analysis Setting**

**Flow Cytometry Analysis (Beta) / FLOCK / FLOCK Analysis Setting**

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

**FLOCK Detail**

Project:	Test research project for v2.8	FLOCK Version :	2.0
Name:	<input type="text"/>	FLOCK Parameters:	<input checked="" type="radio"/> Auto <input type="radio"/> Manual
Description:	<input type="text"/>	Number of Bins (6-30) :	<input type="text"/>
		Density (2-100) :	<input type="text"/>

Use check boxes to select markers to be included in analysis. Markers to be excluded from analysis may include unused markers/channels, limiting multiple scatter parameters (FSC-A, FSC-H, FSC-W) to one parameter (e.g., FSC-H) or pre-gated markers. Please review [the Marker Selection tutorial](#) or [the User Guide](#) for more information.

**Column Selection for 3 Files**

Name	Marker #	1	2	3	4	5	6
Sample 001	6	Forward	Side	FITC	PE	Cy5	Cy7
Sample 003	6	Forward	Side	FITC	PE	Cy5	Cy7
Sample 002	6	Forward	Side	FITC	PE	Cy5	Cy7

**FLOCK: Analysis Status**

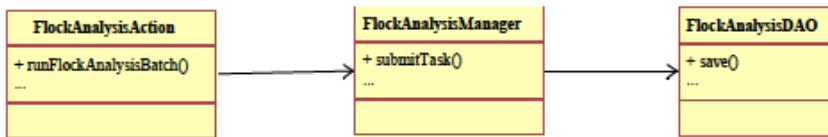
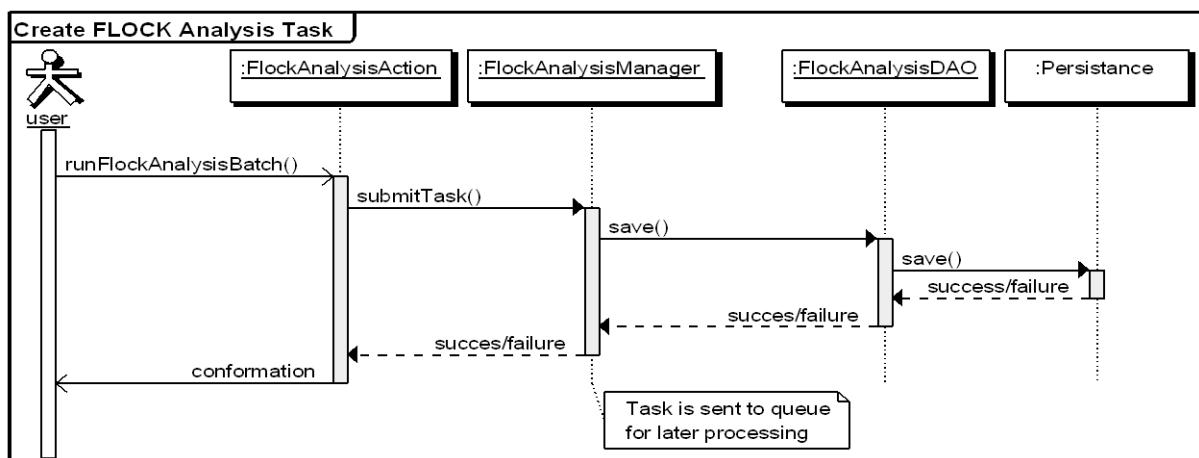
**Flow Cytometry Analysis (Beta) / FLOCK / Analysis Status**

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

Your FLOCK Analysis request was submitted successfully. Your task ID is **2619**.  
You will be sent an email when your task has been completed.

Click the "Show Analysis Results" button to proceed to the FLOCK Analysis History page.

**Show Analysis Results**

**Submit FLOCK Analysis Task Class Diagram****Submit FLOCK Analysis Task Sequence Diagram****Cross Sample Analysis**

Cross Sample analysis automatically maps populations of cells across multiple flow cytometry samples and computes the summary statistics for downstream analysis. The analysis requires a previously created data set and saved centroid file which are presented for user selection in the Cross Sample Analysis UI. The initiating step for Cross Sample Analysis is the creation of an analysis task comprised of the data set and centroid which is followed by a review of the analysis components via Cross Sample Analysis Settings display. The task is placed in a queue where the CrossSampleWorker process picks up the task and calls a Java program for the following:

1. Extracts the text version of the FCS files from the database and stores it in a task folder.
2. The centroid adjustment program written in ‘C’ reads the extracted FCS file and generates detailed and summary results representing cell populations.
3. The result files are processed by the Java program generating overview images.

At the end of the CrossSampleWorker process the task is marked as completed providing the user with access to results which include summary statistics and marker by population two-dimensional displays.

### Cross Sample Analysis: Data set and centroid selection

Flow Cytometry Analysis (Beta) / Cross Sample / Run Analysis

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

Select project: Test research project for v2.9.1

**Choose Centroid File**

Name	Description	Status	Project	Date Created
Cent_16_pop		loaded	Test research project for v2.9.1	2010-09-30
Centroid3		loaded	Test research project for v2.9.1	2010-09-21
Centroid2		loaded	Test research project for v2.9.1	2010-09-21
Centroid1		loaded	Test research project for v2.9.1	2010-09-21

**Choose One or More Analysis Sets**

Name	Description	Status	Project	Date Created
20100923 save all from spw 2	xxxxxx	loaded	Test research project for v2.9.1	2010-09-23
Dataset4	4files 7col	loaded	Test research project for v2.9.1	2010-09-21
Dataset3	save_all_7items	loaded	Test research project for v2.9.1	2010-09-21
Dataset2	4 items 7 col	loaded	Test research project for v2.9.1	2010-09-21
Dataset1	2 items 6 col	loaded	Test research project for v2.9.1	2010-09-21
20100921 SPW set	xxxxx	loaded	Test research project for v2.9.1	2010-09-21

**Continue**

### Cross Sample Analysis: Analysis Setting

Flow Cytometry Analysis (Beta) / Cross Sample / Cross Sample Analysis Setting

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

**Cross Sample**

Project: Test research project for v2.9.1

Name:

Description:

**Marker Selection**

Show Assigned Name | Show Original Name | Help

Type	Name	Marker #	1	2	3	4	5	6	7
Centroid	20101021 test1	7	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
Flow File	Spleen-pLN	7	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
Flow File	Teff pure	7	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
Flow File	CD4 purified	7	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H
Flow File	Treg pure	7	FSC-H	SSC-H	FL1-H	FL2-H	FL3-H	FL2-A	FL4-H

**Cross Sample: Analysis Status**

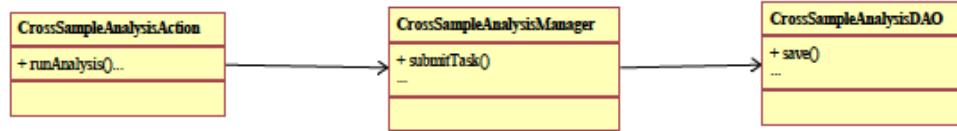
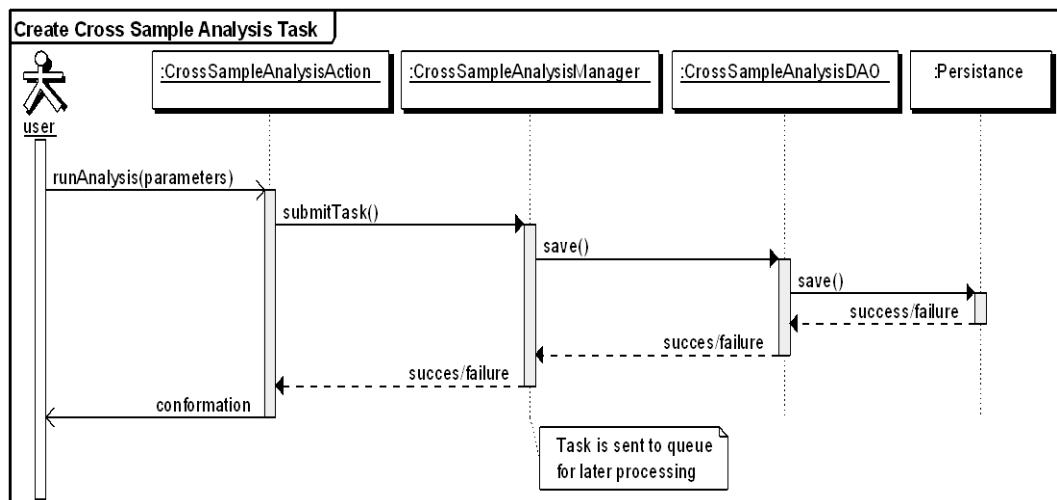
**Flow Cytometry Analysis (Beta) / Cross Sample / Analysis Status**

Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help

Your Cross Sample Analysis request was submitted successfully. Your task ID is **2618**. You will be sent an email when your task has been completed.

Click the "Show Analysis Results" button to proceed to the Cross Sample Analysis Detail page.

**Show Analysis Results**

**Submit Cross Sample Analysis task Class Diagram****Submit Cross Sample Analysis task Sequence Diagram**

## FLOCK Analysis Results

Results of FLOCK analysis tasks are displayed in a tabular grid. Each row in the grid represents one task and contains summary information such as name, description, date completed, etc. Selecting Detail provides greater task detail displayed in two panels

- FLOCK Analysis Task Detail—task ID, name, description, etc.
- FLOCK Analysis Result Files—individual FCS file results with a link to individual results detail

The individual file results contain several panels enabling the user to review both statistics and two dimensional images generated by the analysis.

### FLOCK Analysis Results: tabular view

The screenshot shows a software application window titled "Flow Cytometry Analysis (Beta) / FLOCK / Analysis History". The top navigation bar includes links for Flow Cytometry, Data Management, FLOCK, View/Edit Results, Cross Sample, and Help. A dropdown menu "Select project:" is set to "Test research project for v2.8". Below the header, a message says "Mouse-over column titles, click down arrow to hide or show columns." The main area is titled "FLOCK Analysis History" and displays a table with 8 columns: Details, Name, Description, Task Type, Version, Status, Run By, Start Date, and End Date. The table shows 8 rows of data, with the first row being the header and the last row being the 20th item of 48 total. The "Start Date" and "End Date" columns are sorted in descending order. The "Details" column contains hyperlinks labeled "Detail".

Details	Name	Description	Task Type	Version	Status	Run By	Start Date	End Date
<a href="#">Detail</a>	20100831 test 2		flock	2	completed	Yuen	2010-08-31	2010-08-31
<a href="#">Detail</a>	3 column test		flock	1	completed	Thomson	2010-07-07	2010-07-07
<a href="#">Detail</a>	3 column test		flock	1	completed	Thomson	2010-07-07	2010-07-07
<a href="#">Detail</a>	test TT3454	xxxx	flock	1	completed	Yuen	2010-07-07	2010-07-07
<a href="#">Detail</a>	test edited marker r		flock	1	completed	Thomson	2010-07-07	2010-07-07
<a href="#">Detail</a>	test overview displ		flock	1	completed	Thomson	2010-07-07	2010-07-07
<a href="#">Detail</a>	20100707 TT2073t run run testapp2		flock	1	completed	Yuen	2010-07-07	2010-07-07

**FLOCK Analysis Results: Detail display**

The screenshot shows the ImmPort FLOCK Analysis Result Detail display page. At the top, there is a navigation bar with links for About ImmPort, Admin, Access Data, Tools, Resources, News & Events, Edit Profile, and Sign Out. Below the navigation bar, the title "Flow Cytometry Analysis (Beta) / FLOCK / Analysis Result" is displayed. Underneath the title, there is a breadcrumb navigation: Flow Cytometry | Data Management | FLOCK | View/Edit Results | Cross Sample | Help. The main content area is titled "FLOCK Analysis Task Detail" and displays the following task information:

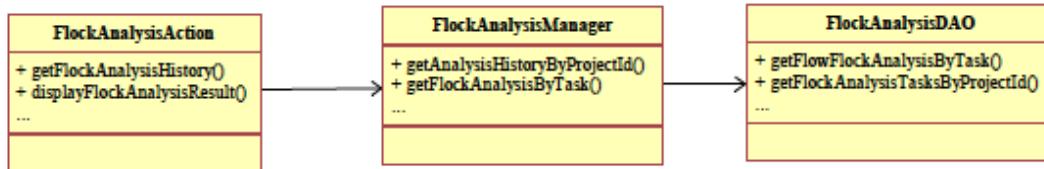
Task ID	1719
Name	test TT3454
Description	xxxx
Algorithm Version	1.0
Run By	Yuen
Start Date	07/07/2010 15:21:01
End Date	07/07/2010 15:21:28
Status	Completed

Below this, there is a section titled "FLOCK Analysis Result Files" containing a table of results:

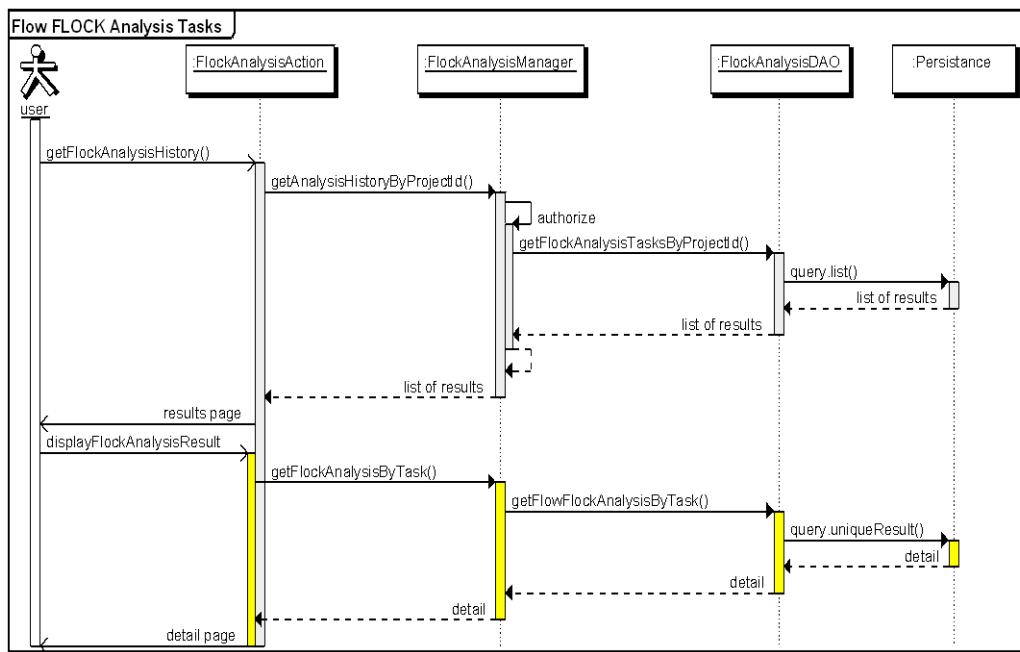
Name	Description	Exp Sample Acc	FCS Text File	Bins	Density	Populations	Status
Sample 001	Multiple file upload example file 1 ES118544		FCS001.txt	52	3	30	<a href="#">Results</a>
Sample 002	Multiple file upload example file 2 ES118545		FCS002.txt	42	3	28	<a href="#">Results</a>
Sample 003	Multiple file upload example file 3 ES118546		FCS003.txt	38	3	27	<a href="#">Results</a>

At the bottom of the table, it says "Displaying 1 - 3 of 3".

**FLOCK Analysis Results Class Diagram**



**FLOCK Analysis Results Sequence Diagram**



### Cross Sample Analysis Task Results

The results of Cross Sample Analysis tasks are displayed in a tabular grid similar to that for FLOCK. Each row in the grid represents one task and contains summary information such as name, description, date completed, etc. Selecting Detail provides greater task detail displayed in three panels:

- Task detail—task ID, name, description, etc.
- Flow Analysis Sets—displays all sets in the analysis
- Cross Sample Analysis Result Files—displays all FCS files in the analysis

Selection of Analysis Results in the task detail panel links to the population proportion table and the marker-by-population two dimensional displays. The Results link within the Cross Sample Analysis Result Files panel links to individual file results.

### Cross Sample Analysis: tabular view

The screenshot shows a software interface titled "Flow Cytometry Analysis (Beta) / Cross Sample / Analysis History". The top navigation bar includes links for "Flow Cytometry", "Data Management", "FLOCK", "View/Edit Results", "Cross Sample", and "Help". A dropdown menu shows "Test research project for v2.9.1". Below the navigation is a search bar labeled "Select project". A note says "Mouse-over column titles, click down arrow to hide or show columns." The main area is titled "Cross Sample Analysis History" and displays a table with the following data:

Details	Name	Description	Task Type	Status	Run By	Start Date	End Date
<a href="#">Detail 1 2598</a>			cross_sample_analys	completed	Thomson	2010-09-30	2010-09-30
<a href="#">Detail 1 2533</a>			cross_sample_analys	completed	Thomson	2010-09-21	2010-09-21

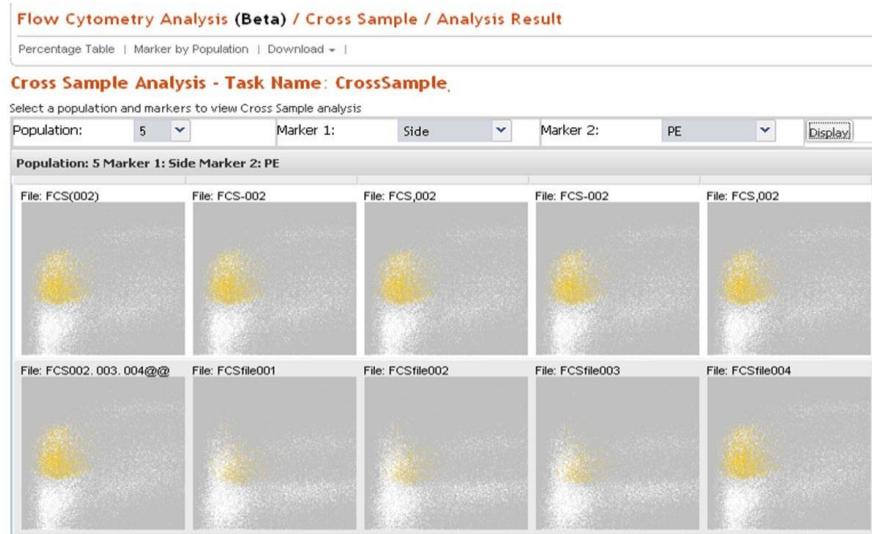
### Cross Sample Analysis: Detail display

The screenshot shows the 'Cross Sample Analysis (Beta) / Cross Sample / Analysis Detail' page. It includes sections for 'Cross Sample Analysis' (Task ID: 2533, Name: CrossSample1, Description: 4 files 7 col, Run By: Thomson, Start Time: 09/21/2010 15:17:33, End Time: 09/21/2010 15:18:13, Status: Analysis Results), 'Centroid Information Detail' (Name: Centroid3), and 'Flow Analysis Sets' (Dataset4: 4files 7col, loaded). Below these are tables for 'Cross Sample Analysis Result Files' (Displaying 1 - 4 of 4) and 'Cross Sample Analysis Result Data' (Displaying 1 - 1 of 1).

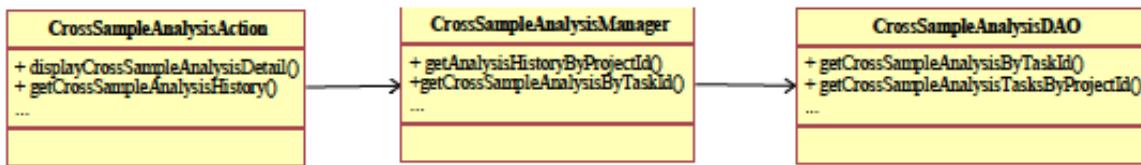
Name	Description	Status
Dataset4	4files 7col	loaded

Name	Description	FCS Text File	Status
CD4 purified		DD082009_012.txt	<a href="#">Results</a>
Spleen-pLN		DD082009_009.txt	<a href="#">Results</a>
Teff pure		DD082009_015.txt	<a href="#">Results</a>
Treg pure		DD082009_017.txt	<a href="#">Results</a>

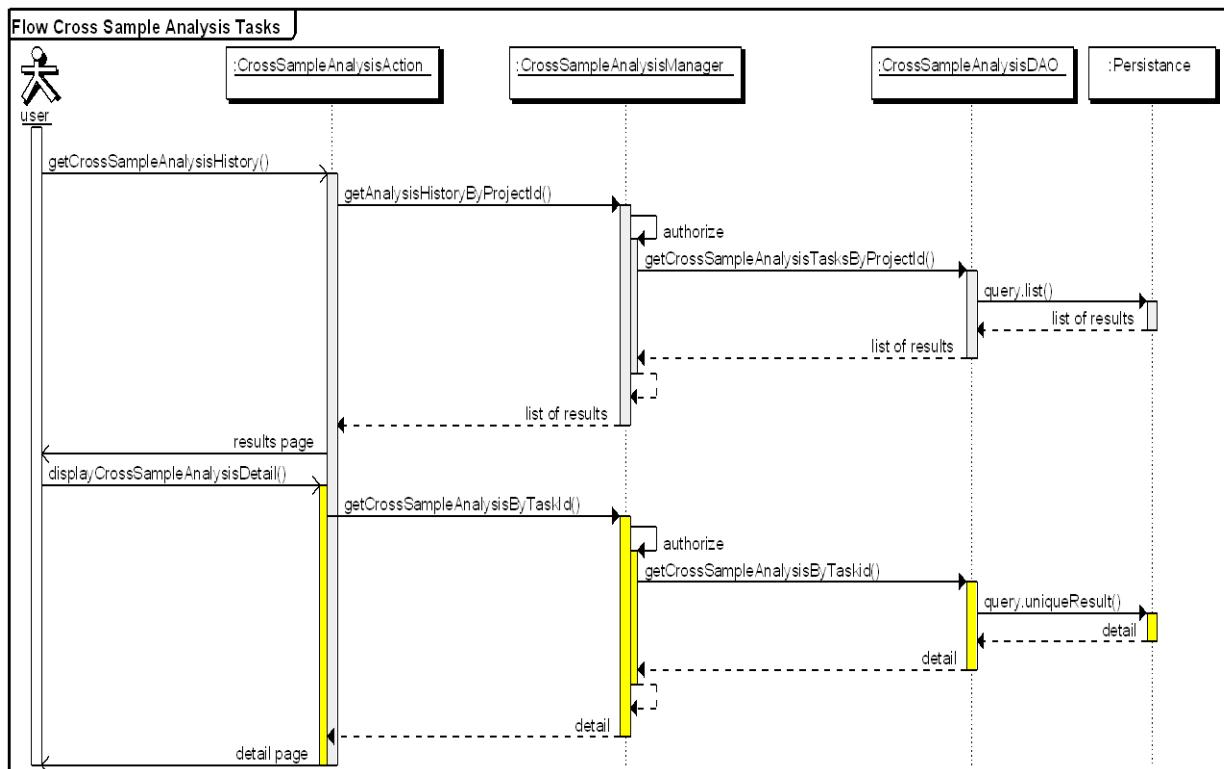
### Cross Sample Analysis: Marker by Marker display



**Cross Sample Analysis Task Results Class Diagram**



**Cross Sample Analysis Task Results Sequence Diagram**



### **FLOCK Results Visualization**

The FLOCK analysis Detail result screen includes 6 displays:

- Overview
- Summary Tables
- Download Tables
- Download Results
- Centroid
- 3D View

Each of the six sections will be reviewed but given the importance of image generation for the Overview and Result Adjustment sections a brief image generation overview is required.

FLOCK analysis produces a tab-delimited text file where rows are events and columns are markers—an exception is the last column which represents the population assigned to the event. The output file is transformed into binary format to facilitate real time image generation. UI requests for images results in a JAVA module reading the binary file, computing the coordinates and drawing dot-plot images in PNG format. Images are written to the file system under the analysis task folder where they are retrieved by the UI.

FLOCK results Overview displays a table of images where each cell represents the dot-plot of the populations assigned based on the expression values for two markers. On the right hand side of the screen there is a panel for selecting which populations to display, and summary information for each population. The user can select/deselect populations using the check box and then click on the Update button to change the display of the images to only include the selected populations. Each image may be selected—this action takes the user to the Result Adjustment screen.



### Result Adjustment

Result Adjustment supports editing results as a means to fine-tune the centroid locations. Images will be regenerated when Save Changes is selected. In addition to moving a centroid the system supports splitting and deleting centroids.



### Summary Tables

Three summary tables are available for each FLOCK analysis and include:

- Results—populations appear as table rows, channels/markers as columns with the population proportion presented in the last column
- Centroid—each cell of this table represents the center for the population/marker combination
- Mean Fluorescent Intensity (MFI)—each cell represents the mean fluorescence value.

### Download Tables

The content of the download tables is the same content that is displayed in the Summary Tables mentioned above. In this case the content can be downloaded in text file format for further processing.

### Download Results

Result files available for download include:

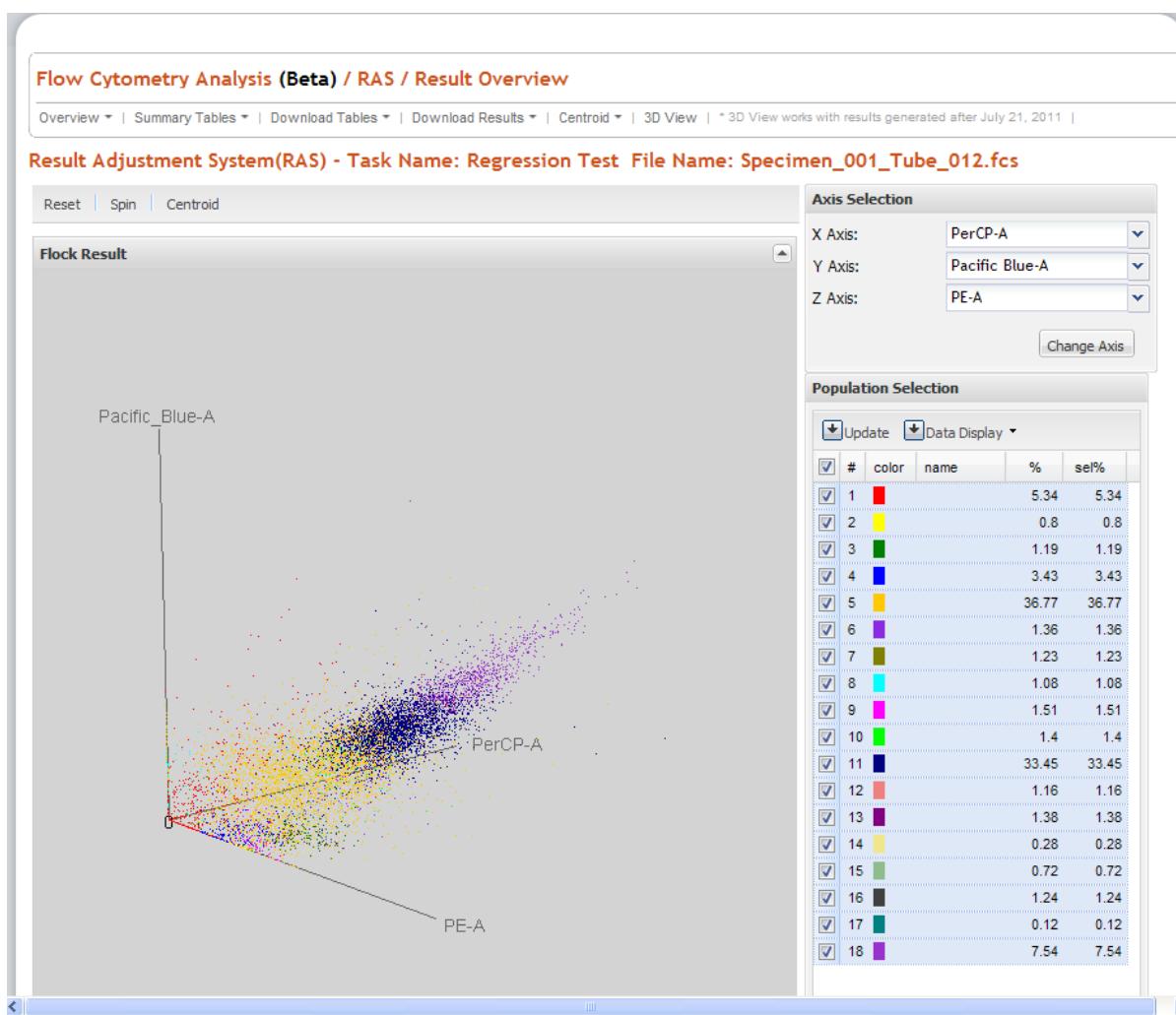
- Results table including event number and population assignment
- A package of files representing all the result files.

### **Centroid**

From this panel the user can select to save the results editing from Result Adjustment or undo any previous adjustments.

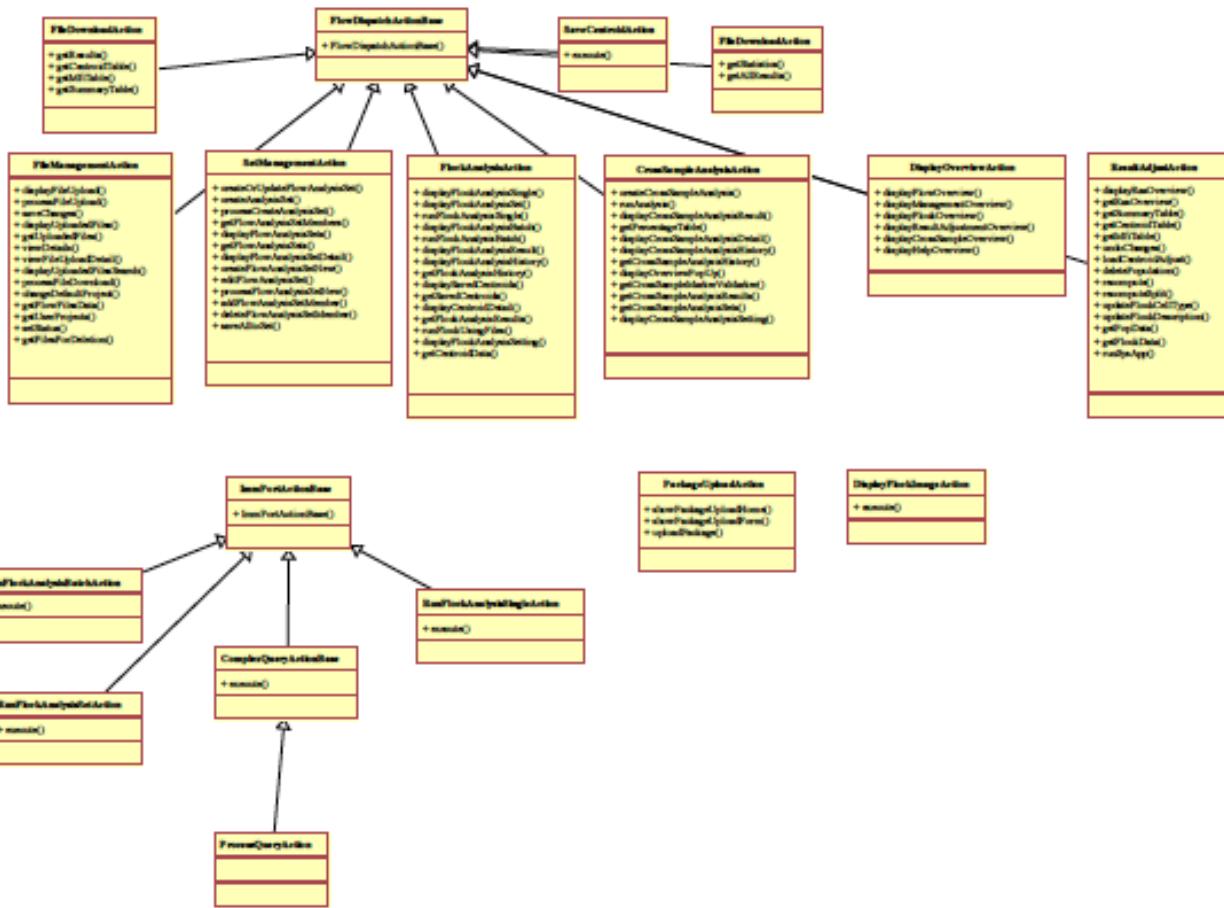
### **3D View**

In this view the results for 3 of the markers can be displayed in a 3D view using the JMOL applet. On the right side panel, the user can select which markers to display and the user can choose which populations to display.

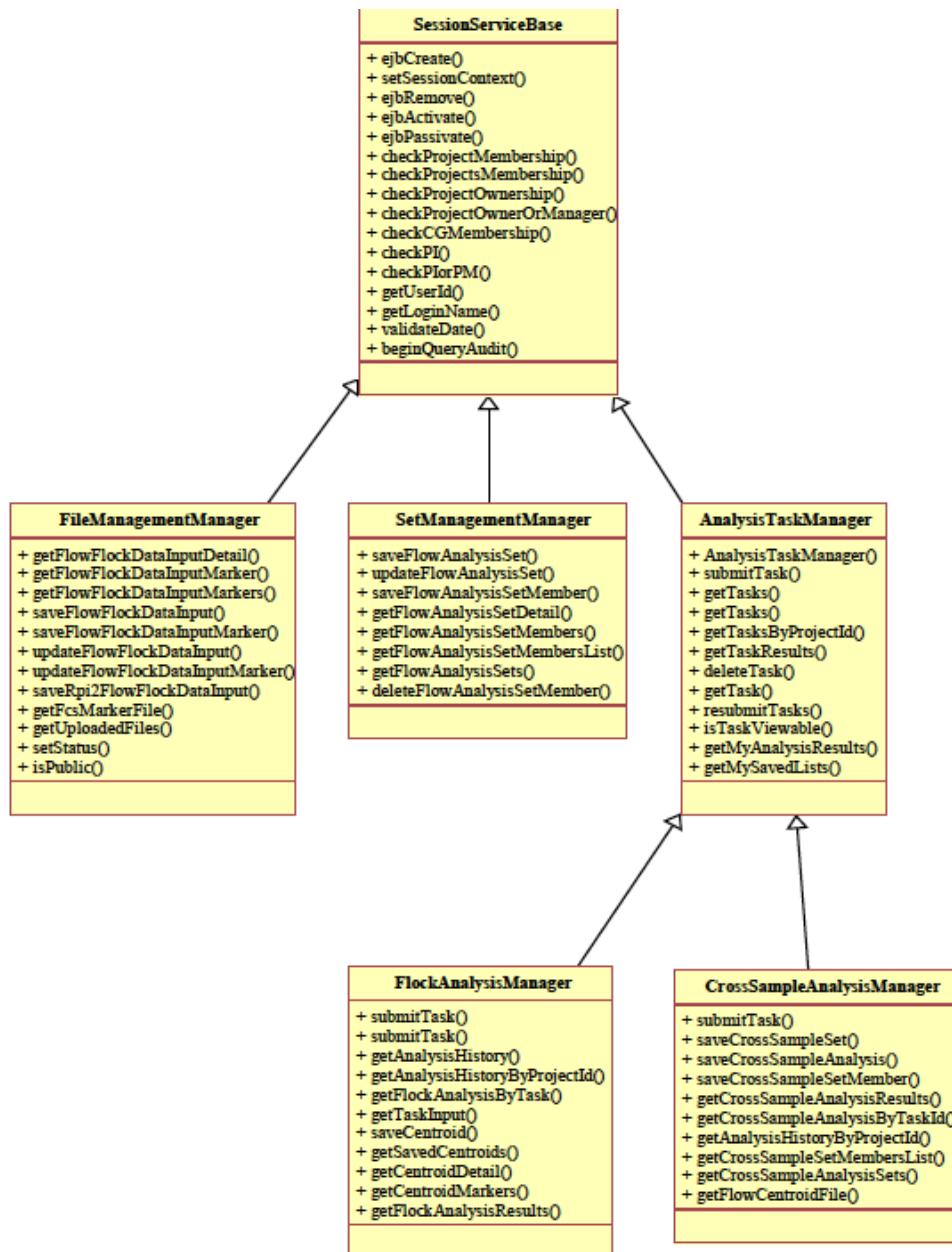


### Flow Analysis Class Diagrams

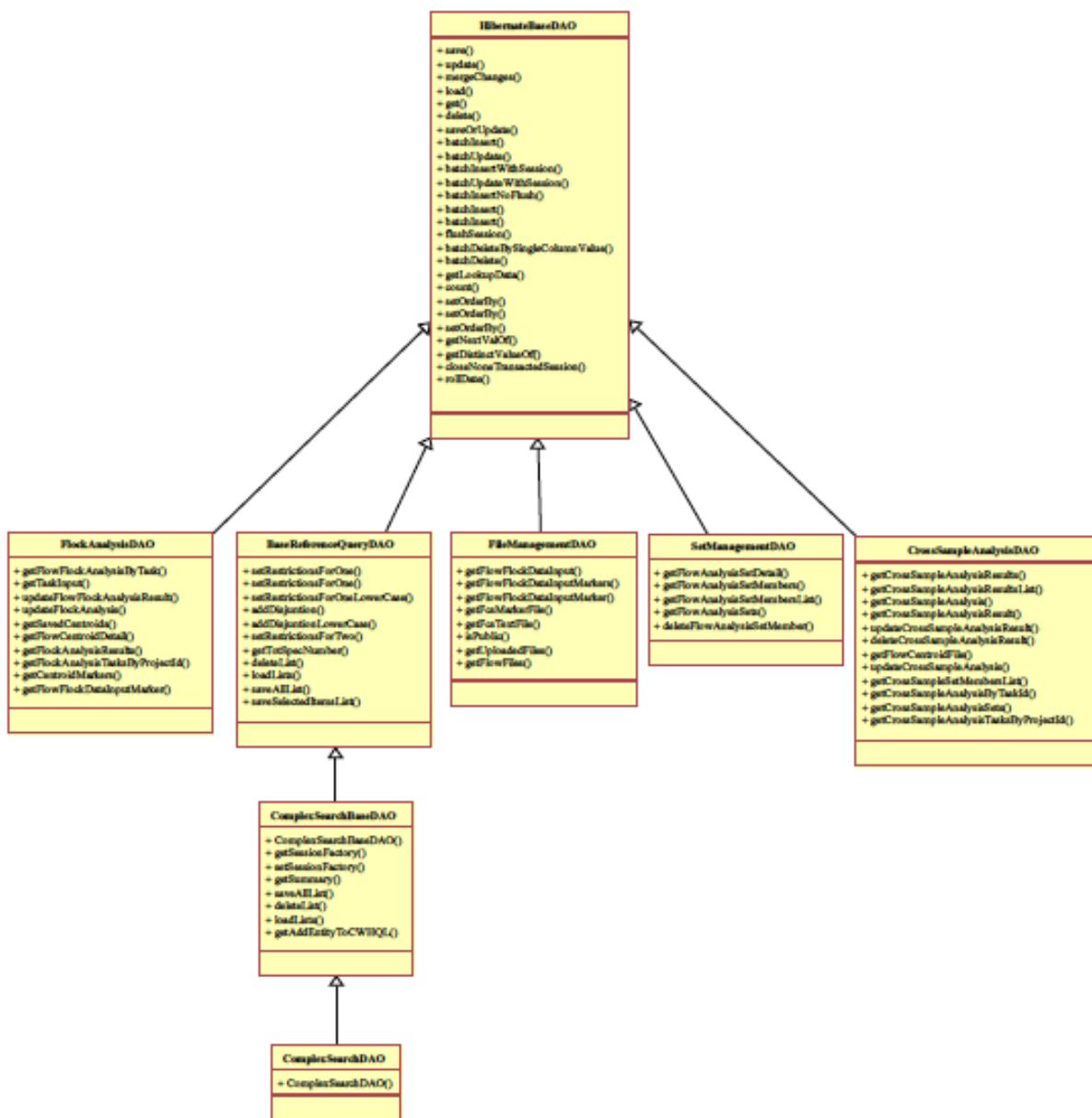
#### Actions



**EJB**



## DAO



### 3.6.5 Data Management

The search module was designed to allow the user to query for summary information about objects by entering ad-hoc queries, based on the properties of each class of object, plus associations between objects. The module was used to provide support for both advanced search and basic search UI's. Examples of the types of objects represented in ImmPort are: subjects, experiments, experimental results, studies, genes, proteins, etc.

During the design process we had several goals and constraints that are outlined below:

- Develop the system to be easily maintained.
- Develop the system to be extensible to other objects within the ImmPort schema.
- Reuse the tabular result page design already developed.
- Reuse the Hibernate object model.
- Make much of the system configurable or meta-data driven.
- Support the use of previously saved query results (lists) in the complex queries.
- Keep the interfaces between the layers clean and simple, allowing reuse in other parts of the application.

Below is a screen shot of the Advance Search module for Research Queries. The screen has 4 major components that will be discussed on more detail later in the document.

- Toolbar.
- Query Tree Panel.
- Criteria Panel.
- Summary Results Panel

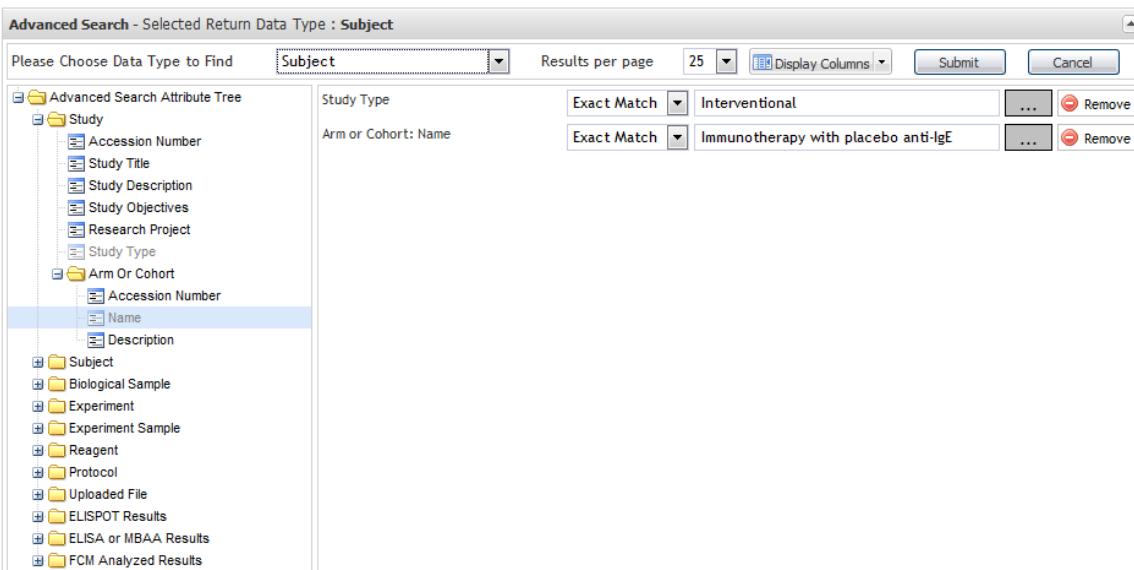
This query and UI framework has been implemented for basic and advanced research search interfaces and for advanced reference query in the same manner. It is also used in selected queries across the system when it is useful to return a subset of objects based on a criteria specific to a given page. This framework re-use allows for greater maintainability and expedites the addition of new search interfaces to the system.

### 3.6.5.1 General Overview

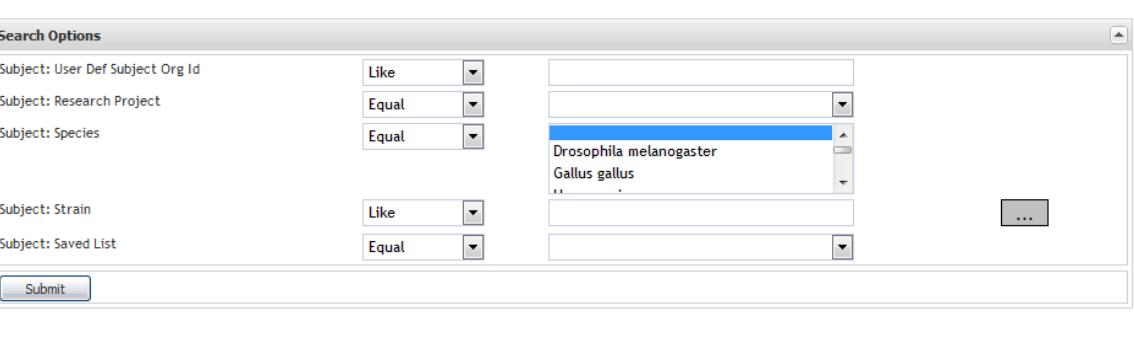
Multiple search screens provide the entry point into the search framework described, but all utilize the same presentation layer technology. The Advanced Search research page appears with a tree view with available search attributes on the left side, the selected search attributes on the right. The user is prompted to return the type of object that they wish to return at the top.

**Using Advanced Search:**

- ▶ Choose the type of data you want to find (e.g. subjects, samples) from the drop down list.
- ▶ Choose a folder to refine your search. It can be the same or different data type from the data you want to find.
- ▶ Click on the data feature(s) you want to search.
- ▶ Enter search terms for one or more features.



Basic Search screens have a fixed and minimal set of attributes with which to query for a fixed object to return, as shown for the Subject Search page.



After a search is submitted, the attribute panels disappear and the query results summary grid appears. This result summary page provides extensive functionality, including:

1. Sorting by each column
2. Re-ordering of columns by drag and drop
3. Hiding and Un-hiding columns
4. Page through result records
5. Selection of individual object records to view further details
6. Save Items or Save All Item rows to a list for later re-use
7. Export of the summary query results to Excel or PDF

Viewing of detail records and list management and use will be discussed in later sections.

#### Using Advanced Search:

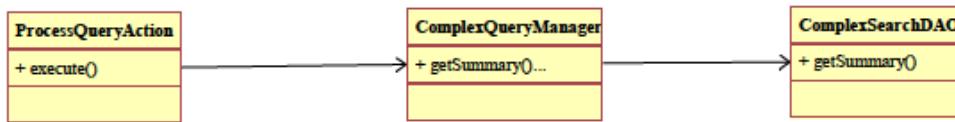
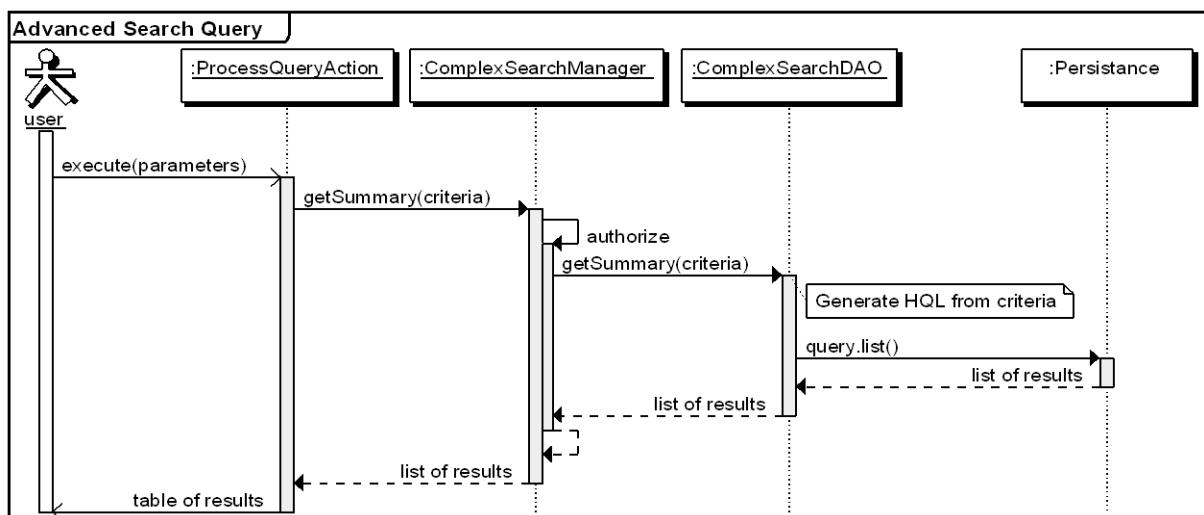
- ▶ Choose the type of data you want to find (e.g. subjects, samples) from the drop down list.
- ▶ Choose a folder to refine your search. It can be the same or different data type from the data you want to find.
- ▶ Click on the data feature(s) you want to search.
- ▶ Enter search terms for one or more features.

To modify your search, please expand the search box ~

Advanced Search - Selected Return Data Type : Subject

Tip: You can select a range of rows by holding down the 'Shift' key and then clicking the first and last rows you wish to select. This works best if you do not click on the check boxes. [More tips.](#)

Sub Org Accession	Species	Race/Stra	Gender	Description	Subject Phenot	Project Title	Study Title	Arm
SUB73370	Homo sapiens	White	Male	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy Co-administered with Omalizumab	Immunotherapy with placebo anti-IgE
SUB73373	Homo sapiens	Asian	Male	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy Co-administered with Omalizumab	Immunotherapy with placebo anti-IgE
SUB73377	Homo sapiens	White	Female	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy Co-administered with Omalizumab	Immunotherapy with placebo anti-IgE
SUB73378	Homo sapiens	White	Female	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy Co-administered with Omalizumab	Immunotherapy with placebo anti-IgE
SUB73384	Homo sapiens	White	Male	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy Co-administered with Omalizumab	Immunotherapy with placebo anti-IgE
SUB73389	Homo sapiens	White	Male	Placebo + Ragweed IT	Ragweed-indu	ITN019AD - Clinical Trial	Allergen immunotherapy	Immunotherapy with placebo anti-IgE

**Class Diagram****Sequence Diagram****UI generation**

The query tree and form attributes are defined in a metadata file using JavaScript Object Notation (JSON). JSON is an emerging standard used to configure JavaScript web components, and used as the data interchange format between components. The metadata file is input into a JavaScript query tree engine, which generates the tree view in the UI. When a property is chosen from the query tree, the node is passed to a JavaScript component to generate the form input field.

**Form processing**

When the form is submitted, an AJAX call is made to a Struts Action sending the form parameters entered by the user. The Struts Action class generates a complex query criteria object that is passed to the DAO layer to construct the query. The complex query constructor generates a Hibernate HQL query that submits a SQL query to the Oracle database.

**Results processing**

The results of the query are packaged into a JSON object and passed back the UI. The results are then displayed using a grid control.

### 3.6.5.2 Presentation Layer

#### ***Technologies***

For the presentation layer we use a mix of JSP, JavaScript, Ajax and metadata files to present the query interface to users. By using metadata files and generic code, we are able to present alternative versions of the query interface, without the need for additional coding. We have chosen the EXT-JS JavaScript libraries, to build the query tree shown in the Advance Search screens, and to build the summary table grids.

#### ***Query UI***

The property query tree on the left panel of the page is built using a combination of JavaScript and metadata file, which represents the entities and the properties that can be queried by the user. As the users clicks on query properties from the tree, an input field is displayed in the right side panel. The input field can be a simple input text box, drop down list for controlled vocabulary terms, or the user can click on a link to display a popup that lists all terms previously entered for this property.

#### ***Query Results***

The results of the query are displayed in a tabular grid, using EXT-JS components. The user has several options:

1. Sort results based on column selection.
2. Selecting rows, for detail view.
3. Selecting rows to save in a result set.
4. Export the results to a file, that can be downloaded

### 3.6.5.3 Form Processing, Business Object Layer

Because the number of different properties that can be chosen by the user to qualify the query is subject to change, we have developed a generic methodology for processing the parameters and constructing the SQL to be submitted. In the Action class, based on the parameters supplied, a ComplexSearchCriteria will be constructed. The ComplexSearchCriteria object has properties to contain the project Id's and the Name of the primary object of interest (Subject, Experiment, etc). In addition, it contains a collection of items that represents the properties and values passed in by the form submission. Each item will contain the value entered for this property, the name of the property and the search operator used to qualify the join to this property. The field name contains information representing the object and the property of the object that is being queried. A query operator can be “Exact Match”, “Like”, “Greater Than”, etc. For example, if the user is looking for Subjects that are “Male” and have Biological\_Samples where the sample\_type\_name is “Tissue”, the action would be passed the following parameters:

StartingObject: Soi  
SoiGenderValue: Male

SoiGenderOperator: “Exact Match”  
BsSampleTypeNameValue: “Tissue”  
BsSampleTypeNameOperator: “Exact Match”

### **3.6.5.4 Query Construction**

The query constructor engine uses metadata that represents the paths between objects and information from the Hibernate mapping files to construct a query similar to the one in the section below. The Hibernate criteria classes are utilized to dynamically build the Hibernate Query language where clauses based on the property, value, and operator passed from the UI through form submission. The names of the properties in the UI tier are named to match the object properties in the Hibernate mapping files to simplify this construction.

#### ***Hibernate Query***

An example of the resulting Hibernate query is show below.

```
select distinct soi.id
  from SubjectOrgInfo soi
inner join soi.biologicalSampleSet bs
  where soi.researchProjectInfo.id in (:projectId)
    and soi.gender = 'Male'
    and bs.bsType = 'Tissue'
```

#### ***SQL Query***

An example of the resulting SQL query generated by Hibernate from the Hibernate query is shown below:

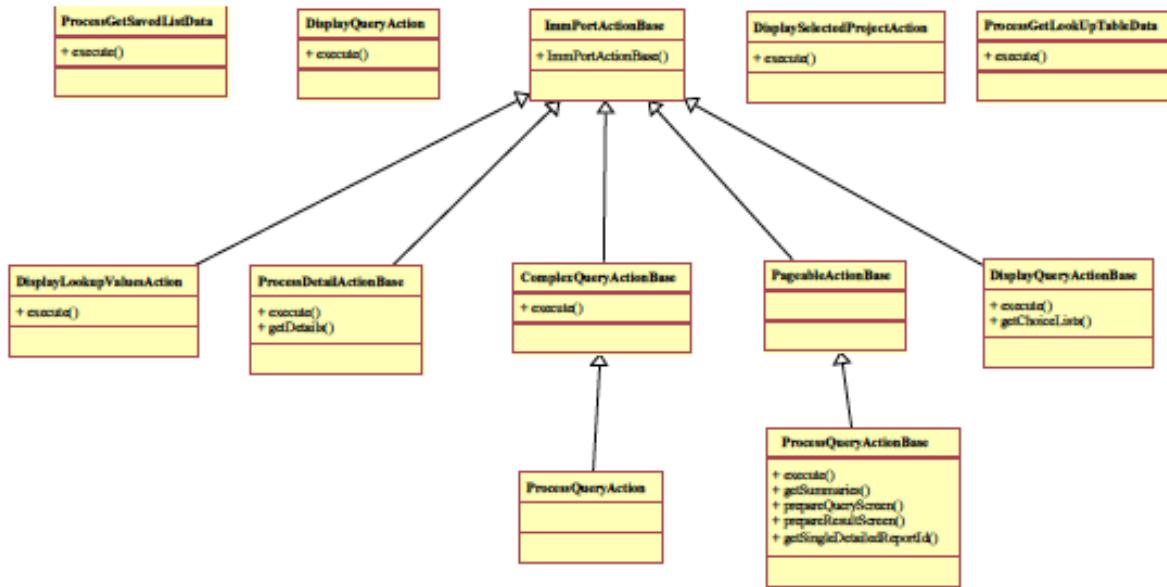
```
select * from
  ( select distinct subjectorg0_.SUBJECT_ORG_ACC_NUM as col_0_0_
    from SUBJECT_ORG_INFO subjectorg0_
   inner join BIOLOGICAL_SAMPLE biological1_ on
     subjectorg0_.SUBJECT_ORG_ACC_NUM=biological1_.SUBJECT_ORG_ACC_NUM
   where (subjectorg0_.PROJECT_ID in (?))
     and (subjectorg0_.GENDER) = 'Male'
     and (biological1_.SAMPLE_TYPE_NAME = 'Tissue') )
where rownum <= ?
```

### **3.6.5.5 Query Processing – Results Display**

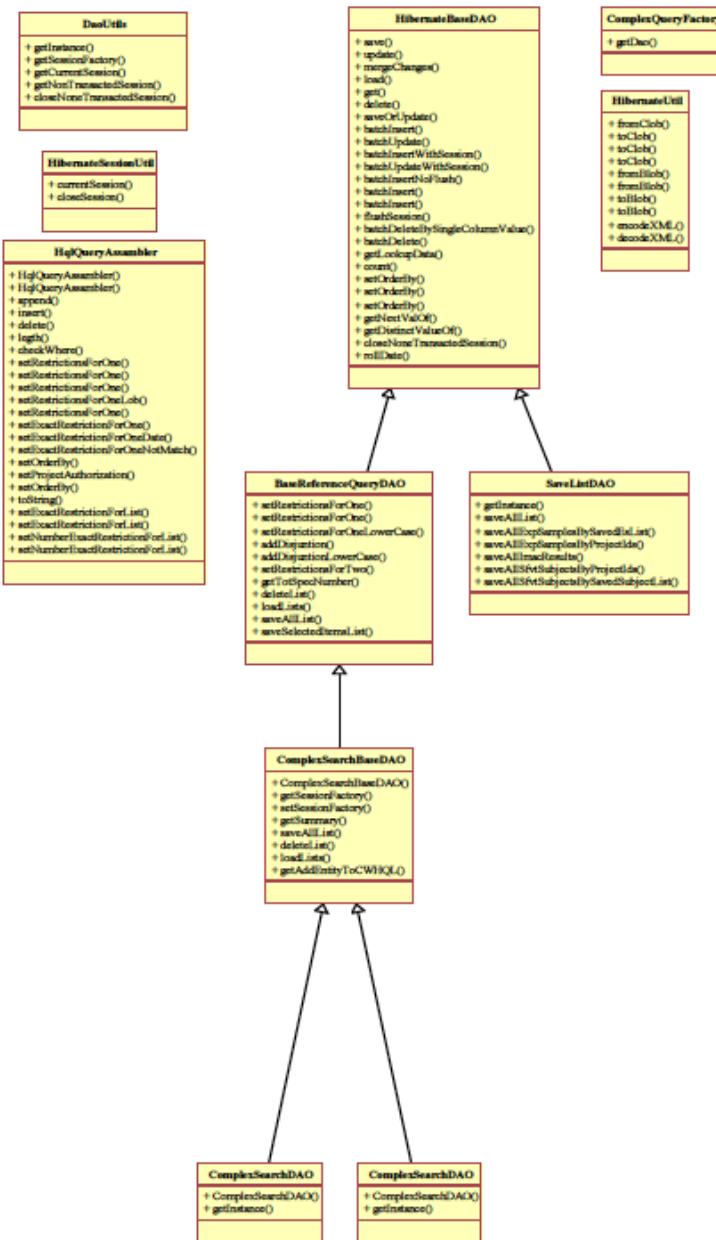
The SQL query generated by query constructor is then submitted to the Oracle database. The results are processed and packaged into a JSON object. The JSON object is then passed back to the AJAX call that initiated the form submission, and displayed using the Table Grid control.

### **3.6.5.6 Search Class Diagrams**

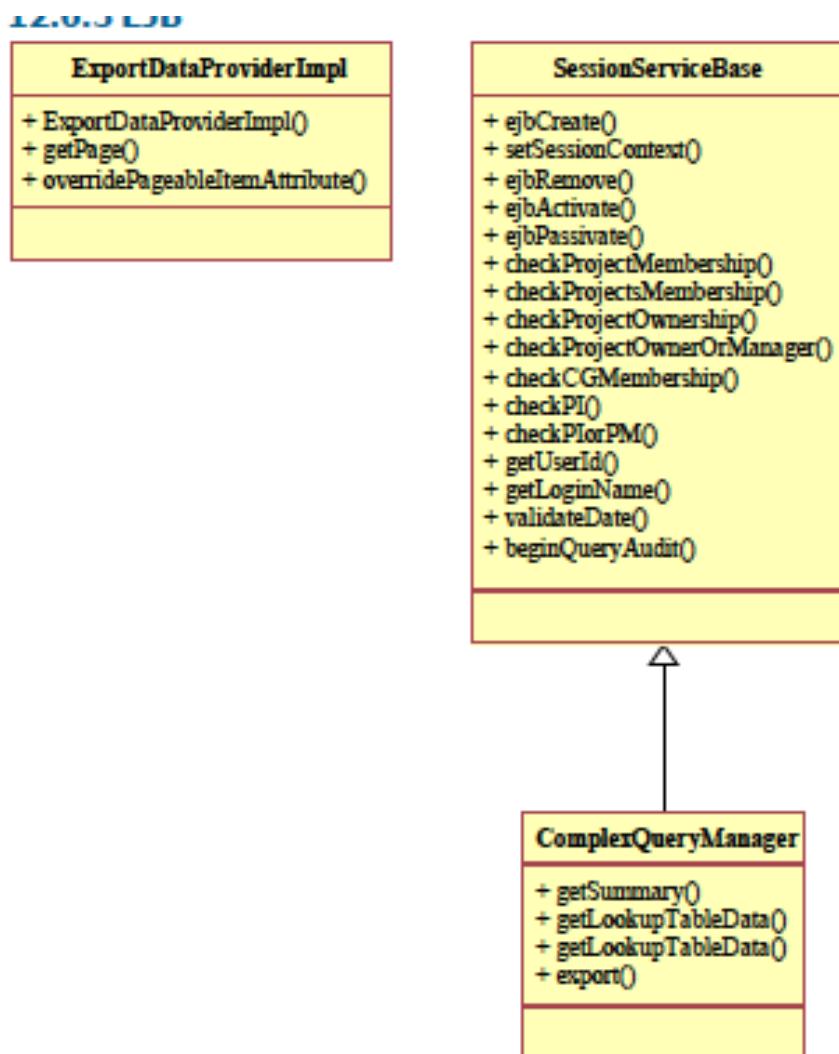
#### ***Action – Presentation Layer***



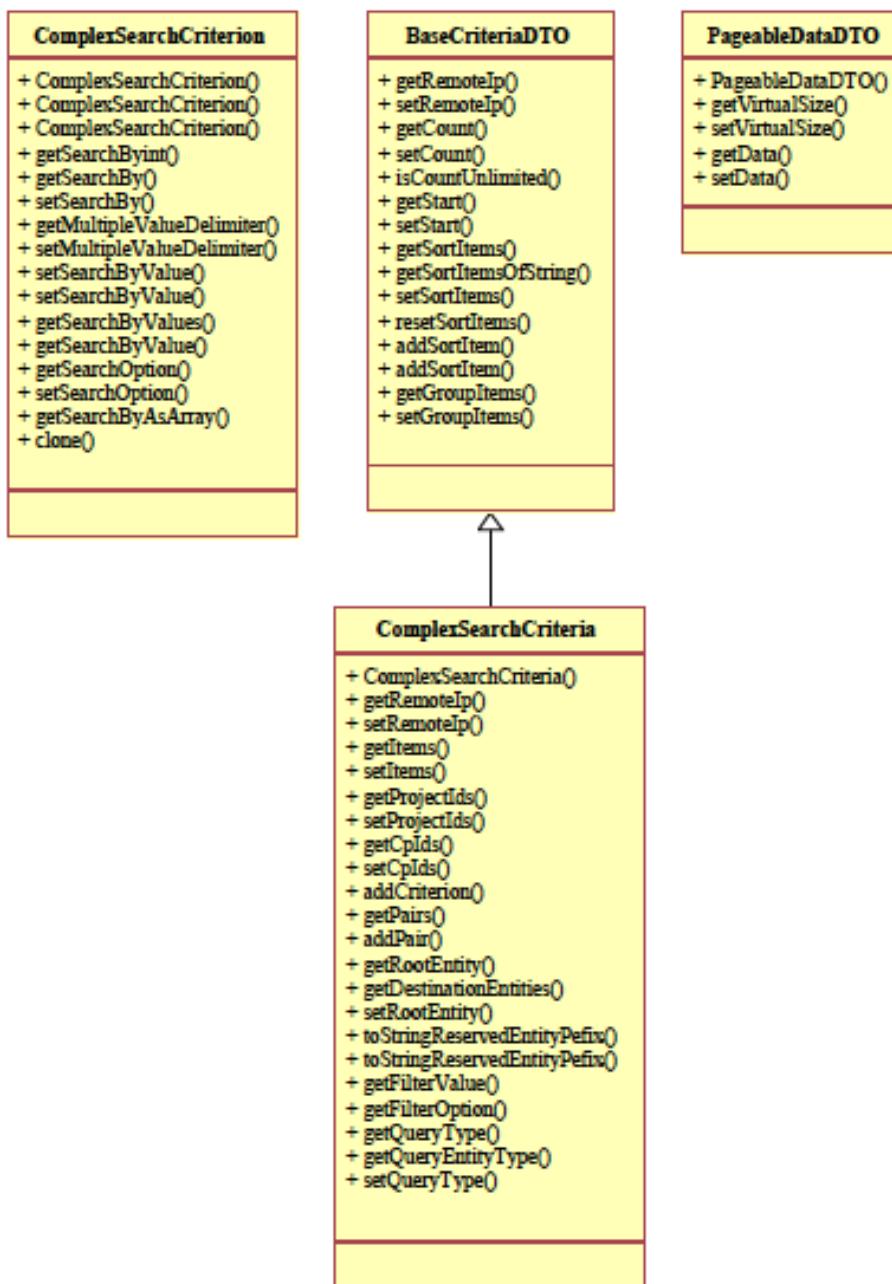
## DAO



EJB



## DTO



### **3.6.5.7 Research Data Detail Page Design Packages**

The detail screens for the research entities, with the exception of the study detail screen, all follow a similar pattern and use the same technology stack. The top panel on the detail screen shows the detailed information for the entity. Below this panel there are a series of panels that can be expanded or collapsed and are loaded on demand. These panels re-use the same technology used to display the summary grids in the advanced search module described previously.

The screenshot displays a research data detail page with the following sections:

- Subject Summary:** Contains fields for User-Defined ID (ITN019AD\_001\_001), Study / ARM / Study PI (Allergen immunotherapy Co-administered with Omalizumab), ARM Name (Immunotherapy with anti-IgE), and Study PI (Thomas Casale). It also lists various subject details like Organism (Homo sapiens), Taxonomy ID (9606), Enrollment Age (47), Gender (Male), Religion, Race (Asian), Ethnicity (Not Hispanic or Latino), Population, Affection Status (0), Subject Phenotype (Ragweed-induced seasonal allergic rhinitis), and Project Title (ITN019AD - Clinical Trial shared to Semi-Public Workspace (SPW) Project).
- Treatments:** A collapsed section.
- Protocols:** A collapsed section.
- Biological Samples:** A collapsed section.
- Assessments:** A collapsed section.
- Clinical Assessments:** An expanded section showing a grid of results. The grid has columns for Subject Acc Nu, Name, Component Name, Clinically Signif, Study Day, Time Of Day, Result, Value, Unit, and Organ. Two rows are visible:
  - SUB73366, Vital Signs, Baseline Weight, -21, 69.4008 KG, 69.4008, KG
  - SUB73366, Vital Signs, Body Mass Index, -21, 23.9633 KG/M<sup>2</sup>, 23.9633, KG/M<sup>2</sup>

Many of the panels in the study detail screen represent aggregated information and cannot be easily satisfied with a simple SQL query. Therefore many of these panels require the aggregation of results to construct in the business layer.

The user may request to view one or more detail pages for a specific entity. For this discussion we will assume the user requested to see the details for 2 subjects: SUB1 and SUB2. This request starts by returning a screen containing “tabs” across the top, where tab 1 one represents SUB1 and tab 2 represents SUB2. As this page is displayed an AJAX call is submitted to an Action

class requesting the detailed information for the first tab. The action class calls an ejb class, which calls a dao class. The results are then merged with the JSP page, and returned to the AJAX call as HTML. This HTML is then placed in the DOM and displayed. The remainder of the page displays collapsed panels that represent entities that can be associated to SUB1. The contents of these panels are not available at this time. To retrieve the content for a panel the user clicks on the expand button, which initiates an AJAX call to retrieve a summary grid representing the associated entities. This AJAX call uses the same technology as outlined in the advanced search portion of this document, to return the summary grid.

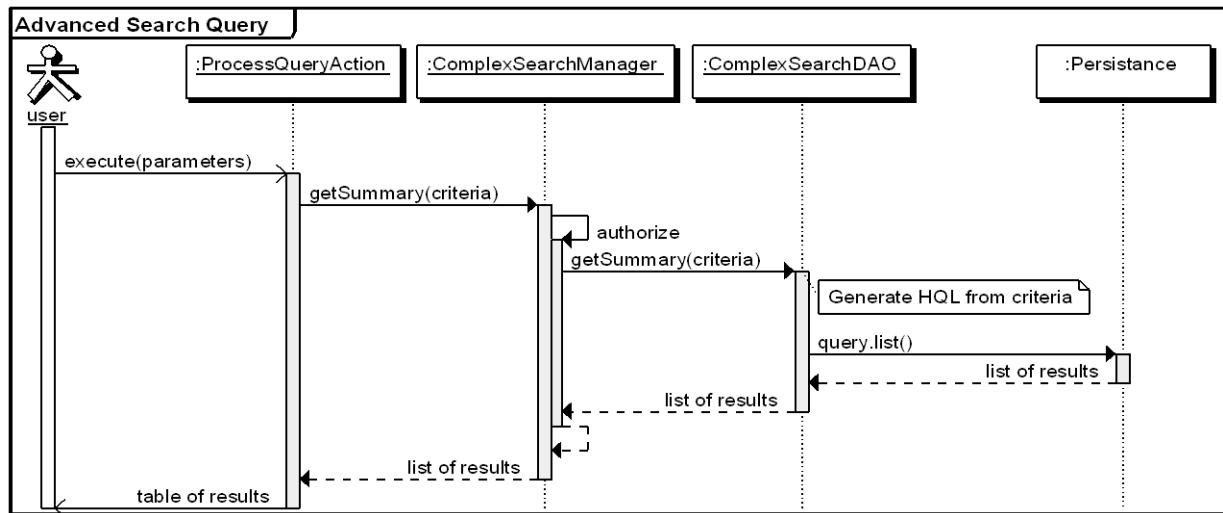
If the user clicks on tab 2, then the whole process described above repeats. If the user then clicks on tab 1 again, results are shown immediately since this content was cached in the DOM.

The design package that follows utilizes the subject object detail page as an illustrative example of how all of the research data detail pages are implemented given that the same search and UI framework is used in all pages. The final section summarizes the detail page entities and the panels that are implemented in each one.

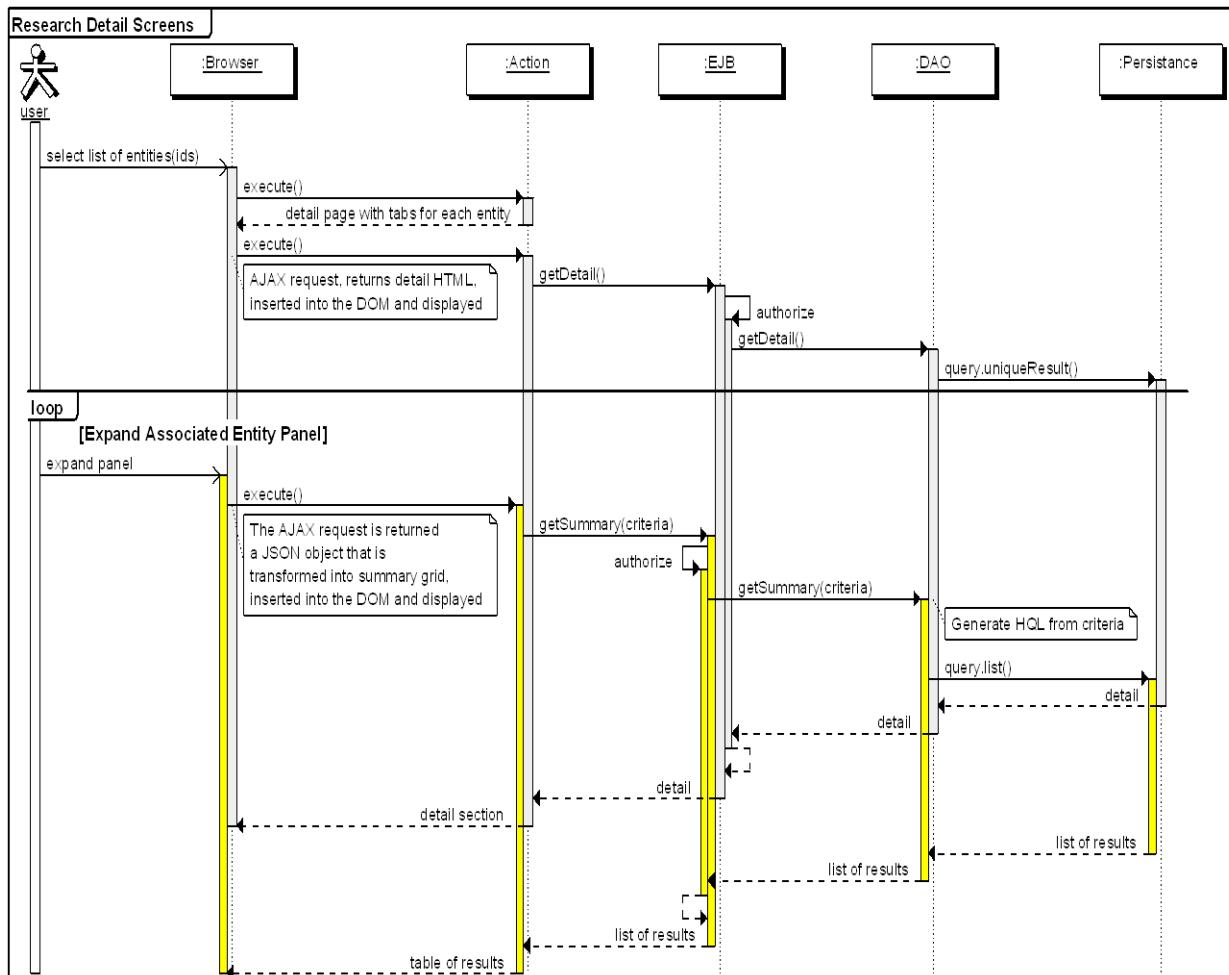


**Class Diagram – Subject Panels**

The classes used to construct the panel grids are the same as the classes used to construct the advanced search grids.



**Sequence Diagram**



**Table of Entities and Panels**

Entity	Panel
Subject	Subject Detail
	Treatments
	Protocols
	Biological Samples
	Assessments
	Lab Tests
	Adverse Events
	Concomitant Medications
Biological Sample	Biological Sample Detail
	Treatments
	Protocols
	Associated Experiments
	Biological samples derived from this Biological Sample
	Clinical Lab Tests
Experiment Sample	Experiment Sample Detail
	Result Files
	Protocols
	Biological Samples
	Reagents
Experiment	Experiment Detail
	Experiment Samples associated with this Experiment
	Protocols
Protocol	Protocol Detail
	Experiments
	BioSamples
	Experiment Samples
	Subjects
	Studies
Reagent	Reagent Detail
	HLA Typing System
	Analytes
	Reagent Sets
Study	Study Detail
	Inclusion Exclusion Criteria
	Demographics
	Assessments
	Concomitant Medications
	Treatments
	Adverse Events
	Clinical Lab Tests
	Mechanistic Assays

	Documentation
	Glossary

### 3.6.5.8 Research Data Management Design Packages

Based on the users permissions they may have the ability to edit the research entities by either updating the primary information of an entity or by creating or removing a relationship to another entity. Because editing of all research entities works in a similar fashion, we will use the editing of the subject entity to illustrate the process.

#### ***Displaying the Edit Screen***

From the subject detail page, the “Edit Subject Information” link is displayed.

The screenshot shows a software interface for managing research subjects. At the top, there are three tabs labeled 'SUB1461', 'SUB1463', and 'SUB1462'. Below the tabs, a blue header bar contains the text 'Subject Summary' and a link 'Edit Subject Information'. To the right of the header are buttons for 'Collapse All:' and 'Expand All:'. The main area displays a table with the following data:

Subject-Organism Accession No:	SUB1461
Description:	Lupus erythematosus with nephritis, ICD-9 710.5
Organism:	Homo sapiens
Taxonomy ID:	9606
Enrollment Age:	40
Gender:	Female

After choosing to update a subject the edit subject detail screen will be displayed. In the top section of the screen the previously entered values will appear in input text boxes or in drop down lists. The user makes the necessary changes and then clicks the Save Changes button.

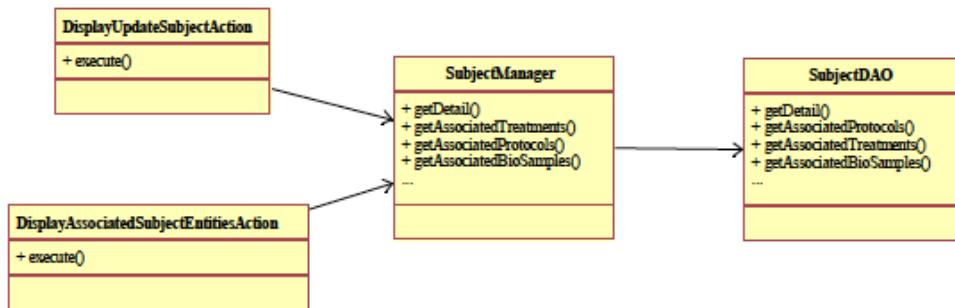
Fields marked with an asterisk \* are required.

Subject Accession No:*	SUB1461
User-Defined ID:*	ELISA Subject 1
Description:	Lupus erythematosus with nephritis, ICD-9 710.5
Organism:*	Homo sapiens
Taxonomy ID:	9606
Enrollment Age:	40
Strain:	
Strain Characteristics:	
Gender:	Female
Religion:	
Race:	Caucasian
Ethnicity:	Non-Hispanic
Population:	
Affection Status:	1 - Unaffected or Control
Subject Phenotype:	
Subject Pedigree ID:	
Family Pedigree ID:	
Mother Pedigree ID:	
Father Pedigree ID:	

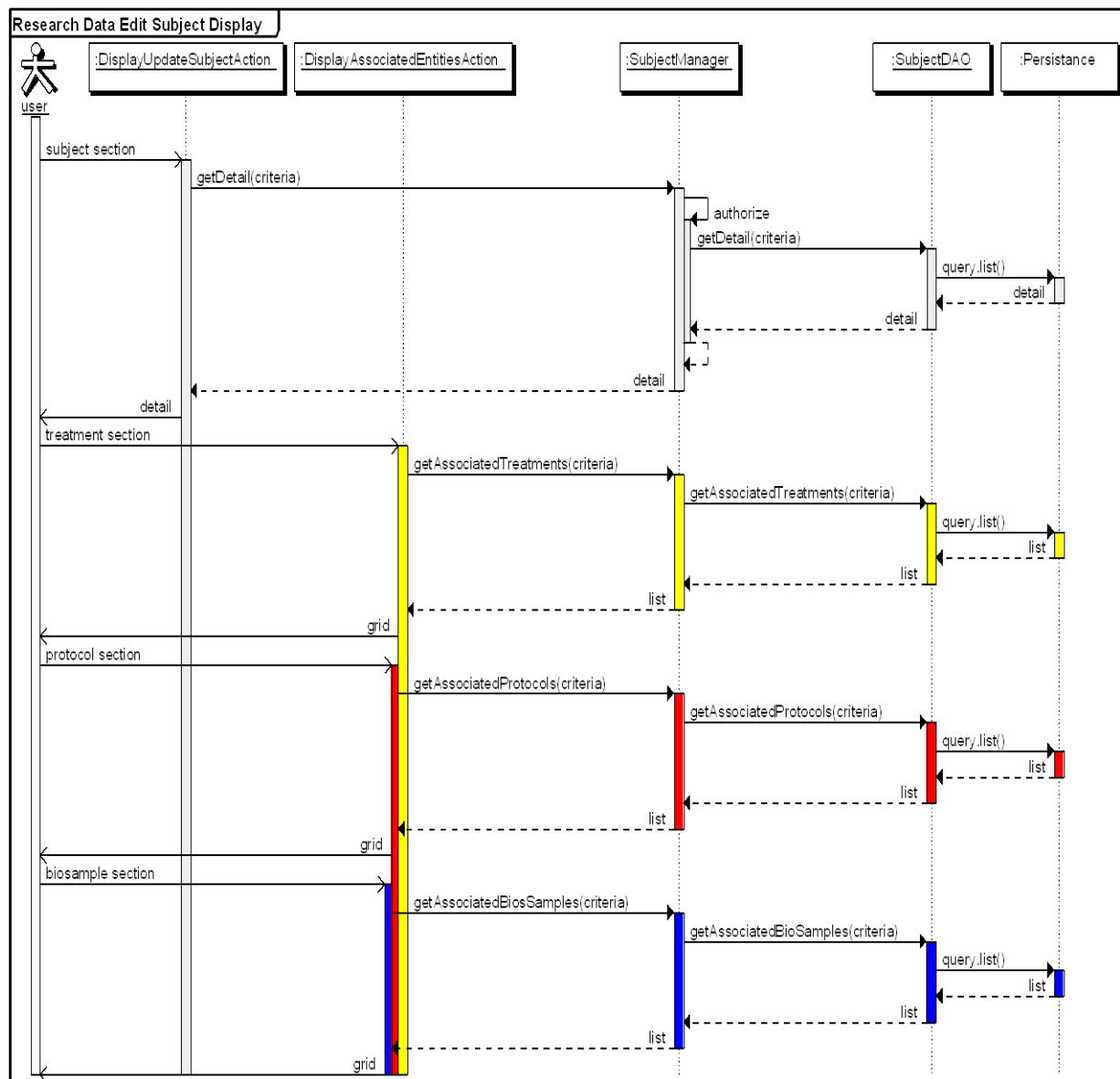
**Treatments**

Page 1 of 1								No items found		
Accession Number	Agent	Weight	Concentration	Volume	Time	Temp	Other	Add	Edit	Delete
None										

### Display Edit Screen Class Diagram



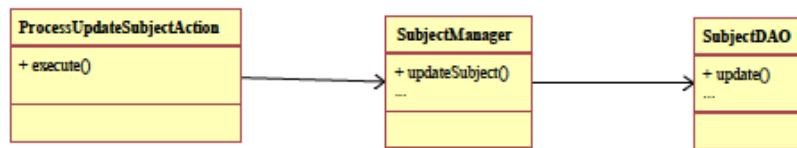
**Display Edit Screen Sequence Diagram**



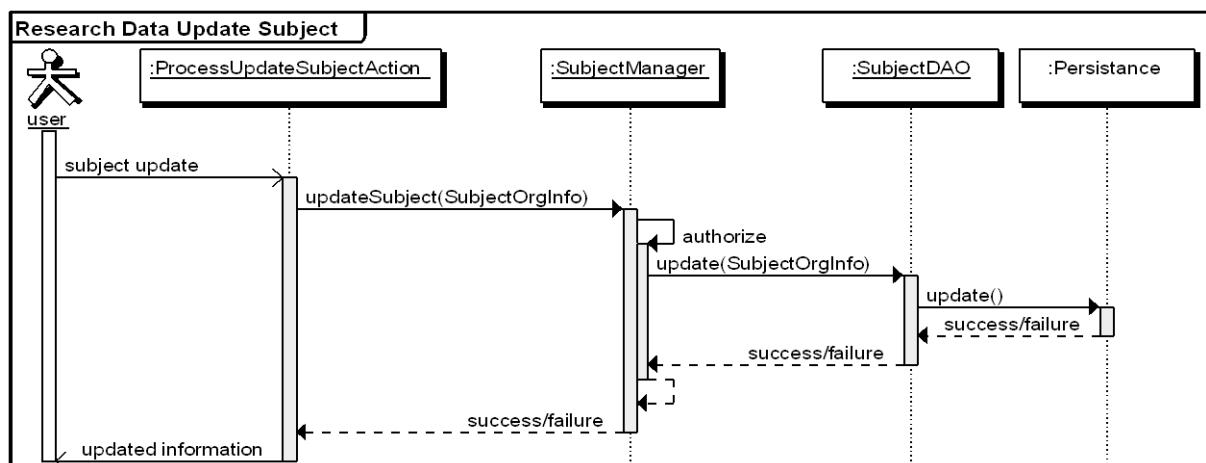
### 3.6.5.9 Update Subject

After the save changes button is clicked, the transaction is sent to the database to update the appropriate fields.

*Update Subject Class Diagram*

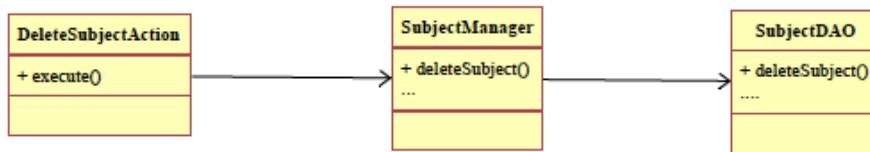
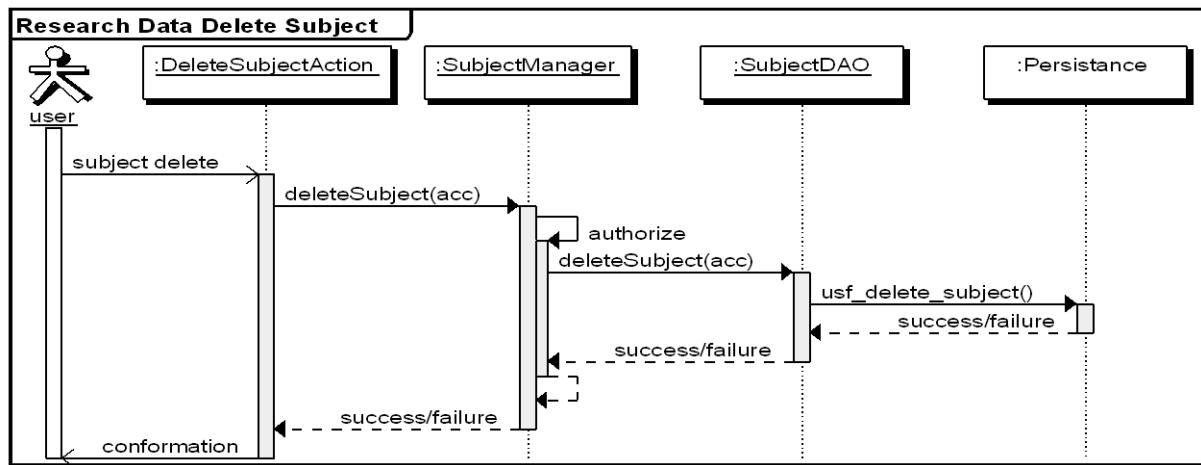


*Update Subject Sequence Diagram*



**Delete Subject**

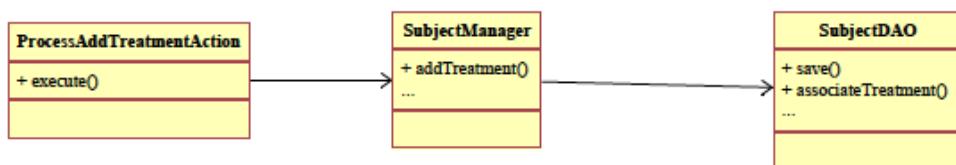
Additionally, if the user has appropriate permissions he may delete the subject from the detail page.

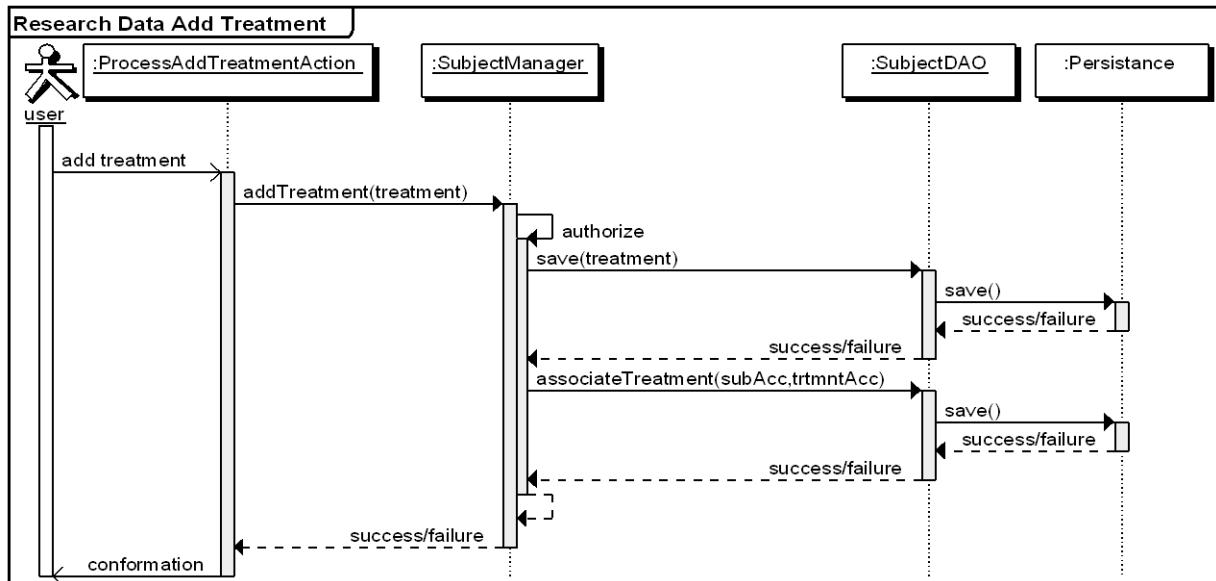
**Delete Subject Class Diagram****Delete Subject Sequence Diagram**

**Add Treatment**

The next section shows any treatments currently linked to this subject. In ImmPort treatments are not considered primary research entities, but are considered as children of subjects. In this section you can add a new treatment, delete a treatment or edit an existing treatment.

A screenshot of a Mozilla Firefox browser window showing the 'Add Treatment' form. The title bar reads 'ImmPort: Immunology Database and Analysis Portal - Mozilla Firefox'. The address bar shows the URL 'immpore.net https://www.immpore.net/immporeWeb/workspace/subject/displayAddTreatment.do?entityPath=subject&'. The main content area is titled 'Add Treatment' and contains ten input fields for various treatment parameters: Agent, Weight Value, Weight Unit, Concentration Value, Concentration Unit, Volume Value, Volume Unit, Time Value, Time Unit, Temperature Value, and Temperature Unit. Below these fields are two buttons: 'Create' and 'Cancel'.

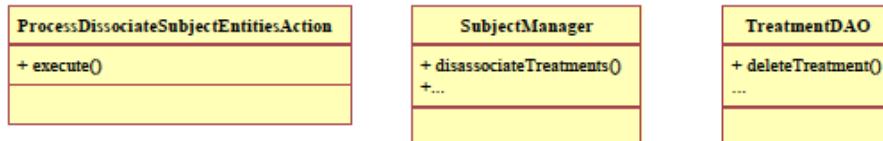
**Add Treatment Class Diagram****Add Treatment Sequence Diagram**



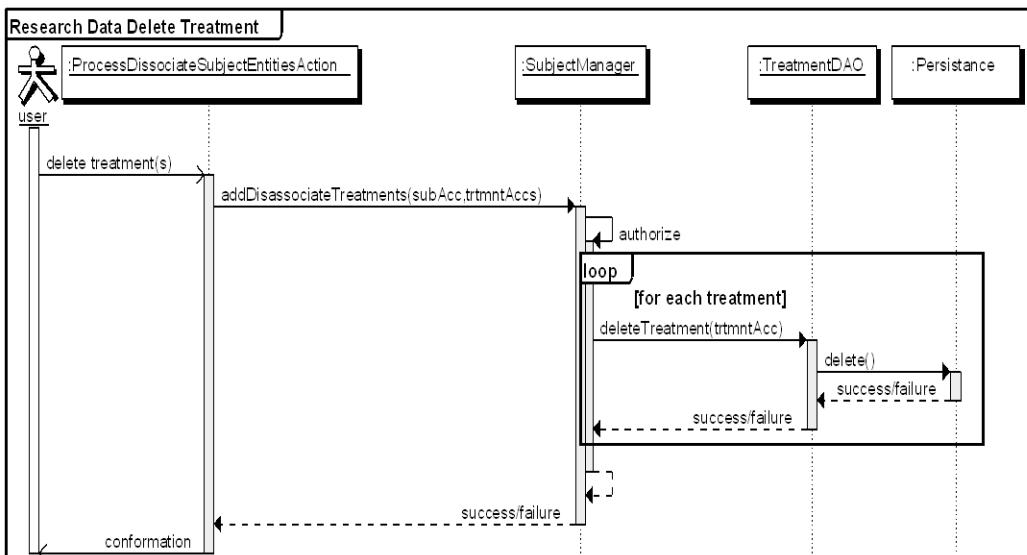
### 3.6.5.10 Delete Treatment

From the summary grid for treatments in the main subject update page, the delete treatment button allows the user to remove a treatment from the database.

**Delete Treatment Class Diagram**



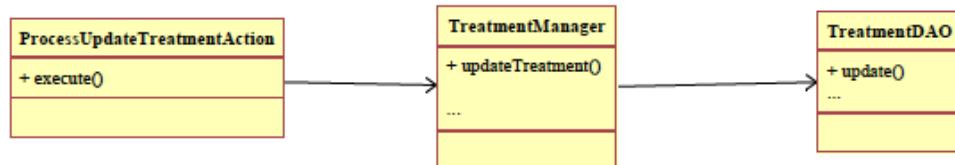
**Delete Treatment Sequence Diagram**



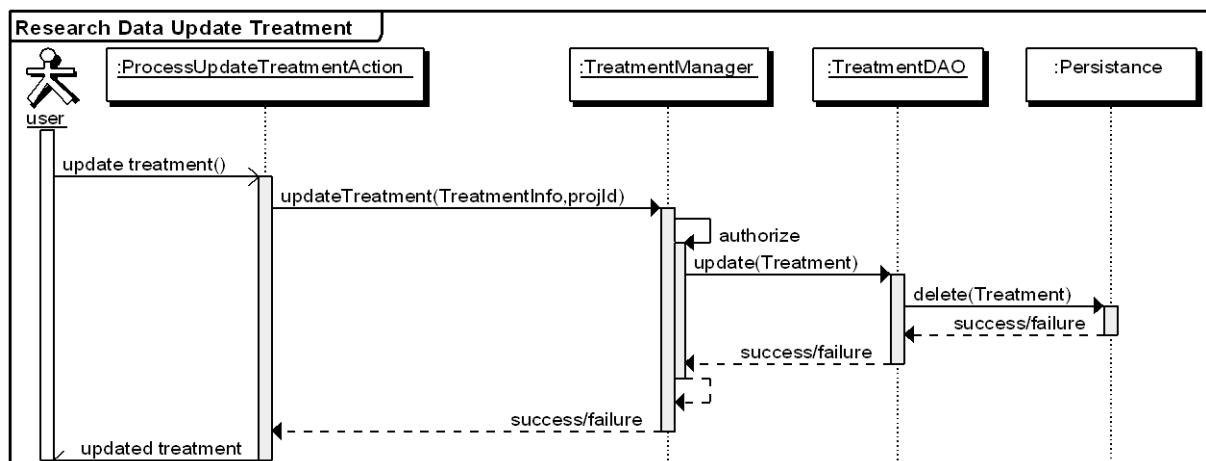
### 3.6.5.11 Update Treatment

Treatments may also be updated in the same interface as was shown earlier for entering a new treatment with the prior values now populated.

**Update Treatment Class Diagram**



**Update Treatment Sequence Diagram**



### 3.6.5.12 Display Non-Associated Entities

A subject record may be associated with many protocols and many biological samples.

The screenshot shows a software interface for managing subject associations. It consists of two main sections: 'Protocols used' and 'Biological Samples Obtained from this Subject/Organism'.

**Protocols used:**

Accession Number	User-Defined ID	Name
PTL29	ELISA Protocol	ELISA Protocol Information

**Biological Samples Obtained from this Subject/Organism:**

Accession Number	Name
BS1602	Blood from Subject 1
BS1603	Lymphocytes from Subject 2

A 'Selected items' message indicates 'BS1602' was selected. A 'Finished' button is visible at the bottom left of the second grid.

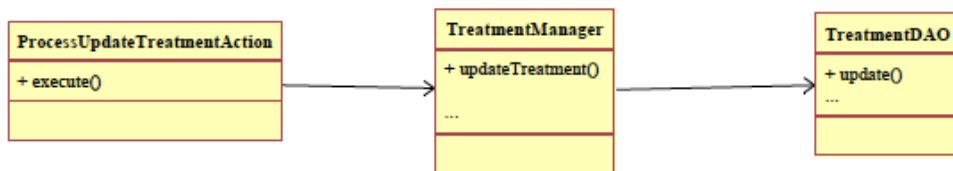
To add additional associations, the user clicks the Add button in the appropriate grid and a window is displayed with a grid containing the available protocols or biological samples that may be associated to the subject. The sequence diagram below will only show how the non-associated biological samples are retrieved and displayed.

Biosamples

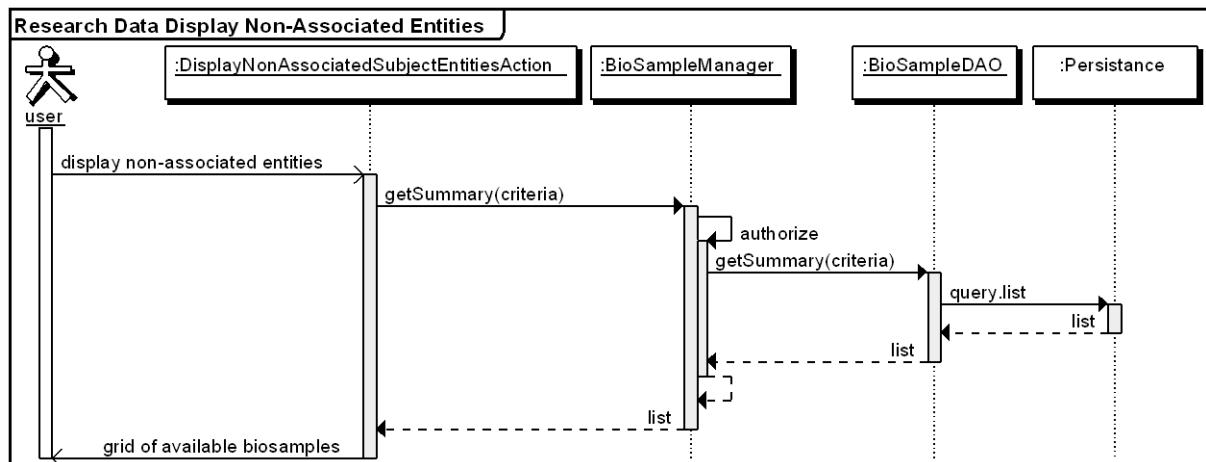
Select one or more Biosamples. Then, click on the Associate Button. Close the dialog box when done.

Accession Number	Study Date	Subject ID	Species	Type	Sub Type	Study Title	Arm
BS1604			Homo sapien	Cell	B cells		null
BS1605			Homo sapien	Tissue	Blood		null
BS1606			Homo sapien	Cell	Lymphocytes		null
BS1607			Homo sapien	Cell	B cells		null
BS1608			Homo sapien	Tissue	Blood		null
BS1609			Homo sapien	Cell	Lymphocytes		null
BS1610			Homo sapien	Cell	B cells		null
BS1611			Homo sapien	Tissue	Blood		null
BS1612			Homo sapien	Cell	Lymphocytes		null
BS1613			Homo sapien	Cell	B cells		null
BS1614			Mus musculus	Cell	B cells		null
BS1615			Mus musculus	Cell	B cells		null
BS1616			Mus musculus	Cell	B cells		null

#### Display Non-Associated Entities Class Diagram



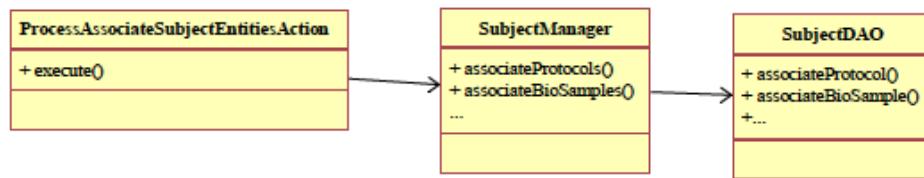
**Display Non-Associated Entities Biological Sample Sequence Diagram**



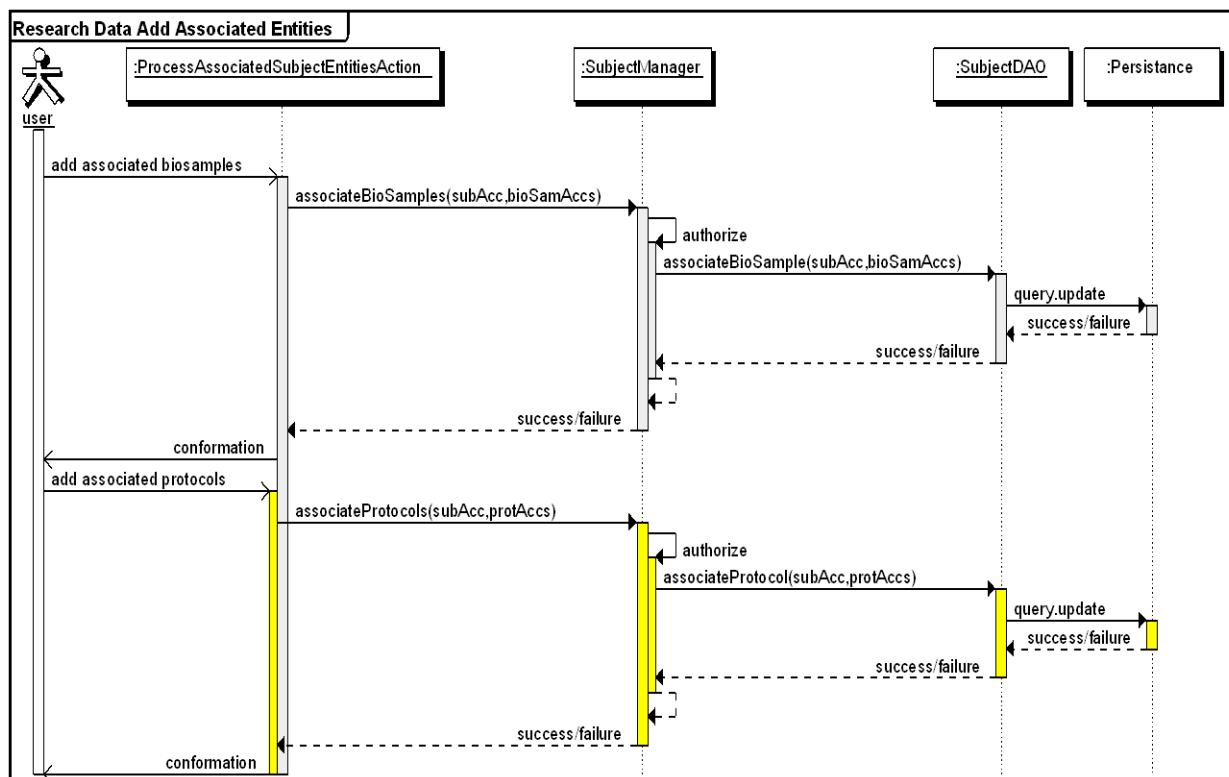
### 3.6.5.13 Associate Entities

After the user has chosen biological sample or protocols to link to the subject they submit the request using the following process.

**Associate Entities Class Diagram**



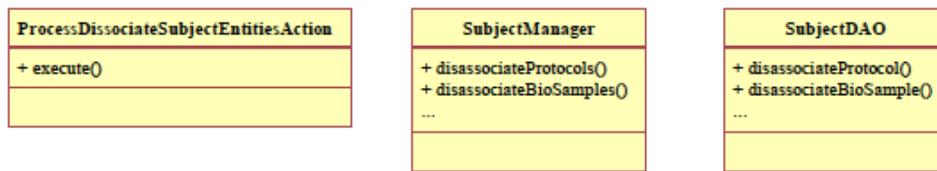
**Associate Entities Sequence Diagram**



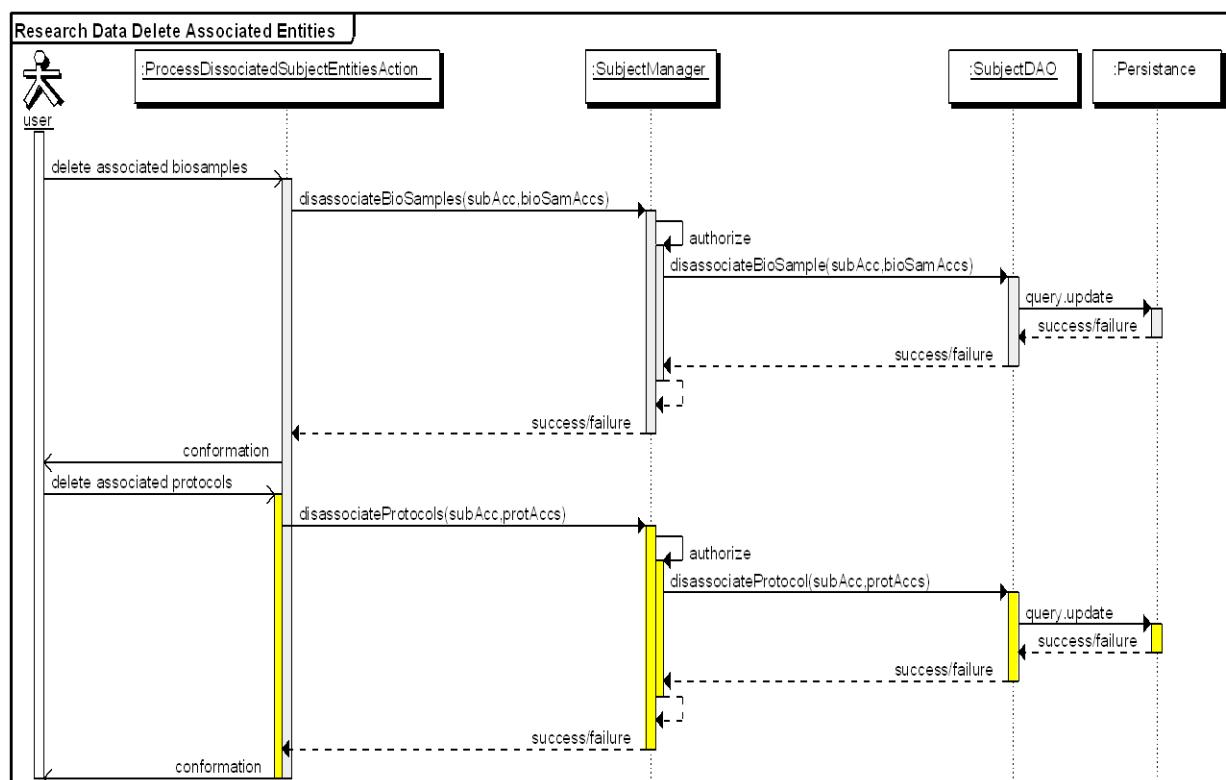
### 3.6.5.14 Disassociate Entities

On the edit summary page the user may choose biological sample or protocols displayed in the associated entities grid to mark them for unlinking from the subject. The process is outlined below.

**Disassociate Entities Class Diagram**



**Disassociate Entities Sequence Diagram**



**Edit Subject Class Diagram**

