

California State University, San Bernardino

CSUSB ScholarWorks

Theses Digitization Project

John M. Pfau Library

2009

Biomedical Informatics Research Network (BIRN) Data Repository

Laurence Jonathan Bouts Bohannan

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd-project>



Part of the [Data Storage Systems Commons](#)

Recommended Citation

Bohannan, Laurence Jonathan Bouts, "Biomedical Informatics Research Network (BIRN) Data Repository" (2009). *Theses Digitization Project*. 3656.

<https://scholarworks.lib.csusb.edu/etd-project/3656>

This Project is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

BIOMEDICAL INFORMATICS RESEARCH NETWORK (BIRN)

DATA REPOSITORY

A Project

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in

Computer Science

by

Laurence Jonathan Bouts Bohannan

December 2009

BIOMEDICAL INFORMATICS RESEARCH NETWORK (BIRN)

DATA REPOSITORY

A Project

Presented to the

Faculty of

California State University,

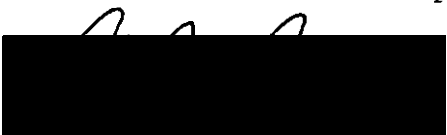
San Bernardino

by

Laurence Jonathan Bouts Bohannan

December 2009

Approved by:



Arturo Concepcion, Chair, Computer Science

03 Dec 2009
Date



Keith E. Schubert



Ernesto Gomez }

ABSTRACT

A great demand has been placed on the resources that researchers have at their disposal by the copious amounts of data that are being created in the domain of biological sciences through the aid of computers, such as through the Magnetic Resonance Imaging. An economical solution to this growing problem is to accumulate this data onto large-scale computer clusters, which are interconnected to other clusters via high-speed Internet backbones to create a scientific computational grid. The introduction of NIH guidelines which outlines specific timelines to the sharing of derived information, tools or data from NIH funded studies have created a need for a common resource to increase the availability of and access to data contributed from the scientific community at large. The Biological Informatics Research Network (BIRN) Data Repository function is to streamline the submission process for researchers to categorize, approve, curate, and publish these data sets. By supporting the sharing and exchange of these data types in fulfillment of the NIH guidelines the BIRN Data Repository will have the ability to grow as the need develops for these researchers to

easily share information. The BIRN Data Repository is intended to incorporate portions of the existing BIRN infrastructure with grid components such as the Storage Resource Broker, Gridsphere Web portal, and existing collaborative portlet environments to reduced the overall development timeline for creating a full featured collaborative scientific data repository.

ACKNOWLEDGMENTS

Saying this project took a little longer than expected would be an understatement. Foremost, I would like to thank all those mentioned below for their patience with my somewhat intermittent work ethic.

I would like to express my gratitude to my advisor, Dr. Arturo Concepcion, for his endless support, patience, and encouragement throughout my graduate studies. He provided a motivating, enthusiastic, and critical atmosphere during the process. It was my great pleasure to conduct this project under his supervision. His technical and editorial advice was essential to the completion of this project. My thanks also goes to the members of my committee, Dr. Keith Schubert and Dr. Ernesto Gomez, for their timely advice, and thoughtful comments. I also want to express my gratitude for their reading of this document and providing many valuable comments that improved the presence and contents of this document.

I would like to express my gratitude to Jeffrey Grathe for giving me the opportunity to work on this project and taking the time to point me in the right

direction. Additionally I would like to thank all my coworkers at BIRN for assisting in the development of the system, it wouldn't had been possible without their domain expertise.

I would like to express my deepest gratitude and indebtedness to my fiancé Jamie Intervalo for her understanding, patience, and love during the past few years. Her support and encouragement was in the end what made this project possible. And of course, my parents, who receive my special thanks and love for their early guidance, support, and persistence.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGMENTS	v
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER ONE: INTRODUCTION	
1.1 Background	1
1.1.1 Biomedical Informatics Research Network.....	3
1.1.2 Tools and Functions of Biomedical Informatics Research Network.....	4
1.2 Purpose of the Project	5
1.2.1 Organization of this Document.....	7
CHAPTER TWO: SOFTWARE REQUIREMENT SPECIFICATION	
2.1 Introduction	8
2.1.1 Purpose	8
2.1.2 Scope	9
2.1.3 Definitions	13
2.1.4 Overview	17
2.2 Overall Description	17
2.2.1 Product Perspective	17
2.2.2 System Interface	23
2.2.3 Product Functions	37
2.2.4 User Characteristics	39
2.2.5 Constraints	40

2.2.6 Assumptions and Dependencies	40
2.2.7 Apportioning of Requirements	40
2.3 Specific Requirements	41
2.3.1 External Interfaces	41
2.3.2 Functional Requirements	73
2.3.3 Performance Requirements	73
2.3.4 Logical Database Requirements	73
2.3.5 Design Constraints	73
2.3.6 Software Systems Attributes	73
CHAPTER THREE: SOFTWARE DESIGN	
3.1 System Design	76
3.1.1 System Overview	76
3.1.2 Portal Framework	77
3.1.3 Portlets	78
3.1.4 Portlet Services	78
3.1.5 User Interface	79
3.1.6 Database	79
3.1.7 Overview	81
CHAPTER FOUR: QUALITY ASSURANCE AND TESTING	
4.1 Introduction	99
4.2 Unit Testing	99
4.3 Integration Test	101
4.4 System Test	102
4.5 Test Samples	105

4.5.1	Creation of Biomedical Informatics Research Network Data Repository Submission.....	105
4.5.2	Upload of Internal Review Board Documents.....	109
4.5.3	Error Handling.....	111
CHAPTER FIVE: MAINTENANCE AND USER MANUAL		
5.1	Maintenance Manual	117
5.1.1	Directory Organization.....	117
5.1.2	Installing and Re-Compiling of the Biomedical Informatics Research Network Data Repository Portal on a Tomcat Server.....	118
5.2	User Manual	121
5.2.1	Browsing Biomedical Informatics Research Network Data Repository Entries.....	121
5.2.2	Downloading Data Sets.....	125
5.2.3	Creating a User.....	128
5.2.4	Initiating a Biomedical Informatics Research Network Data Repository Submission.....	137
5.2.5	Uploading Internal Review Board Documentation.....	138
5.2.6	Sending Notifications to the Biomedical Informatics Research Network Data Research.....	142
5.2.7	Viewing Submission Progress.....	144
5.2.8	Administration of a Biomedical Informatics Research Network Data Repository Submission.....	146

5.2.9 Controlling Submission Progress	147
5.2.10 Setting a User Alert	149
5.2.11 Clearing an Administrative Alert	150
5.2.12 Printing a Submission	152
5.3 Improvements	155
CHAPTER SIX: CONCLUSION AND FUTURE DIRECTIONS	
6.1 Conclusions	158
6.2 Future Direction	161
BIBLIOGRAPHY	166

LIST OF TABLES

Table 4.1. Report of Biomedical Informatics Research Network Data Repository Unit Test.....	99
Table 4.2. Report of Biomedical Informatics Research Network Data Repository Integration Test....	101
Table 4.3. Report of Biomedical Informatics Research Network Data Repository System Test for User Functions.....	102
Table 4.4. Report for Systems Test for Administrative Functions.....	103

LIST OF FIGURES

Figure 2.1. Use-Case Diagram For Biomedical Informatics Research Network Data Repository Administrators/Curators	18
Figure 2.2. Use-Case Diagram For Data Contributors And Data Consumers/Guests	22
Figure 2.3. Deployment/Component Diagram For Biomedical Informatics Research Network Data Repository	25
Figure 2.4. Activity Diagram of the New Submission Workflow Process for the Biomedical Informatics Research Network Data Repository	28
Figure 2.5. Login Page	42
Figure 2.6. My Account Page	43
Figure 2.7. My Submissions Overview Page	44
Figure 2.8. Submission Specific Overview Page	46
Figure 2.9. New Submission Initiation Page	47
Figure 2.10. General Information Page	49
Figure 2.11. General Information Page (Continued)	50
Figure 2.12. General Information Verification Page	51
Figure 2.13. Accession Number Confirmation Page	52
Figure 2.14. High Level Information Page	54
Figure 2.15. High Level Data Page	56
Figure 2.16. High Level Data Page (Continued)	59
Figure 2.17. High Level Data Confirmation Page	60
Figure 2.18. Thank You Page	61
Figure 2.19. Internal Review Board Instruction Page	63

Figure 2.20. Internal Review Board Template Upload Page.....	65
Figure 2.21. Browse Biomedical Informatics Research Network Data Repository Page.....	66
Figure 2.22. Publish Biomedical Informatics Research Network Data Repository Entry Page.....	69
Figure 2.23. Publish Biomedical Informatics Research Network Data Repository Entry Page (Continued).....	70
Figure 2.24. My Profile Page.....	71
Figure 2.25. Administration Page.....	72
Figure 3.1. Biomedical Informatics Research Network Data Repository System Diagram.....	77
Figure 3.2. Biomedical Informatics Research Network Data Repository Database Schema.....	80
Figure 3.3. Biomedical Informatics Research Network Data Repository Database Schema (continued).....	81
Figure 3.4. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 1).....	83
Figure 3.5. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 2).....	84
Figure 3.6. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 3).....	85
Figure 3.7. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 4).....	86
Figure 3.8. Class Diagram of the Package net.birn. services.resources.datarepository.impl.....	89

Figure 3.9. Class Diagram of the Package net.birn. services.resources.datarepository.impl (Continued)	90
Figure 3.10. Class Diagram of the Package net.birn. services.resources. datarepository.impl (Concluded)	91
Figure 3.11. Class Diagram Of The Package net.birn. services.resources.datarepository.....	92
Figure 3.12. Class Diagram Of The Package net.birn. services.resources.datarepository (Continued)	93
Figure 3.13. Class Diagram of the Class net.birn. services.resources.datarepository. OldSubmission (Part 1)	94
Figure 3.14. Class Diagram of the Class net.birn. services.resources.datarepository. OldSubmission (Part 2)	95
Figure 3.15. Class Diagram of the Class net.birn. services.resources.datarepository. OldSubmission (Part 3)	96
Figure 3.16. Class Diagram of the Class net.birn. services.resources.datarepository. OldSubmission (Part 4)	97
Figure 4.1. Launch Page For Beginning A New Submission	105
Figure 4.2. Launch New Submission	107
Figure 4.3. High Level Information Verification	108
Figure 4.4. Accession Number Confirmation	109
Figure 4.5. Internal Review Board Upload Before Submit	110
Figure 4.6. After Completion of the Internal Review Board Document Upload, this is Displayed...	111

Figure 4.7. Error Handling on a Required Field, Submission Title	112
Figure 4.8. Invalid Input - Launch New Submission.....	113
Figure 4.9. Invalid Input - New Submission Information Verification	114
Figure 4.10. Source Code for the Submission Title Section of the Table Figure 4.9.....	115
Figure 4.11. Improperly Formatted - Launch Submission with Malformed Email Value.....	115
Figure 4.12. Improperly Formatted - Launch Submission with Error Message.....	116
Figure 5.1. Default View of the Browse Biomedical Informatics Research Network Data Repository Page	121
Figure 5.2. Expanded View of Browse Biomedical Informatics Research Network Data Repository Page	123
Figure 5.3. Download Button Located at the Bottom of the Detailed View of a Published Submission	124
Figure 5.4. View of a Typical Resource Link That Would be Provided by the Biomedical Informatics Research Network Data Repository	126
Figure 5.5. Download Information Form Page.....	127
Figure 5.6. Default View of the Portal to a User Who is Not Logged In	129
Figure 5.7. New User Registration Form.....	130
Figure 5.8. Verification Page of the New User Registration Process	131
Figure 5.9. Email That Would be Received From the Portal	132

Figure 5.10.	View of a Normal Account Activation Email.....	133
Figure 5.11.	Prompt for the Users Email During Activation.....	134
Figure 5.12.	Prompting the User to Enter a Password....	135
Figure 5.13.	View of the Confirmation Message That Would be Generated for the Successful Activation of the Account.....	136
Figure 5.14.	View of the Error Message That Would be Displayed if the Account Creation had an Error.....	136
Figure 5.15.	Location of the Create Submission Link From the Default View.....	137
Figure 5.16.	Launching a New Submission.....	138
Figure 5.17.	Adding Additional Files to be Uploaded....	140
Figure 5.18.	Submitting the Files to be Uploaded to the Biomedical Informatics Research Network Data Repository.....	141
Figure 5.19.	Clicking the Magnifying Glass to View the Uploaded File.....	142
Figure 5.20.	Clicking the Link to Lead to the Contact Biomedical Informatics Research Network Data Repository Page.....	143
Figure 5.21.	Sending the Notification to the Biomedical Informatics Research Network Data Repository.....	144
Figure 5.22.	Click on a Submission to See a More Detailed View of the Information.....	145
Figure 5.23.	Detailed View of a Submission.....	146
Figure 5.24.	Links That Correspond to Controlling the Progress of the Submission.....	148

Figure 5.25. Link to Create an Administrative Alert for a Submission.....	149
Figure 5.26. Default View of the Admin Portlet.....	151
Figure 5.27. Clearing an Alert From the User to the Admin.....	152
Figure 5.28. The Link That Leads to the Print Feature in the Admin Overview Page.....	153
Figure 5.29. View of the Print Submission Page.....	154

CHAPTER ONE
INTRODUCTION

1.1 Background

We see that there have been great advances in scientific research in the field of biology, most notably those that were derived through the aid of computers. Whether it is high throughput genomic sequencing, mathematical population genetics modeling, proteomics, or high resolution Magnetic Resonance Imaging (MRI), all these techniques present a common attribute. They produce vast amounts of data, which then needs to be manipulated in various ways to render useful conclusions. This places a high demand on the volume of storage, the speed of accessing the data, and the large amount of computational processing power required to analyze the data. Additionally, all these hurdles must be address in an economically responsible manner both resource wise and financially. A solution to this problem is through the use of large-scale computational clusters, these clusters are usually made up of common off the shelf components that can easily and cheaply obtainable. Besides enabling scientist to possess a large amount of computational processing power

there are some additional benefits that accompany the use of some of these types of systems.

More specifically one sub-category of computational clusters is data grid computing; these computing clusters, which are geographically dispersed, are connected via the Internet to form a grid. Though connected together they do not fully trust each other, it is only through the use of a Grid Security Infrastructure, which is based off of public key digital signatures, that an authentication of the users true identity can be established to allow access to other resources on the grid. One feature of a grid is that it allows all the users to be mutually connected to all the resources that the grid has available This enables the users to possess a higher degree of freedom over the data resources that are stored on the grid. The distributed security, which is inherent in the grid, allows users to easily share access to the data that they have obtained, actions like these promote the growth of a basic collaborative environment.

The Biomedical Informatics Research Network is a grid-based system of 29 geographically separated sites. Each of which consisted of racks of standardized hardware that were preloaded with a default software stack. The Biological

Informatics Research Network Portal, which serves as the gateway for access to the network's tools and data was housed at the University of California San Diego. For the past three years I had worked as developer on the Biomedical Informatics Research Network, which was a project under the Center for Research of Biological Systems at University of California San Diego. Our team was responsible the development of new and existing features for Biomedical Informatics Research Network. Jeff Grethe gave the task of developing the BIRN Data Repository to me.

1.1.1 Biomedical Informatics Research Network

The goal of the Biomedical Informatics Research Network (BIRN) is to expand on this inherent feature of a grid by creating a collaborative portal environment to allow greater ease of access, more powerful features, and to promote the growth of the collaborative efforts among biomedical researchers. BIRN was developed at the University of California San Diego under various grants from the NIH. The current application serves as the framework for existing studies in neurological diseases.

1.1.2 Tools and Functions of Biomedical Informatics Research Network

The BIRN infrastructure is build upon the Gridsphere framework that is an open source jsr-168 compliant grid enabled portal. Gridsphere provides built in functionality for user administration, persistence framework, and portlet layout management. Additionally, grid job management is achieved through the use of the gridportlets, a Web application package that is also developed by the creators of gridsphere portal. Gridportlets utilizes Gird Account Management Architecture (GAMA) and Storage Resource Broker (SRB) functionality through specialized portlets to manage user credentials and distributed file system access. Along with these middlewares are additional portlet Web applications like the collaborative Bluesquid package that enables the users of the BIRN portal to create 'projects' that act as collaborative work environments. These projects allow the creators to control access through a group membership based user privilege system, administrative access can be granted by assigning special administrator roles to members of the group. The package also has support for forums, blogs, lists, charts, email, and RSS feeds. Besides providing the basic features of the collaborative

work environment, the BIRN portal system also provides client tools for accessing the various resources of the data grid.

1.2 Purpose of the Project

In 2003, the NIH published a guide that contained an extension of the NIH policy on sharing research resources and reaffirmed its support for the concept of data sharing. The implications of this publication are that it now requires all derived information, tools, or data from NIH funded studies despite form or medium of which it may be stored to be released in a timely manner. This has created a need for a repository tool of sorts, which would be able to access multiple legacy systems and handle various heterogeneous data types.

The primary goal for the BIRN Data Repository is to provide a common resource to increase the availability of, and access to, data contributed from the scientific community at large, rare data sources, complicated imaging data resources, and data from existing well defined data sharing efforts. The BIRN Data Repository encourages data sharing both for investigators wanting to contribute to the established data storage and informatics infrastructure, as

well as those wishing to extend the reach of existing data resources through collaboration with the BIRN Data Repository. The BIRN Data Repository function is to streamline the submission process for researchers to categorize, approve, curate, and publish their data sets. By supporting the sharing and exchange of these data types in fulfillment of the NIH guidelines the BIRN Data Repository will have the ability to grow as the need develops for these researchers to easily share information. There are currently no widely accepted system for clinical imaging that are available to the public that provides data repository functionality, which are integrated with a workflow to guide the user through the publishing process and collaborative features. Previously, there was a static version of existing data sets that have been accepted by BIRN for sharing to the public [<https://bdr-portal.nbirn.net/gridsphere/gridpsphere>].

The BIRN Data Repository is intended to incorporate existing portlet Web applications that were developed by BIRN for use in the BIRN portal, this will also reduce the overall development timeline, the system was presented for use to existing BIRN enabled research projects for use

after the 1.0 release is made. The BIRN Data Repository is the first implementation of a full stand-alone system to unify various heterogeneous data resources of researchers and guide them through all the steps while allowing for the sharing and analysis of the data sets that are available to the public.

1.2.1 Organization of this Document

The remaining chapters of this documentation will be organized as follows: Chapter 2 describes the software requirements specification. Chapter 3 provides a description of the system architecture and detailed design. Chapter 4 is a system test that shows test cases that were used for the acceptance of the project. Chapter 5 is the maintenance and users manual. Finally, Chapter 6 concludes the project and lists suggestions for future developments. Throughout this document I will refer to the system as the "BDR" or "BIRN Data Repository", they will be used interchangeably.

CHAPTER TWO
SOFTWARE REQUIREMENT SPECIFICATION

2.1 Introduction

2.1.1 Purpose

The purpose of this document is to form a software requirements specification (SRS) that details the features that are going to be developed for the BIRN Data Repository. This document will establish the basis for understanding and communicating the requirements between the users and the developers on what the BIRN Data Repository is, and what are its functionalities, provide a basis for validation and verification of the BIRN Data Repository features, and serve as a basis for enhancement. The users for which the BIRN Data Repository is intended includes:

- Data Contributors - Data contributors may be existing BIRN users or external researchers who have collected data sets from their own studies and are interested in sharing the information so that it may be incorporated into other studies;
- Data Repository Curators - The system is also designed for the data repository curators, which will enable

the curators to view, monitor, and help in the assimilation of the datasets in the data repository;

- BIRN Data Consumers - Once information is published in the system, those data sets will be freely available via the Internet to the whole scientific community to use.

The purpose of the BIRN Data Repository is to promote the sharing and publication of neurological imaging data sets to other entities via a publicly accessible Internet portal. The system will facilitate the curation process by allowing users to download existing or newly obtained datasets in various formats. The system will create a sustainable archive of data that is freely available for sharing and exchange. Curators will work with investigators to enable sharing of their existing data, following NIH guidelines. This will provide a means to capture, curate, store, query, view, and download biomedical data, which includes biomedical human and animal model data such as imaging, clinical, micro-array, and associated data.

2.1.2 Scope

With recent advances in medical and scientific research, technology has greatly increased the amount of data that can be collected from a single subject. Several

decades ago an "in vivo" study of a subject would only yield the obvious and superficial physically observable attributes. An in-depth or more detailed view of all the comprehensive features of an organ from the larger superstructure down to the basic building blocks of the cell were only in the wildest imagination of scientists of the time. However, with the development of methods like Magnetic Resonance Imaging the amount of information that can be obtain has exploded. The development of techniques like MRI is a double edged sword, along with the extensive amounts of data that are provided on a subject came the challenge of presenting the data in a understandable manner, storing the mass amount of data, and making the data available to researchers in a convenient readily accessible mode. Researchers utilizing the BIRN are addressing these challenges plus more.

The BIRN infrastructure aids in the storage, retrieval, analysis, and documentation of biomedical imaging data. Using open source toolkits and grid technologies built on top of the high-speed Internet-2/Abilene backbone, BIRN is a virtual community of shared resources consisting of a central coordinating center, four test beds and other associated collaborative projects.

This overall infrastructure consists of distributed data collections hosted on dedicated storage and computational resources located at each participating site, a federated data management system and data integration environment, an XML schema for data exchange, and analysis pipelines designed to leverage both the distributed data management environment and the available grid computing resources, and visualization environments to enable data discovery and interpretation by domain scientists. The BIRN hardware infrastructure consists of physical racks deployed at each participation site and a collection of services and servers hosted by the BIRN Coordinating Center. Currently there are over 25 BIRN racks installed. Each rack supports a standardized configuration of BIRN applications and data storage. The main mode of interaction for users with BIRN is through a Java based JSR-168 compliant grid enabled portal (Gridsphere). The BIRN portal has a variety of existing portlet-based applications that allow the researchers to collaborate and view the data in various methods.

The new application discussed in this document is the BIRN Data Repository, one of the main functions of the system is to further increase the availability of data to

researchers, not just existing users of the network but to public researchers that may have no affiliation with BIRN. This publicly available data must be cleaned and prepared prior to its release on the Internet especially if the data is derived from human subject. The system will aid in the submission process for the publication of datasets in the Internet. By streamlining the submission workflow, the researchers will be guided through the formalities and the steps, which are necessary to prepare the data and receive the appropriate approvals for the publication. Data contributors will be able to track various download statistics on the data. Additionally, the system will provide an interface in which researchers can easily search for information of interest, with capabilities like searching for keywords, favorite topics of interest, and recent popular data set releases.

The system is intended to aid researchers who are interested in contributing and sharing their obtained data to the public community, as would researchers who are interested in obtaining data on a specific topic would find the site very useful. The system will be developed on the Gridsphere portal framework backed by a Postgres database and incorporate several BIRN package libraries to aid in

utilizing some of the preexisting functionalities; mainly these include the Storage Resource Broker (SRB) portlets and the collaborations (Bluesquid) portlets.

2.1.3 Definitions

- BIRN

Biomedical Informatics Research Network is a federated and distributed infrastructure for the storage, retrieval, analysis, and documentation of biomedical imaging data.

- BIRNLex

Provides a controlled terminology for annotation of BIRN data sources which currently include image databases, containing data from structural and functional magnetic resonance imaging (MRI) on human subjects involved in studies on Alzheimer's disease and schizophrenia, and multiscale imaging of mouse models of human neurological disease using MRI, light and electron microscopic imaging

- Collaboration Portlets (BlueSquid)

Is a package of portlets that provides basic collaborative functions. Users can create chats, blogs, lists, forums, Emails, and RSS feeds that are commonly associated to a project. Users can also control access and visibility to the resources by other users.

- Data Grid

Is a grid computing system that deals with data.

- GAMA

Grid Account Management Architecture is a GSI credential management and integration system tailored for use Cyber Infrastructure through web portals and web service-based clients.

- Grid Computing

Is a phrase used to describe multiple independent computing clusters which act like a "grid" because they are composed of resource nodes not located within a single administrative domain.

- GSI

Grid Security Infrastructure is an asymmetric encryption specification for secure, tamper-proof, communication between software in a grid computing environment.

- Gridsphere

Is a grid enabled, open source JSR 168 compliant service bases portal framework, which enables developers to quickly develop and deploy their third party portlets.

- Hibernate

Is an open source object-relational mapping (ORM) solution for the Java language: it provides an easy to

use framework for mapping an object-oriented domain model to a traditional relational database.

- IRB

Institutional Review Board is a group that has been formally designated to approve, monitor, and review biomedical and behavioral research involving humans with the aim to protect the rights and welfare of the subjects

- JSR 168

Version 1.0 of the Java portlet API

- BIRN Mediator

Is a data integration infrastructure provided by the BIRN Coordinating Center can make distributed and heterogeneous databases accessible

- Persistence Manger

Is used in the Gridsphere framework to add or remove objects to the database. Generally this is a wrapper around the Hibernate object relational mapping (ORM) library and provides basic methods for create/restore/update/delete (CRUD) operations.

- Portal

Aggregates one or more portlets into web pages, which are usually personalized or customized for individual users or groups of users.

- Portlet

Is a pluggable user interface components that are managed and displayed in a web portal. Portlets produce fragments of markup code that are aggregated into a portal page.

- RSS

An acronym for Really Simple Syndication, which is a Web content syndication XML format that is published to a website.

- SRB

Storage Resource Broker is a logical distributed file system based on a client-server architecture, which presents the user with a single global logical namespace or file hierarchy.

- UMLS

UMLS is a consolidated repository of medical terms and their relationships. Each biological concept in UMLS is associated with semantic classes such as gene or genome and amino acid, peptide, or protein.

- XCEDE

XML-based Clinical Experiment Data Exchange Schema is an XML-based data exchange schema designed to facilitate metadata transfer between databases, within and between software tools, and between the BIRN and external sites.

- XNAT

Is an open source java-based software platform designed to facilitate common management and productivity tasks for neuroimaging and associated data (www.xnat.org).

2.1.4 Overview

Section 3 follows the guidelines of IEEE Std. 830-1998 IEEE Recommended Practice of Software Requirements Specification. This section provides product perspective, a summary of product functions, a description of the characteristics of the expected users, and a list of assumptions and dependencies. Section 4 presents the specific requirements for this system. They are organized by mode, following the SRS Section 3 template as shown in IEEE Std 830-1998, Annex A, Paragraph A.3.

2.2 Overall Description

2.2.1 Product Perspective

There are three different users that were taken into consideration during the design phase; these are the data contributors, public browsers, and specialized BIRN Data Repository administrators (curators).

2.2.1.1 Biomedical Informatics Research Network Data Repository Administrator. The BIRN Data Repository

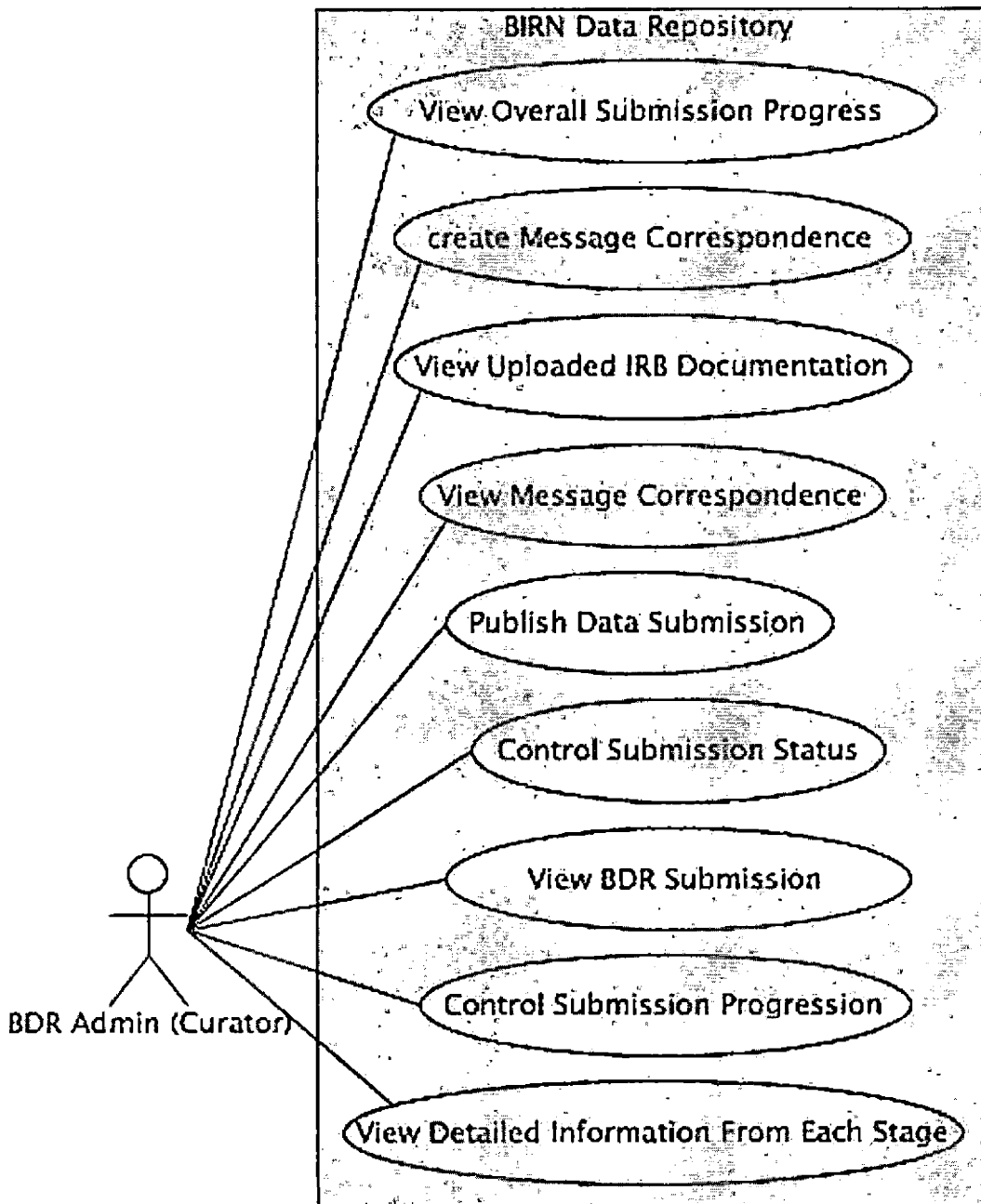


Figure 2.1. Use-Case Diagram For Biomedical Informatics Research Network Data Repository Administrators/Curators

administrators will have nine different functional use cases that must be implemented to meet the administrative needs of processing a data submission up to the published stage (Figure 2.1). These cases can be broken down into three different types of subsets for primary functions that can be performed on the submission: management, correspondence and viewing.

2.2.1.1.1 Management Functions. The three primary management functions that an administrator will have are to publish a submission, control submission progression and status. The status denotes whether or not a submission is on hold, suspended, deleted or rejected. The administrator will be able to change the status of a submission by placing and removing the appropriate flag attributes to the entry. When the administrator chooses to control the progression of a submission they are actually moving the submission from one stage to another. This movement can be to the next stage in the workflow or back to the previous one. Once the user has completed all the requirements deemed necessary for the submission to be made public the administrator will 'publish' the submission.

2.2.1.1.2 Correspondence. The administrators will have the ability to message the data contributors on recent

developments or problems that may arise. These messages will be kept during the duration of the submission processing and may be used to refer back to past conversations; each submission will have its own conversation log. If needed the administrator can view the conversation log, which will contain all the correspondences between the contributor and administrator.

2.2.1.1.3 Viewing. The administrator will be able to view the general progress of the submission in respect to which stage the submission is in the workflow. From the previous view the administrators can view the more detailed information from each of the different stages that has already been completed by the user. These detailed views will contain all the information that was entered by the data contributor at the time they completed the stage. Once the contributor has completed the requirements and uploads digital copies to the BIRN Data Repository the will administrators can view and download the various IRB documentation.

2.2.1.2 Biomedical Informatics Research Network Data Repository Contributor. The BIRN Data Repository contributor will have seven specialized functional use cases that must be implemented to be able to supply all the

pertinent information to fully process a data submission. Of course the primary use case of a contributor is that for which the primary purpose the portal was created, creation of a BIRN Data Repository submission, after which the contributor may proceed through the workflow. Similar to the administrator the contributor will be able to view and add to a conversation log, which will contain all the previous correspondences on the submission. Any communication that is preformed through the portal in regards to the submission will be recorded. The contributor will also be able to view the overall progress of the submission in respect to its location in the workflow, from this view they can view a detailed summary of the information that was entered for each section that has been completed. If the contributor's datasets contain information that was derived from human subjects they must complete the Internal Review Board section of the workflow. For this section there must be a feature to upload specific documentation for review, which can be multiple file of various formats. Once the contributor has selected a particular database, they must use the corresponding tools to upload their data to the database. Additional use cases

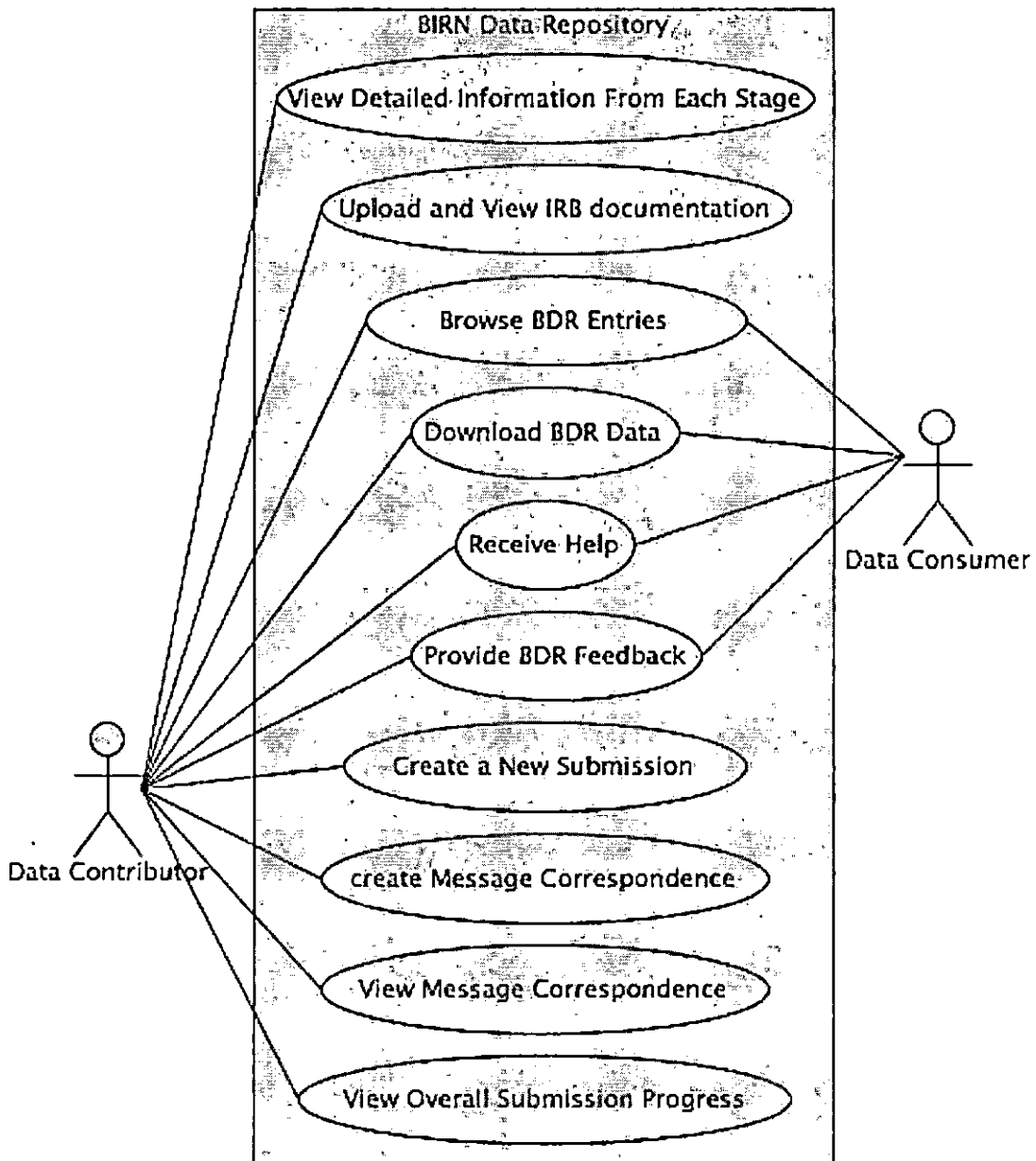


Figure 2.2. Use-Case Diagram For Data Contributors And Data Consumers/Guests

are common functions that are shared with the guest user role (see Figure 2.2).

2.2.1.3 Data Consumers and Guests. The third role that a user may have is as data consumers, they will interact with BIRN Data Repository as anonymous guests or researchers with an existing account. If a data consumer has interest in a dataset that has been published to the data repository they are not required to have an account to browse or download the information. Prior to downloading a dataset the user must agree to the terms of use of the BIRN Data Repository. The data consumer must also be able to receive help on any particular topic that is related to the BIRN Data Repository and leave feedback on possible improvements or problems that they experienced during their use of the BIRN Data Repository (see Figure 2.2).

2.2.2 System Interface

The general system architecture of the BIRN Data Repository is that of the typical client and server model. However the BRIN Data Repository is not limited to a single stand-alone server. The BIRN Data Repository consists of a federated system that is split into several components besides the main application server, the Storage Resource Broker server, the Mediator server, and the Grid Account Management Architecture server. The main application server runs the Gridsphere framework along with the accompanying

portlet applications in the tomcat servlet container. The application server also houses the Postgres database and clients for the appropriate services (SRB and GAMA). The Grid Account Management Architecture (GAMA) server handles authentication services and credentialing for Gridsphere. Data integration services are provided by the BIRN Mediator server, which uses a 'data mediation' type of architecture to access resources allowing for the integrated querying across several heterogeneous databases at one time. These services are used when searching for datasets tagged with a particular keyword. The Storage Resource Broker (SRB) Server is a data grid management system that provides a hierarchical logical namespace to manage the organization of the heterogeneous data resources.

The application server contains several additional portlet application packages (see Figure 2.3). SRB and GAMA application packages both provide portlet services to utilize the corresponding grid or authentication service. Gridportlets is a portlet toolkit that was developed by the creators of GridSphere to manage the credentials that are obtained from the GAMA server and assist in developing grid enabled portlets. Bluesquid is a portlet application that

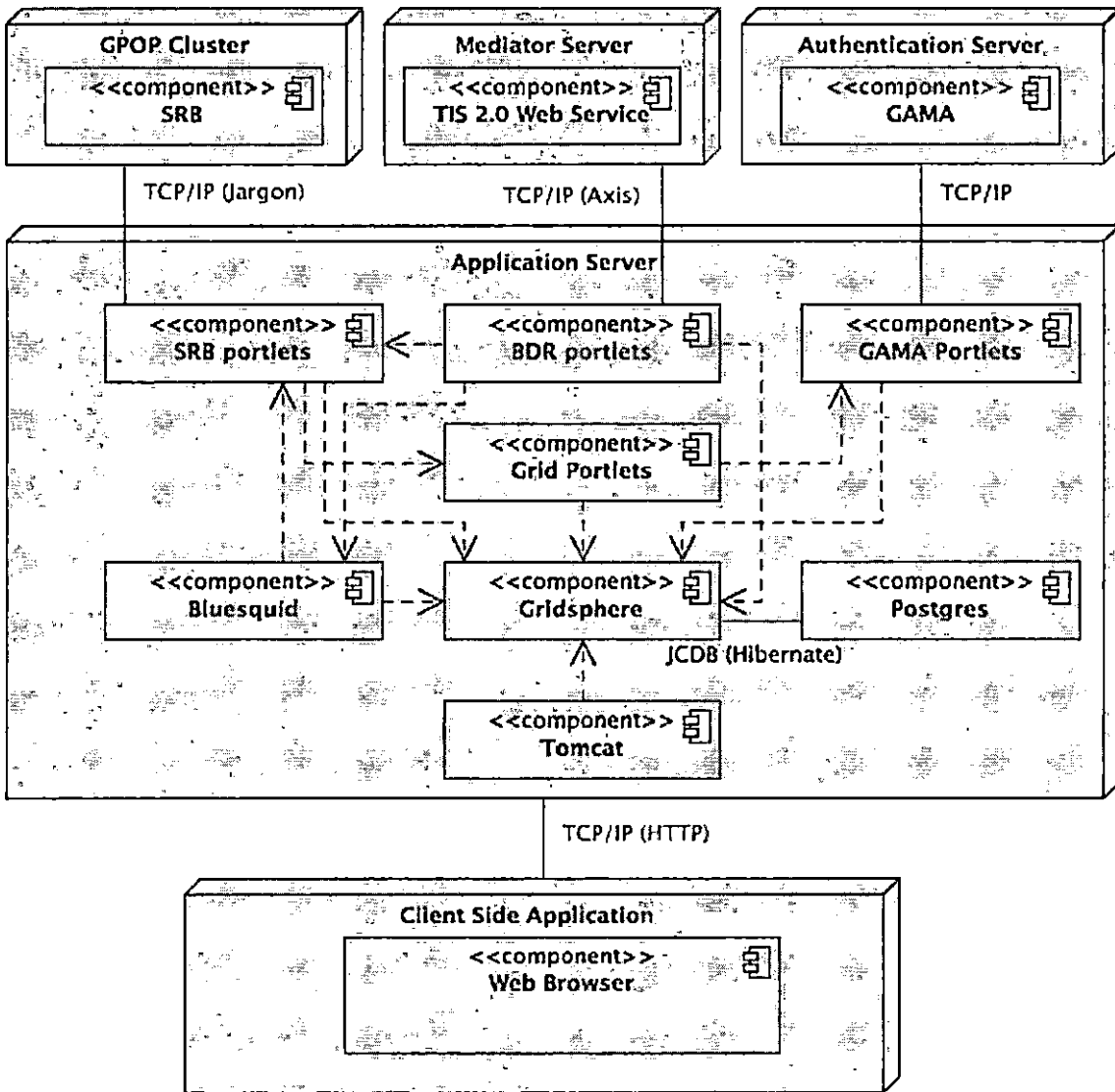


Figure 2.3. Deployment/Component Diagram For Biomedical Informatics Research Network Data Repository

was developed by BIRN to provide more collaborative features in the portal, this packages introduces in the concept of collaborative project work environments with group permissions.

2.2.2.1 User Interface. User interfaces for the system will be designed as html and incorporate JavaScript for more dynamic client side functionality. Portlet and JSP generate the contents dynamically in response to the user's requests. The following features will be incorporated to produce a more descriptive representation of the interface:

2.2.2.1.1 Login Page. The user will enter a Login ID and Password to gain entry to the system. Additionally there will be a link to lead to a new account creation (Figure 2.5). The user can also obtain help through a link to the standard Gridsphere password recovery page.

2.2.2.1.2 Account Page. This page will be the default page that is viewed by a logged in data contributor (Figure 2.6). The user can:

- View the current submissions that are in progress and those which have been published
- Display user created searches on topics of interest and favorite search terms
- A welcome message with recently released BIRN Data Repository data topics and additional news.
- Click on help links to aid in their orientation and usage of the system's functions.

2.2.2.1.3 Submissions Overview Page. This page is a more detailed view of the My Submissions Portlet view. It displays a progress timeline for each of the submissions, shows all the stages of the submission process, and indicates in what stage is the submission request (Figure 2.7).

2.2.2.1.4 Submission Specific Overview Page. This page is similar to the overview page. It displays the submissions progress timeline and indicates in what stage is the submission request. Additionally, the user can view a log of the progress of the submission, related documents, and links of interest. The user can also continue the submission from the point where they last left off (Figure 2.8).

2.2.2.1.5 New Submission Workflow. This section covers the interfaces of the pages that are used throughout the new submission creation process. An activity diagram of the overall workflow can be referenced, see Figure 2.4.

2.2.2.1.5.1 Submission Initiation Page (Launch New Submission Stage). This page is the starting page for a user who would want to begin the submission

process, it will contain general high level instructions and requirements at the bottom of the frame where there

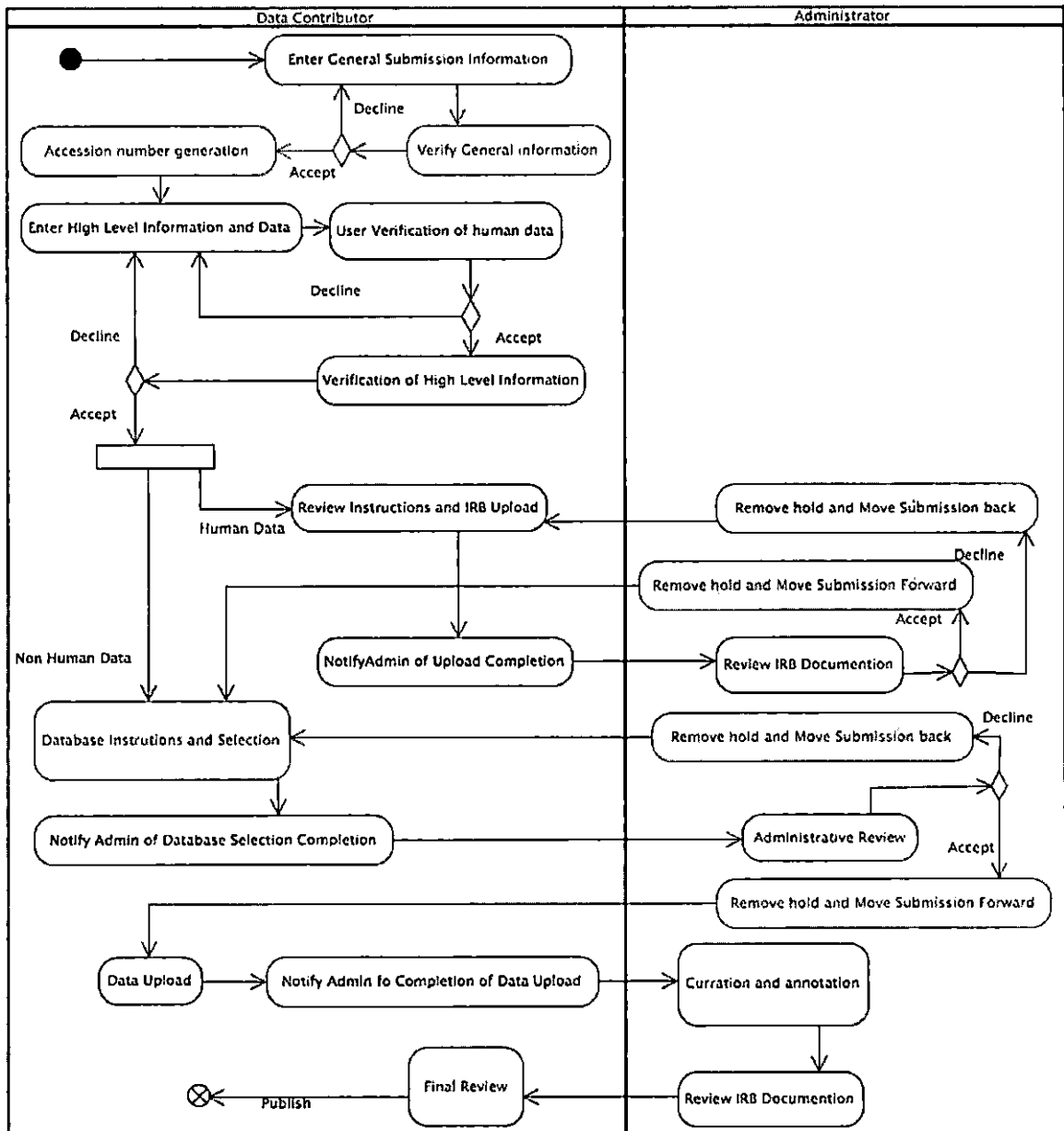


Figure 2.4. Activity Diagram of the New Submission Workflow Process for the Biomedical Informatics Research Network Data Repository

will be a start button to send the user to the first form in the submission workflow (Figure 2.9). There will be menu of links containing:

- Overview
- Detailed Instructions
- Minimum Requirements
- Data Contributors Agreement
- Data Use Agreement
- IRB Template for Sharing Human Data
- Glossary

2.2.2.1.5.2 General Information. This is the first form for the submission workflow. This page will allow the contributor to give general information about the submission such as: a title, a description or abstract of the study, and primary contact information for the contributor if different from the current creator. There is a pre-selected list of keywords that allows the user to select the relevant terms under which the submission can be indexed. The page will also display the users profile information such as name, email, and organization that the user belongs to and phone number (Figure 2.10). At the bottom of the page is a text area window that will display the contributors agreement, below the window there is a

check box for agreeing with the terms stipulated in the contributors agreement. Additionally if the checkbox specifying if the data is part of a BIRN project is selected, there will be a drop down list to select which project the data originates from (Figure 2.11).

2.2.2.1.5.3 General Information

Verification. This is the second page in the workflow. It would allow the contributor the opportunity to review the information that they have provided up to the point (Figure 2.12). There is a link to edit the new submission information.

2.2.2.1.5.4 Accession Number Confirmation

Page. This is the third page in the submission workflow. This page display a thank you to the contributor that informs them that an entry has been made for their submission in the system resulting in the assignment of an accession number. The accession number is assigned in the format of YYYY-(ORG)-XXXXX where YYYY stands for the current year ORG is the organization of the data repository (always BIRN in this case) and the last five characters is a randomly generated alpha numeric character sequence (Figure 2.13). The contributor is then presented the

option of continuing on with the workflow or continuing it later.

2.2.2.1.5.5 High Level Information Page

(Describe Your Data Stage). This is the fourth page in the workflow. This page will allow the contributor to add to various lists of investigator/co-authors, institutions / URLs, sponsors, Funding Sources/Grant Numbers, original and additional publications. Additionally there will be a field to add a list of affiliated institutions/URLs, sponsors, funding sources/grant number and citations or publications (Figure 2.14).

2.2.2.1.5.6 High Level Data Information

Page. This page will have fields to select whether or not the data is human or non-human derived, the number of cases in the sample, to specify if it is gender specific, the age range of the subjects, and the diagnostic group represented in the sample (specified in BIRNLex). The page will have an option selection boxes which lists the available data types such as imaging, demographics, behavioral / cognitive measures, clinical measures, microarray, etc (Figure 2.15). Additionally, an option list that represents the user selected data types, which is modified through the use of add and remove buttons. When the user submits the

information to the server a simple JavaScript alert box notifying the user that the data they are submitting does or doesn't contain any human data and asks them to confirm this with a I agree or I disagree button selection.

2.2.2.1.5.7 High Level Data Confirmation

Page. This is a simple screen that displays the entered values and verifies with the user that the data they are submitting is correct. The user will confirm this by selecting 'Submit' or 'Edit Information' to go back and make changes (Figure 2.17).

2.2.2.1.5.8 Thank You / Notification Page.

Besides the normal thank you text, this page will contain instructions on the next steps of the submission process. This will depend on whether or not human data was included in the data set (Figure 2.18).

2.2.2.1.5.9 Internal Review Board

Instructions (Internal Review Board Information Stage).

This page will give the user links to the IRB forms that must be downloaded and returned completed. Additionally it will contain instructions and links to other help pages pertaining to this topic (Figure 2.19).

2.2.2.1.5.10 Internal Review Board Template

Upload. This page allows the user to upload multiple

documents of various formats. They can be accompanied with titles and an upload event description (Figure 2.20).

2.2.2.1.5.11 Data Specific Upload. These pages will contain instructions that list the different methods of data upload, storage and simple database selection list-box for the user to enter their method of choice and a button to notify the curators that they have completed the data upload.

2.2.2.1.5.12 Curation, Annotation, and Final Review. These pages will have instructions on what is happening to the submission during this period that may be taking longer to process. Once this stage is completed the user may not edit any of the uploaded information on the page, additionally movement in the submission workflow is restricted. The user may only advance to the curation pages in the next phase of the workflow.

2.2.2.1.6 Published. This page is the last page in the submission process that the data contributor will see. This page will lead the user to another page to view statistical summary information on the project.

2.2.2.1.7 Browse Biomedical Informatics Research Network Data Repository Page. This is the general page that will enable the user to view all the published data

submissions and search through the listings of the projects. This page will be available to all the users including guest users who may not have an account (Figure 2.21).

2.2.2.1.8 View Published Biomedical Informatics Research Network Data Repository Entry Page. This page re-displays all the data that was obtained throughout the course of the data contributor's submission to the BDR, once the submission has made it all the way through the submission workflow and approved for publication. The submission BDR entry link is published to the browse BDR page and viewable on this page (Figure 2.22).

2.2.2.1.9 My Profile Page. This page is available to log in users and it displays all the users profile information plus providing a link to a page to edit the profile information (Figure 2.24).

2.2.2.1.10 Help Page. This page provides basic knowledge for the submission process, documents that may be needed, and contact information to the system administrators and the curators to aid in any questions the user may have.

- This page is a general contact page that allows the user to contact either the curators or the system

administrators if there is a persistent problem.

Contact is through email or phone communication.

- Overview Instructions - This is a general outline of the submission process.
- Minimum Requirements Checklist - This is a list of different requirement for the submission to be processed correctly. This will be a document in tabular form that is retrieved from the database.
- IRB Template Overview - This page contains all the pertinent documents that are related to obtaining approval from the IRB and also instructions on submitting the documentation back to the BIRN Data Repository curators.
- Data Contributor Agreement Page - This is a simple formatted text page that enables the users to review the terms and conditions of the agreement at their own convenience.

2.2.2.1.11 Administration Page. This page allows the curator to print out a formatted submission report to present to the IRB, create alerts for a submission, and control the movement of the submission through the phases of the workflow (Figure 2.25).

2.2.2.2 Hardware Interface. The system will have a server with an NIC connected to the Internet. This will supply the necessary interfaces to connect with the software of the client. The client will need an NIC or modem in order to connect to the Internet to access the application server.

2.2.2.3 Software Interfaces. Software interfaces are provided in Java 1.5 APIs, Portlet API, Servlet API, JSP APIs, and GridSphere 2.1 APIs. In addition, users need a Java Script compatible web browser.

2.2.2.4 Communication Interfaces. The communication will take place within Web browsers, either Mozilla Firefox or Safari.

2.2.2.5 Memory. The system when deployed on the server has the minimum memory requirement of 512MB to run and a memory requirement of 1GB for optimal performance. The memory requirement of the client users computer, 128MB, is the same as for all Internet applications that needed to run. This includes the Web browser, either Mozilla Firefox or Safari, and any plug-in that is needed in the system for the user to access the system.

2.2.2.6 Operations. The user will access the BIRN Data Repository through the Internet. The session can be

terminated if the user is inactive for a certain period of time, and subsequently accessing any pages will be redirected to the initial login page. The system will be maintained on an Intel based server and will need to be up 365 days of the year.

2.2.3 Product Functions

2.2.3.1 Create a New Biomedical Informatics Research Network Data Repository Submission of Data Sets From an Existing Biomedical Informatics Research Network Project or From an External Source. Users will be guided through a elaborate workflow that will aid the user in supplying all relevant information to the project that will reduce the time required to review the data prior to publication. The User will be asked in the initial pages of the submission workflow if the data they are submitting is related to an existing BIRN project. The following are fields that can be filled in the submission process:

- Submission Title
- Abstract
- Related Keywords
- Investigators & Co-Authors
- Affiliated Institutions
- Sponsors

- Funding Sources/Grant Numbers
- Primary Publications
- Additional Publications
- Declare if the study contains human data
- Data Types
- Image Types
- Sample Size
- Represented Sex
- Age range of subjects from high to low plus the time metric
- Diagnostic groups represented in the sample

2.2.3.2 Submission Tracking. Allow a logged in user the ability to check the progress of their submissions. This includes a general overview of each submission and a more detailed drill down which includes the submission project log, alerts, etc.

2.2.3.3 Communicate With the Biomedical Informatics Research Network Data Repository Curators. Allow the data contributor and the curators to communicate back and forth with ease on the various aspects of the submission this it may be problems or progress topics.

2.2.3.4 Browse the Published Biomedical Informatics Research Network Data Repository Entries. Public users

without a login for the BDR can browse all the published submissions.

2.2.3.5 View the User Profile. A user may view the different attributes of their account and edit the values to the fields.

2.2.3.6 Obtain Help Information. Users can obtain help information from the website via download of instructional documentation or through phone/email support.

2.2.3.7 Allow Curators to Perform Administrative Functions. An administrative user should have the ability to print submission reports, correspond with a owner of a submission, create alert messages for those users, and move a submission forward or backwards in the latter half of the workflow.

2.2.4 User Characteristics

2.2.4.1 Biomedical Informatics Research Network Data Contributors. These are existing users of BIRN who want to contribute their data. They already have a login and possibly existing projects that are already present.

2.2.4.2 Guest Data Contributors. These users may want to contribute their data however they are not an existing member of BIRN and are required to register before having the ability to create any submissions.

2.2.4.3 Guest Users. Anonymous Internet users that may want to browse through the publicized datasets or with simply supplying the BDR with the users email address the user may download the full data set.

2.2.4.4 Curators. These are the users that will have privileged accounts, which will enable them to print out reports on the submission, alter the stage the submission is in to move it forward or backwards in the process, and perform various administrative tasks.

2.2.5 Constraints

Researchers must provide their full name, email, phone number, and affiliated organization in order to utilize the full capabilities of the BDR. There are no constraints to browse the repository.

2.2.6 Assumptions and Dependencies

The assumption is made that the user who utilizes the system enters the correct data, has all necessary client-side scripting enabled, and there are no applicable hardware limitations.

2.2.7 Apportioning of Requirements

No requirements specified in this system can be delayed until a future version without prior permission of the client.

2.3 Specific Requirements

2.3.1 External Interfaces

This is a detailed description of all the inputs and outputs of the BIRN Data Repository. Throughout this section there are references to a default user layout this is a standardized layout that is used on most of the pages in the system when a user is logged in. The default user layout is intended to aid the users so that they will be familiar with the most common actions and have quick access to navigate the site to perform these actions. The default user layout divides the page into two columns, the first column will contain the 'My Submissions' and 'BDR Information' portlets and does not vary at all. The wider second column will contain a portlet specific to that page being used (Figure 2.6). The 'My Submissions' portlet contains a table that displays the title, progress, and alert information for all the submission that have been created by the user. Each individual row will link to the corresponding 'Submission Specific Overview' page for that submission. Below the 'My Submission' portlet is another portlet frame that contains the 'BDR Information' portlet. This portlet simply displays four links that lead the user to the corresponding documentation for the BDR Data Use

Agreement, the BDR Data Contributors Agreement, the Data Submission Overview, and a Glossary of terms used throughout the system.

2.3.1.1 Login Page. This is the start page for the system, it is divided into two main columns. The wider first column contains the welcome text that explains the function of the BDR. To the right is a portlet frame that contains a login button and two text input fields, which are for the username and the password. If the login is

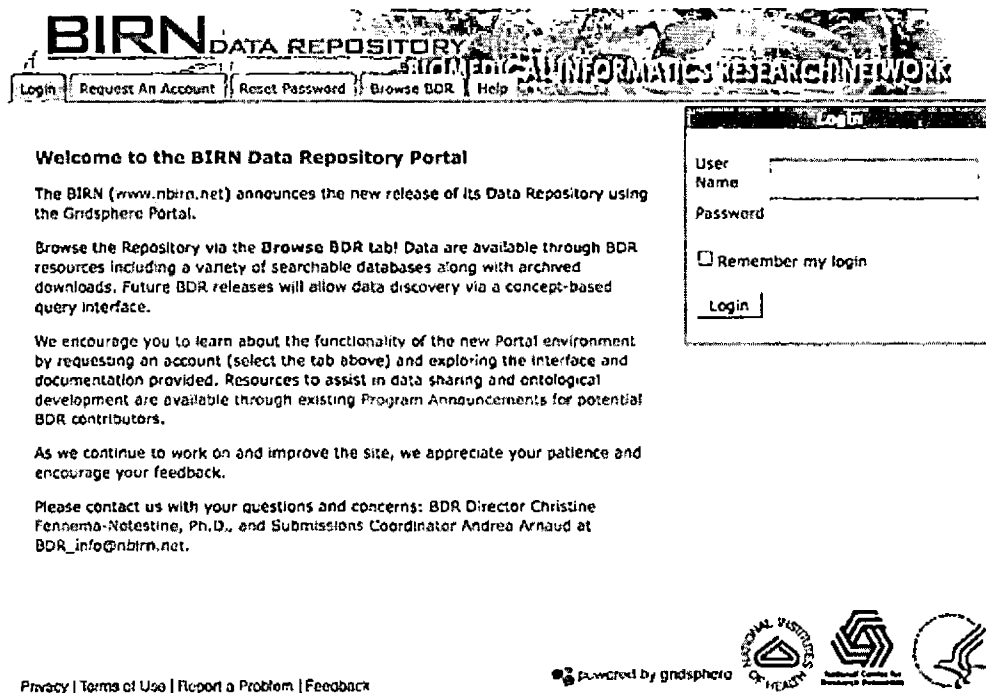


Figure 2.5. Login Page

successful the My Account page will be displayed otherwise a error message will be displayed above the username field (Figure 2.5).

2.3.1.2 My Account Page. This is the first page that would be displayed to a logged in user. The user will see three portlets arranged into two columns, this page follows the default user layout. In the left column the user will find the 'My Submissions' and 'BDR Information' portlets. The second column contains the 'Welcome' portlet, it

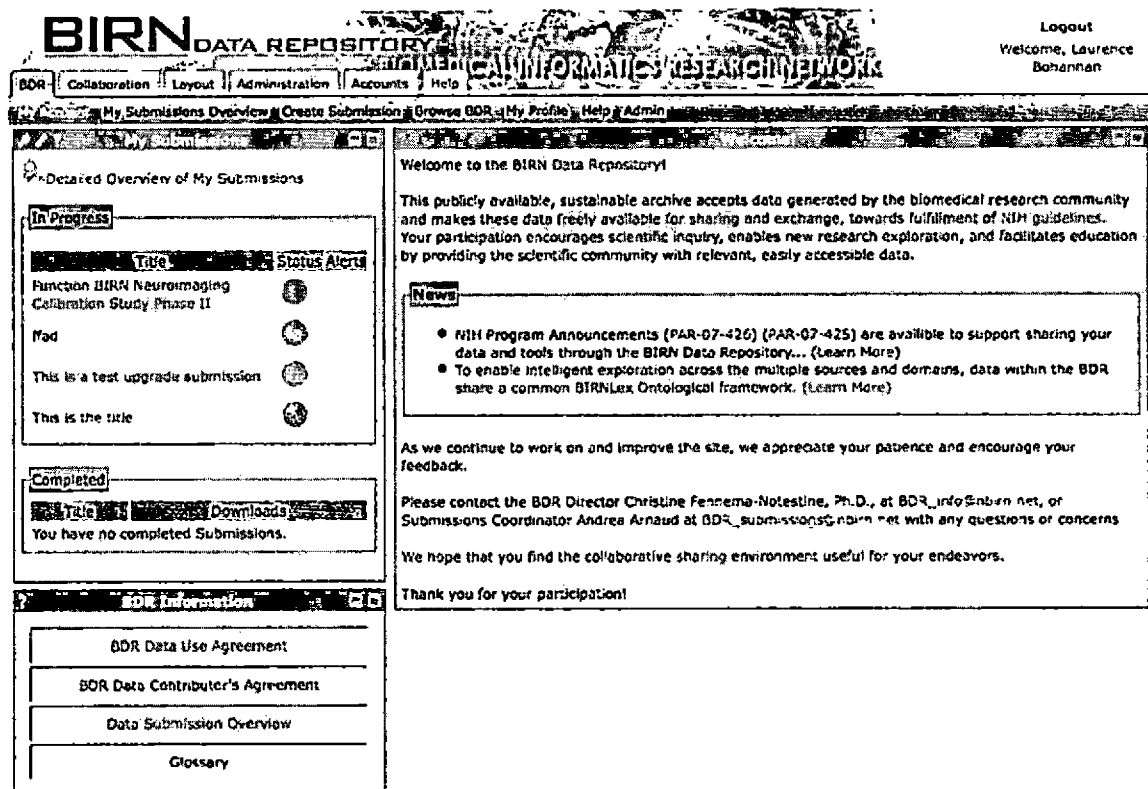


Figure 2.6. My Account Page

display's some more detailed information of the purpose and helpful contacts that a user may find useful (Figure 2.6).

2.3.1.3 My Submissions Overview. This page will be a single portlet that will display in a list of all the submissions that the user has created. Each individual listing will contain general information for that particular submission entry; title, accession number, primary contact, contributing investigator name and email. Additionally a graphical display of the current progress

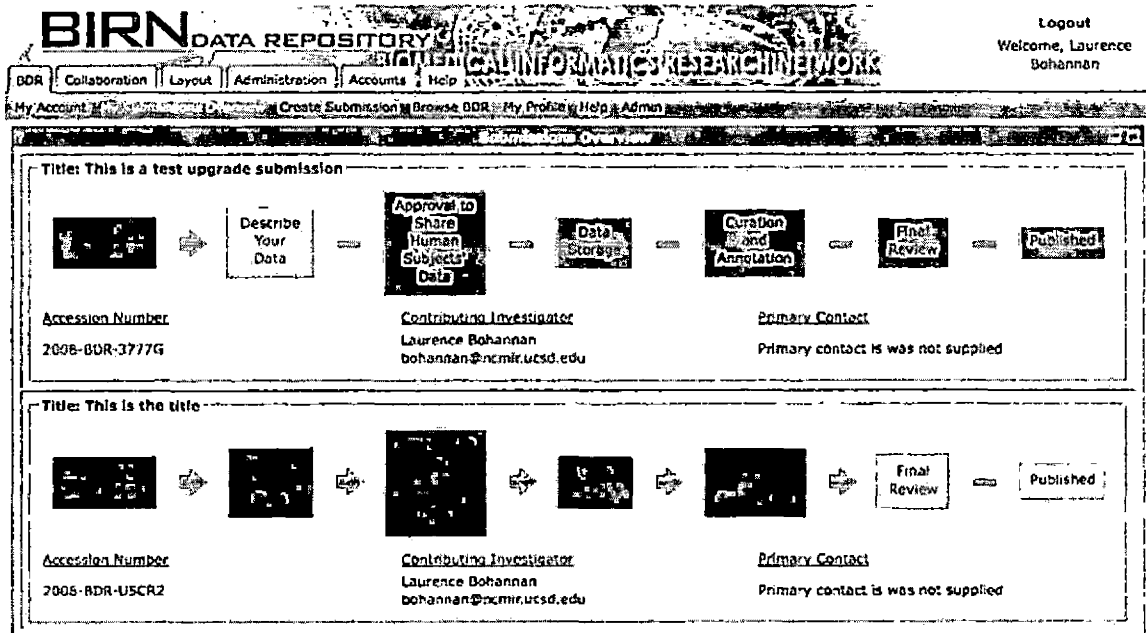


Figure 2.7. My Submissions Overview Page

and status of the submission throughout the pipeline is shown above the general information (Figure 2.7). The user can click on one of the entries to lead them to the 'Submission Specific Overview' for that particular submission.

2.3.1.4 Submission Specific Overview. This page will contain a single portlet that will display a graphical representation of the progress and status of the selected submission through the pipeline. To the bottom left of the graphic will be where the submission history and alerts are displayed. To the bottom right will be a link of actions that the user can perform, Continue Submission, Contact BDR, provide feedback, and a link to the collaborative workspace (Figure 2.8).

BIRN DATA REPOSITORY **CLINICAL GENOMICS RESEARCH NETWORK**

Logout
Welcome, Laurence Bohannon

BDR | Collaboration | Layout | Administration | Accounts | Help

My Account | Create Submission | Browse BDR | My Profile | Help | Admin

Title: This is the title

Accession Number	Contributing Investigator	Primary Contact
2008-BDR-U5CR2	Laurence Bohannon bohannon@ncmir.ucsd.edu	Primary contact is was not supplied

Alerts:
There are no user alerts

History of Submission:

- 10/22/2008 at 02:22:53 PM PDT: BDR Admin has advanced the submission to the 'Final Review' stage
- 10/22/2008 at 02:22:46 PM PDT: Admin has received your notification message
- 10/22/2008 at 02:22:43 PM PDT: BDR Admin has advanced the submission to the 'Data Curation and Annotation' stage
- 10/22/2008 at 02:22:24 PM PDT: 'Data Storage' has been completed
- 10/22/2008 at 02:22:24 PM PDT: Submission has completed the database upload
- 10/22/2008 at 02:21:43 PM PDT: BDR Admin has advanced the submission to the 'Data Storage Completion' step, the last step of the 'Data Storage' stage
- 10/22/2008 at 02:21:40 PM PDT: Admin has received your notification message
- 10/22/2008 at 02:21:11 PM PDT: 'Requirements for Data Storage' have been submitted.

Continue Submission

Contact BDR about this submission

Please provide feedback to the BDR

Collaborative Workspace

Figure 2.8. Submission Specific Overview Page

2.3.1.5 New Submission Initiation Page. This page follows the default user layout. In the second column is the 'Create New Submission' portlet, the user will see the 'New Submission Initiation' view by default (Figure 2.9). When the user clicks on the 'Begin New Submission' button they are taken to the second step of the process, which is the 'General Information' view.

BIRN DATA REPOSITORY **BIOINFORMATICS RESEARCH NETWORK**

Logout
Welcome, Laurence Bohannon

BDR Collaboration Layout Administration Accounts Help

My Account My Submissions Overview Browse BDR My Profile Help Admin

My Submissions

Detailed Overview of My Submissions

In Progress

Title	Status Alerts
Function BIRN Neuroimaging Calibration Study Phase II	
This is a test upgrade submission	
This is the title	
ifad	
Test Submission for 11/5/2009	
test 11/9/2009	

Completed

Title Alerts Downloads

You have no completed Submissions.

Submission Information

- Overview
- Detailed Overview
- Associated Metadata
- BDR Data Contributor's Agreement
- BDR Data Use Agreement
- IRB Template for Sharing Human Data
- Glossary

Create New Submission

During the submissions process, as the Contributing Investigator responsible for the data to be shared, you will:

- create a New Submission from your BDR Portal Account, although you may choose to designate someone else as the Primary Contact who will oversee the submission,
- provide a description of your data and the original study,
- provide verification of IRB approval for the sharing of human data, with the guidance of the BDR staff,
- provide associated metadata to sufficiently describe imaging data, to ensure that shared data is meaningful and useful,
- assist the BDR staff to ensure that data are provided in readable formats that, when necessary, are sufficiently de-identified,
- work with the BDR to select an appropriate database resource and to annotate shared data to the BIRNlex/Ontological framework for intelligent search and integration with other data sources, and
- make these data publicly available in a reasonable time frame, as determined by the nature of the data and collaborative investigations.

[Begin New Submission](#)



Figure 2.9. New Submission Initiation Page

2.3.1.6 General Information Page. This page adheres to the default user layout. The second column contains the 'General Information' view of the 'Create New Submission' portlet. The top of the portlet contains a section with

instructions for the page. Below the instruction section are text entry fields for the submission title, study description, keywords, and primary contact information for the submission. Additionally, there is a selection list of pre-selected keywords, an option box to select if the submission is associated to a previously existing project, a text box to display the data contributors agreement, and a checkbox for the user to acknowledge that they have read the contributing data agreement (Figure 2.10). Once the user has completed all of the required fields and agreed to the data contributor's agreement they may click on the 'Next' button (Figure 2.11). This will lead to the third step of the process, which is the General Information verification view.

Launch New Submission

Instructions: Page 1 of 7

Please provide submission title, study description, keywords that describe your data, and agree to the terms and requirements outlined in the **BDR Data Contributor's Agreement**. You may also designate someone else as a **Primary Contact**.

- Required Fields

• Submission Title: _____

• Study Description / Abstract (Maximum of 5000 characters):

Keywords:

- Select at least 3 keywords that best describe your data:

Which dictionary would you like to use? **BIRNLex**

Search
 Keywords: _____

Note: Please wait while the ontology loads

Unselected		Selected
(R)N6-phenylisopropyladenosine		
11-dehydro-thromboxane B2		
129		
129P1/RaJ	<input type="button" value="Add >>"/>	
129P3/J	<input type="button" value="<< Remove"/>	
129P3/J/EmsJ		
129P4/RrRKJ		
129S1/SvImJ		
129S6/SvEvTac		
129T2/SvEms		

Don't see all of the keywords you like? Add more of your own. _____

Contact Information:

Name of Contributing Investigator: Laurence Bohannon

E-mail of Contributing Investigator: bohannon@ncmir.ucsd.edu

Institution/Organization: UCSD BIRN-CC

Phone Number: 858 822-0718

Name of Primary Contact (if different from Contributing Investigator): _____

Primary Contact e-mail: _____

Project Information:

- Is this study related to an existing submission or BIRN project? (if 'Yes' please select the project from the list below)

No / Yes

Figure 2.10. General Information Page

BDR Data Contributor's Agreement:

* **BDR Data Contributor's Agreement:** (PDF Format)

BIRN Data Repository

Data Contributor's Agreement

To further enable a collaborative research environment, the BDR archives data generated by the biomedical research community and makes these data freely available, towards fulfillment of NIH guidelines (http://grants.nih.gov/grants/policy/data_sharing/). Your participation will encourage scientific inquiry, enable new research exploration, and facilitate education by providing the scientific community with relevant, easily accessible data. Shared data are available to the public with provision of an active e-mail address and acceptance of the standard BDR Data Use Agreement.

I have read, understand, and agree to the terms and requirements outlined in the **BDR Data Contributor's Agreement.**

Cancel
NEXT

Figure 2.11. General Information Page (Continued)

2.3.1.7 General Information Verification Page. This page adheres to the default user layout; the second column contains the 'General Information Verification' view of the 'Create New Submission' portlet. The top of the portlet frame has a section that contains instructions for the user. Below the instructions is another section that contains all the values that were entered by the user from the previous page. The user has the option to click one of two buttons at the bottom of the page, 'Edit Information' and 'Launch Submission' (Figure 2.12). The 'Edit Information' button will lead the user back to the previous 'General Information' view, the 'Launch Submission' button

will lead the user to the fourth step in the process, which is the 'Accession Number Confirmation' page.

2.3.1.8 Accession Number Confirmation Page. This page adheres to the default user layout. The second column contains the Accession Number Confirmation view of the

The screenshot shows a web browser window with the title 'Create New Submission'. The page content is as follows:

Instructions: Page 2 of 7
Please verify the information you have provided for accuracy. To make changes, click **Edit Information**. Once you have verified the information, click **Launch Submission** to proceed.

New Submission Information Verification:

Submission Title: BDR Demo Submission

Study Description / Abstract:
This is the description.

Keywords: keyword 3
keyword 1
keyword 2

Contributing Investigator: Laurence Bohannon
Contributing Investigator e-mail: bohannon@ncmir.ucsd.edu
Organization: UCSD BIRN-CC
Phone Number: 858 822-0718
Primary Contact:
Primary Contact e-mail:
Related to an existing BIRN project: No

Note: Please contact us if you need to modify any information from this stage after you have completed it.

[Edit Information](#) [Launch Submission](#)

Figure 2.12. General Information Verification Page

Create New Submission portlet. The page contains four short paragraphs of information pertaining to the tracking of the

submission using the accession number that was provided on the page. Below the information section are two buttons, 'Exit' and 'Continue' (Figure 2.13). The 'Exit' button will

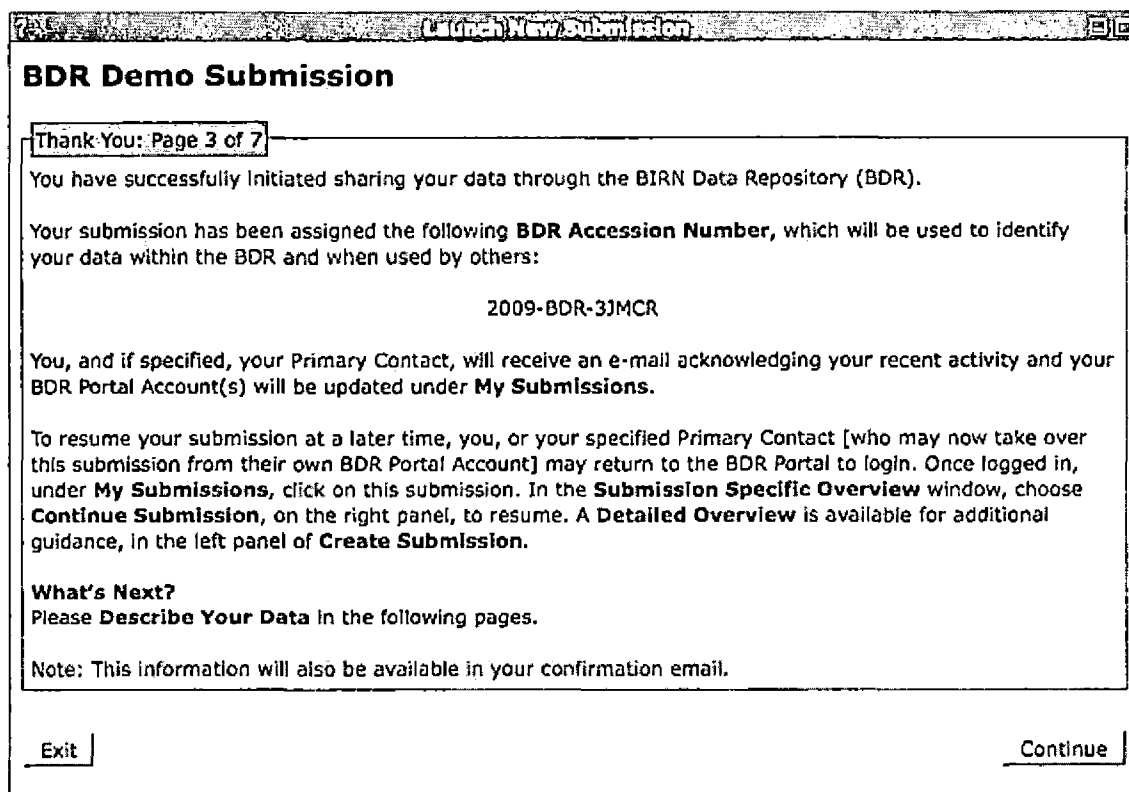


Figure 2.13. Accession Number Confirmation Page

lead the user back to the initial page view of the 'Create New Submission' portlet. The 'Continue' button will lead to the fifth step in the process, which is the 'High Level Information' page.

2.3.1.9 High Level Information Page. This page adheres to the default user layout. The second column contains the 'High Level Information' view of the 'Create New Submission' portlet. The top of the portlet frame has a section that contains instructions for the user. Below the instructions are seven sections: affiliations, investigators, sponsors, funding sources, primary publication, additional publication, and figure. The figure section contains a file upload field, a text input field for a caption, and a single button that is labeled 'Upload'. When the user clicks the 'Upload' button the supplied image file will be displayed in the place of the text input fields in that section. The affiliations section contains an empty three column table. The columns are for the affiliation name, URL, and a space provided to perform a delete action on the particular row in the table. Below the table are the two text input fields which allow the user to populate the table for that section, one for

BDR Demo Submission

Instructions: Page 4 of 7

Please list affiliations, investigators and contributors, sponsors, funding sources, and relevant citations. To load each entry under each category, click the **Add** button. Repeat for additional entries. If desired, add a figure and caption, then choose **upload**.

* Required Fields

Affiliations:

Name	URL	delete
UCSD BIRN-CC		

Name: URL: **Add**

Investigator(s) & Contributor(s):

Name	Email	Affiliation
There are no entries.		
Laurence Bohannon (Contributing Investigator)	bohannon@ncmir.ucsd.edu	

Add

Sponsor(s):

Name
There are no entries.

Name: **Add**

Funding Source(s)/Grant Number(s):

Source	Grant Numbers
There are no entries.	

Source: Grant Numbers: **Add**

Figure/Caption:

Notice: Do not upload copyrighted material; all data submitted will become public, without any restrictions for use other than those stated in the BDR Data Use Agreement.

Add Figure File:

Browse...

Add Caption (Maximum of 250 characters):

upload

Primary Publication(s)/Original Article(s):

Full Citation	ID	Type of ID
There are no entries.		

Full Citation:

ID: Type of ID: **Add**

Additional Related Publication(s):

Full Citation	ID	Type of ID
There are no entries.		

Full Citation:

ID: Type of ID: **Add**

Back **Exit** **Next**

Figure 2.14. High Level Information Page

affiliation name and one for the URL. Following the text input fields is a button labeled 'Add', once this is clicked the values in the accompanying fields will be present in the table above. All other sections on this page behave in a similar manner. Each just differs on the number of input fields for the particular table in that section. The sponsor section contains a two-column display table above one text input field for the sponsor name and a add button. The funding sources section contains a three-column display table above two text input fields for the source and grant numbers. The primary and additional related publications sections are identical in format, each contain a four column display table above text input fields for the full citation, ID number, and a selection box for the type of ID (Figure 2.14).

2.3.1.10 High Level Data Page. This page adheres to the default user layout. The second column contains the 'High Level Data' view of the 'Create New Submission' portlet. The top of the portlet frame has a section that contains instructions for the user. Below the instructions are seven sections: human subjects declaration, data types, image types, data information, diagnostic group represented

Describe Your Data

BDR Demo Submission

Instructions: Page 5 of 7

Please specify if data is human or not; specify data and image types to be shared; and provide requested information to describe the sample. To save your entered information and advance to the next screen, choose **Next**.

- **Required Fields**

Human Subjects' Declaration:

• Does this submission contain human subjects' data?

Data Types:

Unselected Data Types

- Imaging
- Demographics
- Behavioral/cognitive measures
- Clinical measures
- Microarray
- 2D Images
- Cell-fills

Selected Data Type(s)

Other Data Type:

Image Types:

Unselected Image Types

- Anatomical MR
- Functional MR
- Diffusion
- Microscopy

Selected Image Type(s)

Other Image Type:

Data Information:

• Number of individual cases in sample:

• Sex represented in sample:

• Specify range of age in sample (low - high): -

Note: To specify embryonic and/or postnatal, add an "e" or "p" to the low and high values. (e.g., 16e-17p days)

Other:

Diagnostic groups represented in sample:

Unselected Diagnostic Groups

- Healthy Individuals
- Schizophrenia
- Alzheimer's disease
- Mild Cognitive Impairment

Selected Diagnostic Group(s)

Other Diagnostic Group:

Figure 2.15. High Level Data Page

in sample, technical contact, and usage recommendations and requirements. The human subject declaration section has a prompt asking the user if this submission contains any human subject data and options box with the value of 'Yes' or 'No'. The 'Data Types' section contains two multi-select list option boxes, 'Selected Data Types' which is empty and 'Unselected Data Types' is pre-populated with the values; 'Imaging', 'Demographics', 'Behavioral/Cognitive Measures', 'Clinical Measures', 'Microarray', '2D Images', and 'Cell-Fills'. Below the list box is a text input field for other user entered values, these are populated to the 'Selected Data Types' list box once the add button is clicked. The 'Image Types' section contains two multi-select list option boxes, 'Selected Image Types' is empty while 'Unselected Image Types' is pre-populated with the values; 'Anatomical MR', 'Functional MR', 'Diffusion', and 'Microscopy'. Below the list box is a text input field for other user entered values, these are populated to the 'Selected Image Types' list box once the add button is clicked (Figure 2.15). The Data Information section contains text input fields for the number of individual cases in the sample. A selection box for the sex represented in the sample, two text input fields to specify the high and low of the age range in that

sample, a selection box for the units of the age range, and an additional text input field to allow the user to specify any embryonic or parental values. The Diagnostic Groups Represented in the Sample section contains two multi-select list option boxes, 'Selected Diagnostic Groups' which is empty and 'Unselected Diagnostic Groups' which is pre-populated with the values; 'Healthy Individuals', 'Schizophrenia', 'Alzheimer's disease', and 'Mild Cognitive Impairment'. Below the list box is a text input field for other user entered values, these are populated to the 'Selected Data Types' list box once the add button is clicked. The 'Technical Contact' section contains two text input fields to enter the name and email of the users technical contact. The 'Usage Recommendations and Requirements' section contains a text area for the user to enter specific instructions. Below the usage section are three buttons 'Back', 'Exit', and 'Continue' (Figure 2.16).

Technical Contact:		
Name:	<input type="text"/>	
E-mail:	<input type="text"/>	
Usage Recommendations and Requirements:		
Please describe any recommended tools, any system requirements, or any additional materials that may be needed. (Maximum of 5000 characters):		
<input type="text"/>		
<input type="button" value="Back"/>	<input type="button" value="Exit"/>	<input type="button" value="Next"/>

Figure 2.16. High Level Data Page (Continued)

2.3.1.11 High Level Data Confirmation Page. This page adheres to the default user layout. The second column contains the 'High Level Data Confirmation' view of the 'Create New Submission' portlet. The top of the portlet frame has a section that contains instructions for the user. Below the instructions is another section that contains all the values that were entered by the user from the previous page. The user has the option to click one of three buttons at the bottom of the page 'Edit Information', 'Submit', and 'Exit' (Figure 2.17). The 'Edit Information' button will lead the user back to the previous 'High Level

Data' view; the 'Submit' button will lead the user to the fourth step in the process, which is the 'Thank You' page.

Describe Your Data

BDR Demo Submission

Instructions: Page 6 of 7

Please verify the information you have provided for accuracy. To make changes, choose **Edit Information**. Once you have verified the information, click on **Submit**.

Information Verification:

Affiliated Institution(s) & URL(s):	UCSD BIRN-CC
Investigator(s) & Contributor(s):	Laurence Bohannon (Contributing Investigator) bohannon@ncmir.ucsd.edu UCSD BIRN-CC
Sponsor(s):	
Funding Source(s)/Grant Number(s):	
Figure:	
Caption:	
Primary Publication(s)/Original Article(s):	
Additional Related Publication(s):	
Data Types:	Microarray
Image Types:	Functional MR
Number of individual cases in sample:	32
Sex represented in sample:	Female
Specify range of age in sample (low - high):	23 - 32 Months
Diagnostic groups represented in sample:	Mild Cognitive Impairment
Name of Technical Contact:	
Technical Contact e-mail:	
Usage Recommendations and Requirements:	

Note: Please contact us if you need to modify any information from this stage after you have completed it.

[Edit Information](#) [Submit](#)

[Exit](#)

Figure 2.17. High Level Data Confirmation Page

2.3.1.12 Thank You Page. This page adheres to the default user layout. The second column contains the 'Thank You' view of the 'Create New Submission' portlet. This page

The screenshot displays the BIRN Data Repository user interface. At the top, the header includes the BIRN logo, navigation tabs (BDR, Collaboration, Layout, Administration, Accounts, Help), and a user profile section with 'Logout' and 'Welcome, Laurence Bohannon'. Below the header, a navigation bar contains links like 'My Account', 'My Submissions Overview', and 'Browse BDR'. The main content area is divided into two columns. The left column features a 'Detailed Overview of My Submissions' portlet with an 'In Progress' section listing submissions like 'Function BIRN Neuroimaging Calibration Study Phase II' and 'BDR Demo Submission', and a 'Completed' section stating 'You have no completed Submissions.'. Below this is a 'Submission Information' portlet with a list of links: Overview, Detailed Overview, Associated Metadata, BDR Data Contributor's Agreement, BDR Data Use Agreement, IRB Template for Sharing Human Data, and Glossary. The right column displays a 'BDR Demo Submission' 'Thank You' message, indicating successful completion of the 'Describe Your Data' stage and that the submission is under review. It includes a 'What's Next?' section advising on IRB approval and data storage options, and buttons for 'Exit', 'Explore Data Storage Options', and 'Continue'.

Privacy | Terms of Use | Report a Problem | Feedback



Figure 2.18. Thank You Page

contains information about what will be the next step in the process, below the information are three buttons 'Exit', 'Continue', and 'Explore Data Storage Options' (Figure 2.18). The 'Explore Data Storage Options' leads to documentation of the various storage options for the user.

2.3.1.13 Internal Review Board Instruction Page. This page adheres to the default user layout. The second column contains the 'IRB Instructions' view of the 'Create New Submission' portlet. This page contains information pertaining to the steps and requirements of the upcoming IRB Documentation Upload. Below the information are three buttons 'Exit', 'Continue, and 'Explore Data Storage Options' (Figure 2.19). The 'Explore Data Storage Options' button leads to documentation of the various storage options for the user.

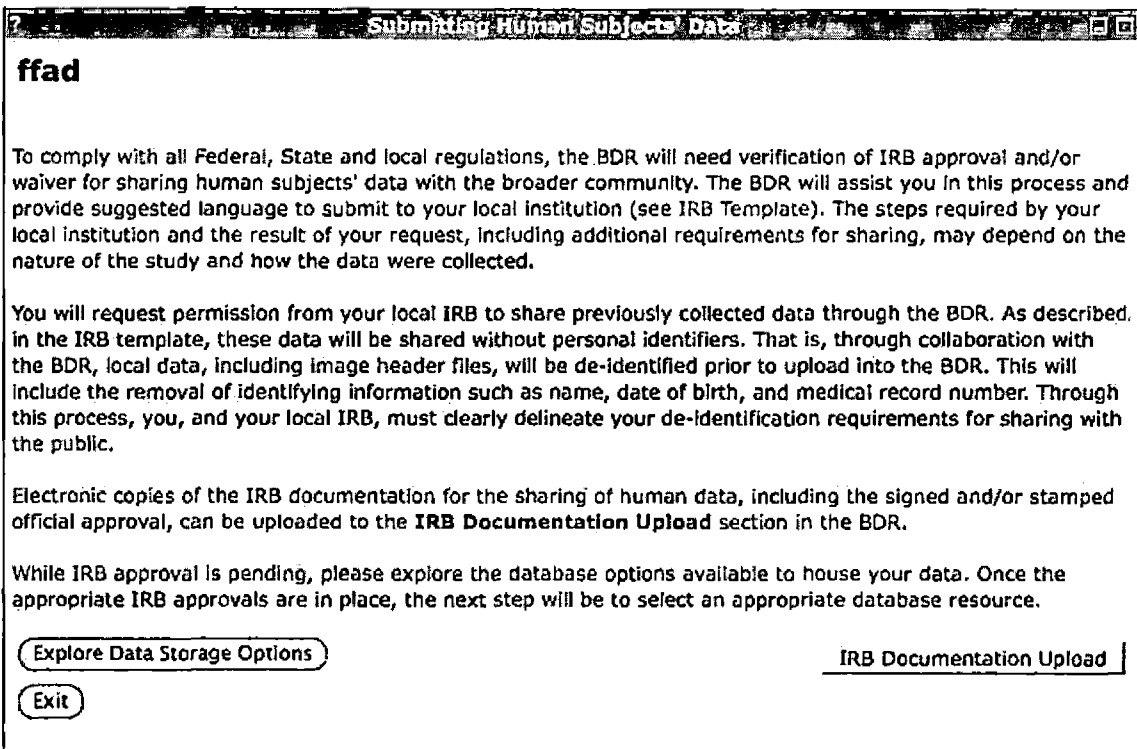


Figure 2.19. Internal Review Board Instruction Page

2.3.1.14 Internal Review Board Template Upload Page.

This page adheres to the default user layout. The second column contains the IRB Template Upload view of the 'Create New Submission' portlet. The top of the portlet frame has a section that contains instructions for the user. Below the instructions is a section labeled 'Documentation Upload' it contains a text area for a description, a text input field to enter the title of a file, a select box to specify the expected file format, and a file browser button to upload the actual documentation. In the event that the user may

the actual documentation. In the event that the user may have more than one file to upload at the time of submission they may add and remove additional file entry fields by using the 'Add More Files' and 'Remove' links located at the bottom of the section. Once the user has completed filling in the information they can complete the upload by clicking on the 'Submit' button. If all inputs that were provided are correct, the files will be displayed below in the 'Submitted Documentation' section. If there was invalid input an error will be displayed. At the bottom of the page are three buttons 'Back', 'Exit', and 'Notify the BDR of Completion' (Figure 2.20). 'Back' will simply move the submission to the previous page in the submission process. 'Exit' will erase the submission session and return the

IRB Documentation Upload

ffad

Instructions:

1. Under **Documentation Upload**, please provide a description of the file(s) to be uploaded, title, file type, and file. Click on **Add More Files** to load additional files under the same description; then click on **Submit** to send them to the BDR.
2. Repeat step 1, for additional file(s) under a different description, as needed.
3. When **ALL** the necessary files for IRB approval have been submitted and your documentation is complete, click on **Notify BDR of Completion**.

* Required Fields

Documentation Upload:

Current Date: 11/18/2009

Description:

Title	File Type	File
_____	_____	_____

Add More Files - Remove

Submitted Documentation:

Time: Apr 21, 2009 7:06:38 PM

Description:

This is the upload description for the file set

Title	File Type	File	View
This is the title of file 2	pdf	file2.pdf	
This is the title of file 1	pdf	file1.pdf	

Figure 2.20. Internal Review Board Template Upload Page

user to the initial page of the new submission process. The 'Notify BDR of Completion' button will freeze the submission from movement, alert the administrators that

this step in the process has been completed and display a notification message at the top of the page.

2.3.1.15 Browse Biomedical Informatics Research

Network Data Repository Page. The public or a logged in user can view this page, if the user is logged in they would see the page separated into two columns. The first narrower column would contain the 'My Submissions' portlet.

The screenshot displays the BIRN Data Repository interface. At the top, the header includes the BIRN logo, navigation links (BDR, Collaboration, Layout, Administration, Accounts, Help), and a user greeting: "Logout Welcome, Laurence Bohannon". Below the header is a navigation bar with "My Account", "My Submissions Overview", "Create Submission", "My Profile", "Help", and "Admin".

The main content area is divided into two columns. The left column, titled "Detailed Overview of My Submissions", contains two sections: "In Progress" and "Completed". The "In Progress" section features a table with columns for "Title" and "Status Alerts". It lists three items: "Function BIRN Neuroimaging Calibration Study Phase II" with a status alert icon, "ifad" with a status alert icon, and "This is a test upgrade submission" with a status alert icon. Below this, the "Completed" section shows a table with columns for "Title" and "Downloads", with the message "You have no completed Submissions." below it.

The right column is titled "Currently Available Data". It contains a paragraph explaining that the BDR is a sustainable archive of data generated by the biomedical research community, available for sharing and exchange. It also states that all data are freely available to the public with the provision of an active e-mail address and acceptance of the standard BDR Data Use Agreement. Below this, there are two sections: "BIRN DATA REPOSITORY SOURCES:" and "CURRENTLY AVAILABLE DATA FROM BIRN-ENABLED COLLABORATIONS:". At the bottom right of the page, there are logos for "powered by gndsphere", "NATIONAL INSTITUTE OF HEALTH", and "National Center for Human Genome Research".

At the bottom left of the page, there are links for "Privacy", "Terms of Use", "Report a Problem", and "Feedback".

Figure 2.21. Browse Biomedical Informatics Research Network Data Repository Page

In the second column, the user would see the same BDR Browser portlet as a guest user but rather it would take up the whole width of the page for the guest. This portlet contains information about the BIRN Data Repository, a list of links to the various data sources, and a listing of all the approved entries in the BRIN Data Repository (Figure 2.21). If the user would click on one or the individual entries it would lead them to the 'Published BIRN Data Repository Entry' page.

2.3.1.16 Publish Biomedical Informatics Research Network Data Repository Entry Page. The public or a logged in user can view this page, if the user is logged in they would see the page separated into two columns, the first narrower column would contain the My Submission portlet. In the second column the user would see the same Published BIRN Data Repository View of the BDR Browser portlet as a guest user will but rather it would take up the whole width of the page for the guest. This portlet displays all the general and high-level information that was gathered by the system throughout the submission process: description, figure, accession number, technical contact, recommended usage, investigators and contributors, affiliated institutions, primary publications, and funding sources

(Figure 2.22). At the bottom of the page is a button labeled 'Download', this will take the user to the download page (Figure 2.23), which will provide a link for them to follow for download.

My Submissions

Detailed Overview of My Submissions

In Progress

Title	Status Alerts
Function BIRN Neuroimaging Calibration Study Phase II	<input type="radio"/>
rad	<input type="radio"/>
This is a test upgrade submission	<input type="radio"/>
This is the title	<input type="radio"/>
BDR Demo Submission	<input type="radio"/>

Completed

Title	Downloads
You have no completed Submissions.	

Function BIRN Neuroimaging Calibration Study Phase II

Description: Technological advances in functional and structural neuroimaging have revolutionized the biomedical investigation of neuropsychiatric illness. The tremendous potential that this methodology brings in advancing diagnostic and prognostic capabilities and in treatment of neuropsychiatric illness has as yet remained largely an unfulfilled promise. This potential has been limited by a number of financial, technological and cultural impediments. These impediments include the expense of acquiring imaging data that limits sample sizes that can be collected by individual investigators and a lack of cross site calibration and standardization methodologies that would enable coordinated multi-site acquisition for basic and clinical trials research. Most sites develop their own acquisition and analysis software algorithms and conduct independent studies that have produced interesting, but often conflicting, results in the literature. It has not been possible to disentangle the sources of these site differences (e.g., are they a consequence of differences in data acquisition and analysis algorithms, due to differences in study populations or a combination of the two). Some of these impediments could be overcome by a body of scientific knowledge describing the contribution of acquisition and analysis algorithms to the variance in the collected neuroimaging data. Some of these impediments could be overcome by creation of a database containing benchmark raw imaging data available to biomedical imaging laboratories developing advanced image analysis software for comparison of their new analysis algorithms to the current state of the art algorithms. More of these impediments could be overcome by access to sufficiently large samples of calibrated structural and functional imaging data from populations of patients and controls needed for matching to the study population of interest. Further impediments could be overcome by the development of consensus recommendations made by the experts in the research community for functional and structural neuroimaging data acquisition and analysis calibrations and standards. The Function Biomedical Informatics Research Network (FBIRN) is a trans-disciplinary consortium of research institutions integrating imaging, clinical, cognitive and genetic data, to achieve the following objectives: 1) Conduct longitudinal studies of schizophrenia in multiple centers of research excellence around the nation, developing within and between the centers the infrastructure and methodologies that seamlessly integrate imaging data, clinical and cognitive studies, and the gathering of genetic data into a commonly shared pool of knowledge; and, 2) Determine how these methodologies can be adapted into the larger research community and disseminated to the broader field of biomedical research.

Figure:



Caption: BOLD activation related to working memory task performance

Accession Number: 2008-BDR-6QF87

Technical Contact: Dingying Wei (dywei@AT@uci.edu)

Recommended Usage: Image data is provided in Nifti-1 format which can be read by most image viewing and analysis software packages. Related demographic and clinical data can be queried in a detailed manner through the FBIRN HID and/or downloaded as a comma separate variable file for import into any standard program that reads text files. Behavioral data is available as downloaded with related fMRI image files and are provided in a text format. Related documentation for all data will be available for download.

Investigators and Contributors: Christine Fennema (Contributing Investigator) [1], Ron Kikinis, M.D.[2], Gregory McCarthy, Ph.D.[null], Aysenil Belger, Ph.D.[null], Bruce Rosen, M.D., Ph.D.[4], Gary H. Glover, Ph.D.[5], Arthur W. Toga, Ph.D.[6], Gregory G. Brown, Ph.D.[7], Daniel S. OLeary, Ph.D.[8], Judy Ford, Ph.D.[11], Kelvin O. Lim, M.D.[9], Steven Potkin, M.D.[12].

Affiliated Institutions:

1. UCSD
2. Brigham and Womens Hospital
3. Duke University/University of North Carolina
4. Massachusetts General Hospital
5. Stanford University
6. University of California, Los Angeles
7. University of California, San Diego
8. University of Iowa
9. University of Minnesota
10. University of New Mexico
11. Yale University
12. University of California, Irvine

Primary Publications/Original Articles:

Figure 2.22. Publish Biomedical Informatics Research Network Data Repository Entry Page

● Friedman L, Stern H, Brown CG, Mathalon DH, Turner J, Glover GH, Gollub RL, Laurienti J, Lim KO, Cannon T, Greve DN, Bockholt HJ, Belger A, Mueller B, Doty MJ, He J, Wells W, Smyth P, Pieper S, Kim S, Kubicki M, Vangel M, Patkin SG. Test-retest and between-site reliability in a multicenter fMRI study. Hum Brain Mapp. 2007 Jul 17.

Funding Sources / Grant Numbers: NIH / NCRR: U24-RR021992

[Download](#)

[Privacy](#) | [Terms of Use](#) | [Report a Problem](#) | [Feedback](#)



Figure 2.23. Publish Biomedical Informatics Research Network Data Repository Entry Page (Continued)

2.3.1.17 My Profile Page. This page simply contains a portlet which display all the user information that is stored in the portal to the current user (Figure 2.24).

Laurence Bohannon
Information
User ID lbohannon
Salutation:
Organization UCSD BIRN-CC
Site BIRN Coordinating Center(BIRN-CC)
Domain ucsd-bcc
CV URL
Job Title Portal Programmer
Contact
Email bohannon@ncmir.ucsd.edu
AIM
Phone 858 822-0718
Fax
Address
Miscellaneous
About Me
Research Interests
Hobbies
Favorite Quote(s)
Edit Profile



Figure 2.24. My Profile Page

2.3.1.18 Administration Page. This page displays all the submissions that are in the system, when the admin

BDR | Collaboration | Layout | Administration | Accounts | Help

My Account | My Submissions Overview | Create Submission | Browse BDR | My Profile | Help

Accession #	Title	Status	Alerts
2008-BDR-M3M9Z	Mouse Model of Parkinson's Disease		
2008-BDR-67AFE	Dopamine transporter knockout mouse MRI		
2008-BDR-6QFR7	Function BIRN Neuroimaging Calibration Study Phase II		
2008-BDR-YBN1H	jdjfaljfal		
2008-BDR-XQGYS	MRM Reeler Mouse		
2008-BDR-WMB4G	MS Mouse Atlas		
2008-BDR-MZGZJ	MRM Reeler Mouse		
2008-BDR-6P5PF	dgedgsfdgs		
2008-BDR-88PCF	Function BIRN Neuroimaging Calibration Study Phase II		
2008-BDR-G3ZA9	MRM Reeler Mouse		
2008-BDR-EGJ3E	fdad		
2008-BDR-3777G	This is a test upgrade submission		
2008-BDR-RZU6E	MS Mouse Atlas		
2008-BDR-MMBNQ	khqjghkgykg		
2008-BDR-HNX89	fd		
2008-BDR-8A4Y7	sfdsg		
2008-BDR-ERXA3	Multi-site calibration		
2008-BDR-HKQWC	fdsgfdgsdfg		
2008-BDR-U5CR2	This is the title		
2008-BDR-6BEN4	Test		
2008-BDR-Y8FRT	Human Functional MRI Data: Traveling Subjects Study		

Figure 2.25. Administration Page

clicks on the submission listed they are lead to an administrative version of the submission specific overview (Figure 2.25).

2.3.2 Functional Requirements

All functions to be added to the BDR are detailed in section 2.

2.3.3 Performance Requirements

The BDR will support approximately all researchers on BIRN plus additional guest users use. The response time to view any page should be less than three seconds when accessing from the Internet. The response time for seeing any result after submitting an input or query usually should not exceed twice the length than viewing a page.

2.3.4 Logical Database Requirements

Postgres will be the primary database used for the BDR application.

2.3.5 Design Constraints

User interfaces will comply with standard Gridsphere UI style and format. The same look and feel will be adapted to the BDR.

2.3.6 Software Systems Attributes

2.3.6.1 Reliability. All contents and logs shall be generated dynamically and automatically so no human intervention is needed. The server shall be up twenty-four hours a day and seven days a week, with the exception that periodical system maintenance needs to be conducted

depending on the reliability of the hardware that the server is deployed on. The system should handle network packet loss smoothly. The system should not save inconsistent data or incomplete data into the application.

2.3.6.2 Availability. The BIRN Data Repository shall be accessible at any moment in time from the Internet. Maintenance time should be negligible due to the common project directory structure, In the current configuration with three commands the server can be shutdown, reinstall the application, and started back up again in less than a minute. This may also be affected by the speed of the server that the BDR is hosted on.

2.3.6.3 Security. To protect the application from malicious attacks from guest users and existing users extensive field validation methods will be preformed on all user inputs to protect against SQL injection and cross-site scripting attacks. All communication between the client and the server will be preformed on a https protocol typically any pages that require the use of a login will utilize this transport protocol.

2.3.6.3.1 Maintainability. The BDR mainly consists of Java classes, JSP pages, and Gridsphere specific XML configuration files which vary from general

portlet descriptors XML to hibernate mapping files. Since the BDR adheres to the basic directory and file layout for a Gridsphere default project important configuration files are placed into a common well known location that is easy to locate.

2.3.6.3.2 Portability. The Server-Implementation of BDR system shall be 100% portable since it will be written in Java, a proven portable language. The only determinant of how easily the BDR can be ported from a particular architecture to another is, if the server has the latest version of the Java Virtual Machine installed. The Client portion of the BDR will also be 100% portable since the system will be presented in HML pages, JavaScript, and style sheet, which are supported by most up-to-date Web browsers.

CHAPTER THREE

SOFTWARE DESIGN

3.1 System Design

The BIRN Data Repository portal is built upon a collection of various other portlet packages and Web services to achieve its functionality.

3.1.1 System Overview

The BDR portal was built with the concept of a widely distributed data grid collaborative project workspace in mind. This collaborative project allows an individual who belong to a project group to access shared data depending on the users privileges. The BDR portal is deployed on the Gridsphere 2.1 framework. Gridsphere enables developers to quickly implement and deploy their third party portlets. Gridsphere is a Java based JSR-168 complaint grid enabled portlet framework, which was developed, by Jason Novotny, Oliver Wehrens, and Michael Russel with initial funding from the European Commissions Gridlab project and Germany's D-Grid initiative [B6]. For Gridsphere to allow end-users to make use of Grid technologies, the Grid portlets Web application, which is also develop by the Gridsphere development team, must be installed. Grid portlets enables

users to upload their Grid credentials that they have obtained from GAMA and use them to access a variety of Grid services. The BDR portal is using a Postgres database to persist all information that is handled by a hibernate-based persistence manager.

3.1.2 Portal Framework

The BIRN Data Repository will use the GridSphere portal framework which is similar to the IBM WebSphere portal model in that it's central paradigm for providing access to resources is through a service based-architecture. There are four main elements in the system:

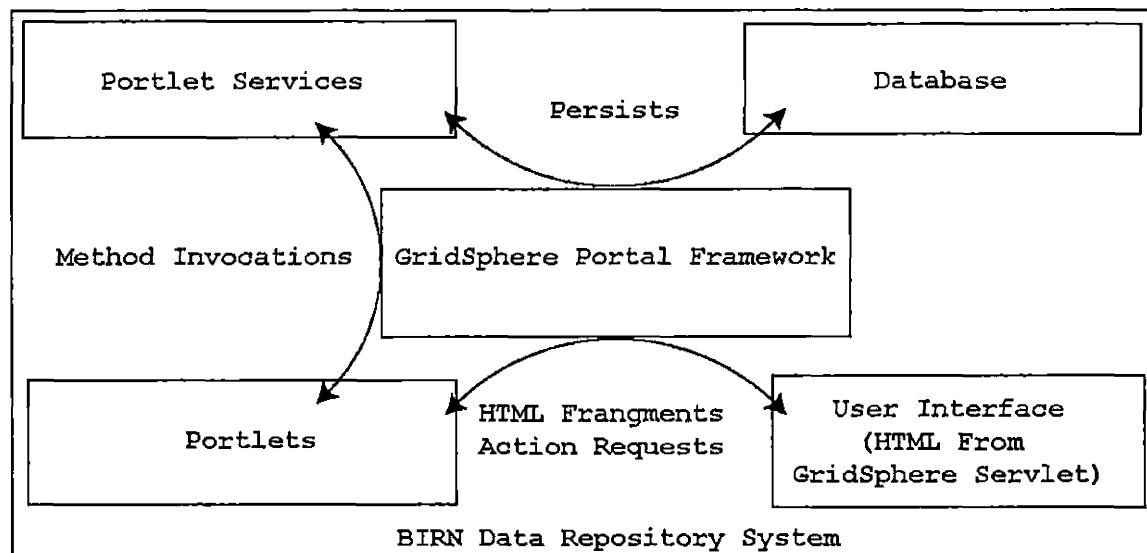


Figure 3.1. Biomedical Informatics Research Network Data Repository System Diagram

user interface, portlets, portlet services, and the database, see Figure 3.1.

3.1.3 Portlets

Portlets define Web application components with a well defined set of lifecycle methods much like the Java Servlet API as well as providing a visual interface to a content of service provider. Portlets are modular, reusable software components that may be developed independently of the general portal architecture and offers a specific set of operations [B10]. The portlets generate small fragments of html that are gathered by the portal to be combined with any other HTML fragments from other portlets on a particular layout view to render the page that is displayed to the user.

3.1.4 Portlet Services

Portlet services are designed to select the functions provided by portlets from the services with which they need to interact with. Portlet services provide an encapsulation of reusable business logic that may be reused by one or more portlets [B10]. For the BIRN Data Repository to function properly, it directly depends on eight services of which three are newly created services: a submission

manager, a specialized logging manager, and a download tracking service.

3.1.5 User Interface

GridSphere provides an advanced user interface library to support the development of portlets. However, it is not widely accepted as a mainstream standard with the portlet development community. When the GridSphere UI Beans are coupled with the GridSphere's ActionPortlet model, which is an extension of the GenericPortlet class, it greatly reduces the complexity for the retrieval and population of elements in request form and responses.

3.1.6 Database

The GridSphere Framework has a native support built in for a persistence manager service, which can be configured to run off a variety of databases, from a simple file in memory instance to a full stand alone Oracle server. Each of the systems three services access the database for storage and updating of objects. The database for the BIRN Data Repository contains 19 application specific tables (Figure 3.2 and Figure 3.3).

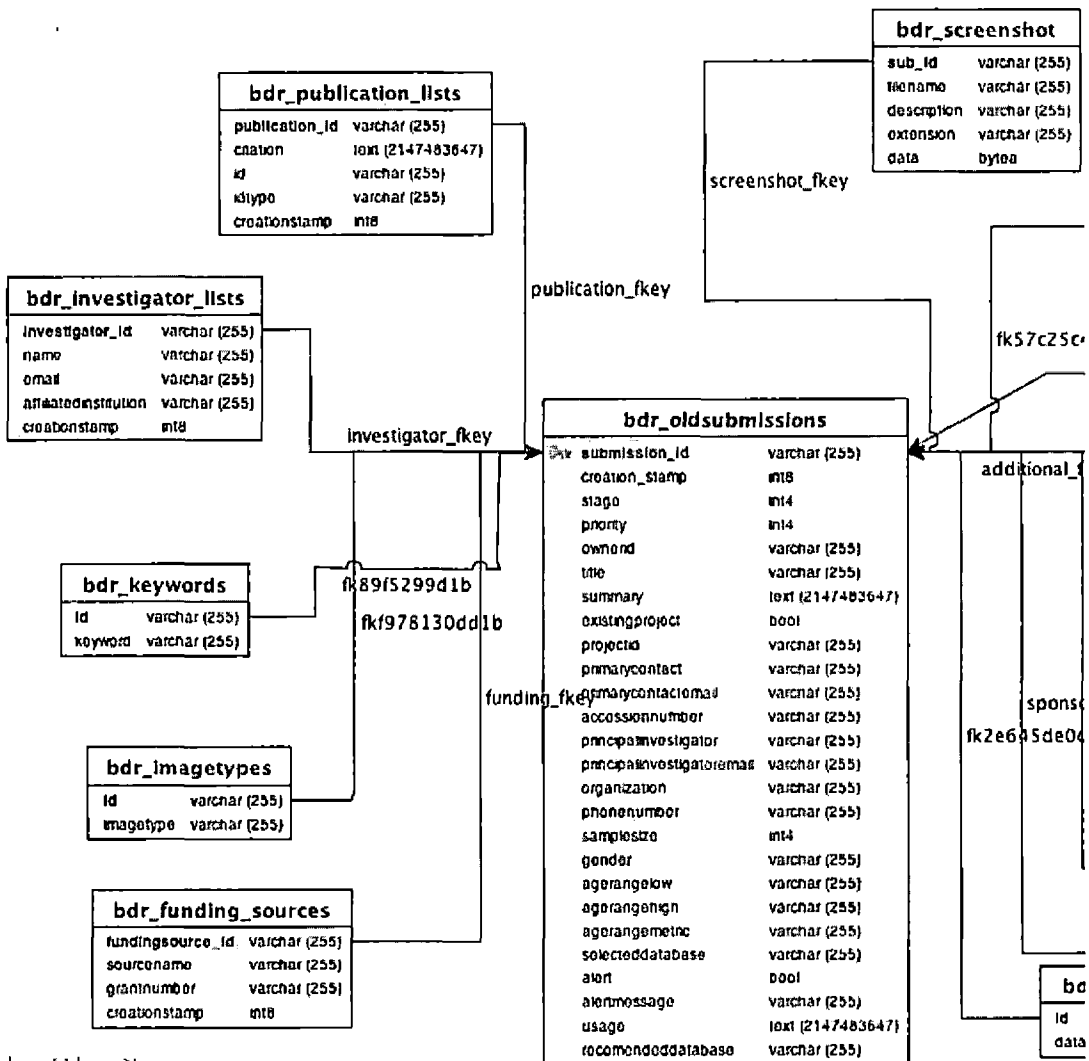


Figure 3.2. Biomedical Informatics Research Network Data Repository Database Schema

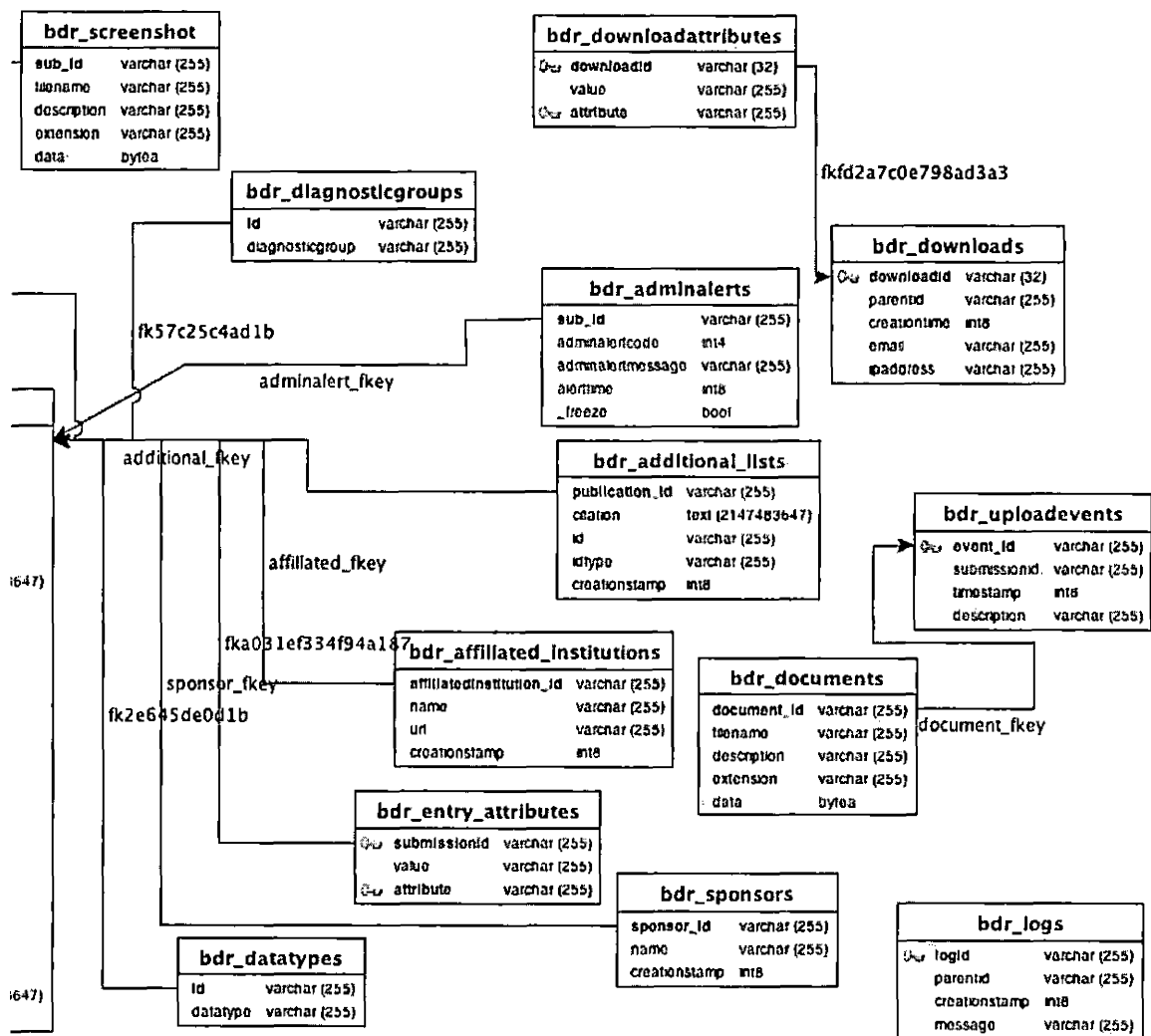


Figure 3.3. Biomedical Informatics Research Network Data Repository Database Schema (continued)

3.1.7 Overview

The BDR portal source consists of three sub packages (portlets, services, and servlets) under package net.birn. All these packages were created to implement the functionality outlined in Chapter 2.

3.1.7.1 Package net.birn.portlets.resources

.datarepository. The data repository portlets package contains nine portlet classes: Administration, BDRFavorites, BDRHelp, BDRLinks, DataRepository, NewSubmission, SubmissionOverview, SubmissionViewer, and Welcome. See Figure 3.4, 3.5, 3.6 and 3.7.

3.1.7.1.1 Administration Class. Administration is a portlet class responsible for handling all the administrative functions that a privileged user may need to perform on a submission that is in the new submission workflow. The initial view it presents to the administrator is an overview of all submissions that are in various stages of being processed.

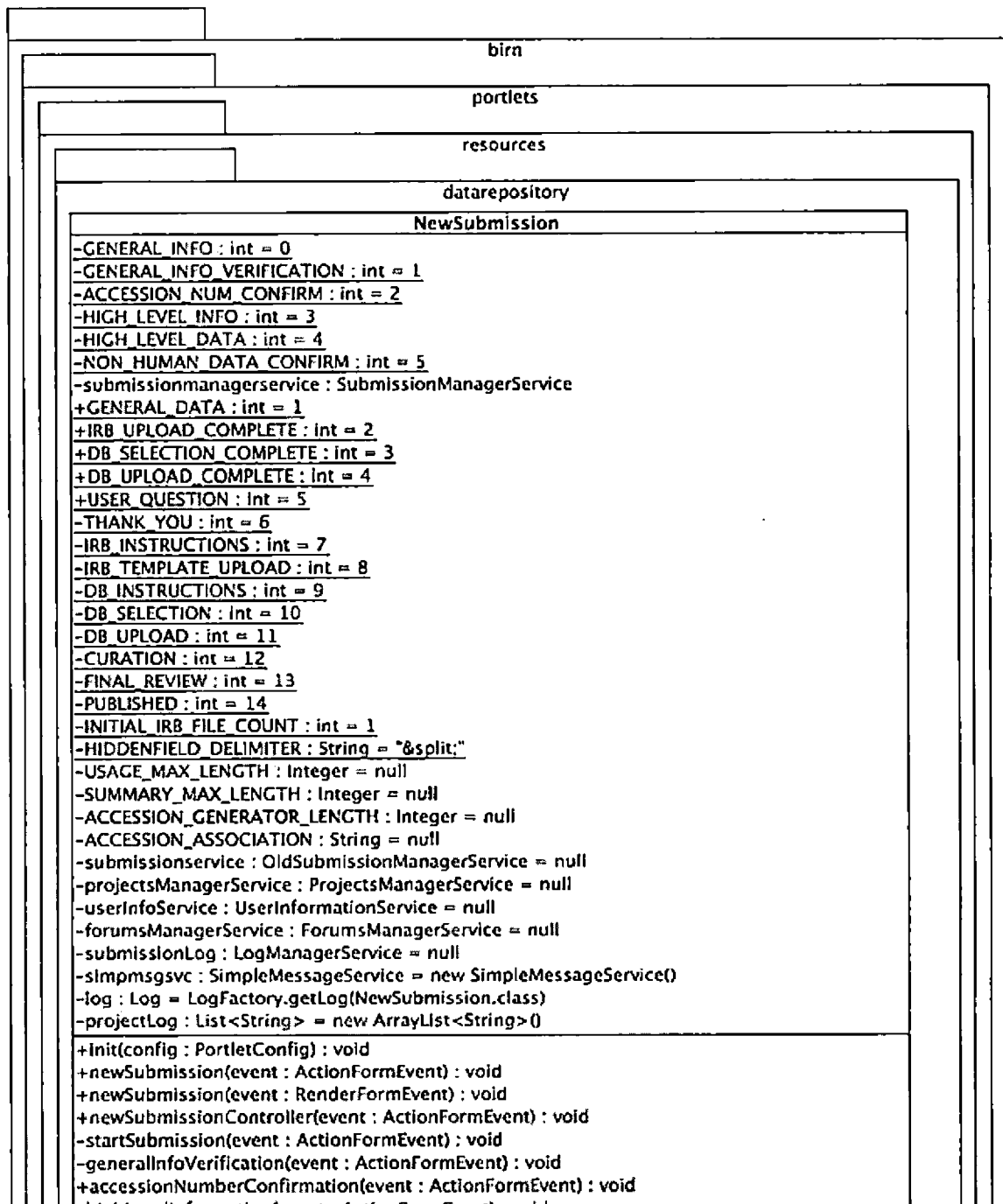


Figure 3.4. Class Diagram of the Package `net.birn.portlets.resources.datarepository` (Part 1)

```

+accessionNumberConfirmation(event : ActionFormEvent) : void
-highLevelInformation(event : ActionFormEvent) : void
+highLevelData(event : ActionFormEvent) : void
+nonHumanDataConfirmation(event : ActionFormEvent) : void
+thankYou(event : ActionFormEvent) : void
+init(config : PortletConfig) : void
+recoverSubmission(event : ActionFormEvent) : void
+recoverSubmission(event : RenderFormEvent) : void
-marshallSubmission(event : ActionFormEvent) : OldSubmission
-validateKey(event : ActionFormEvent) : Boolean
+back(event : ActionFormEvent) : void
+renderController(event : RenderFormEvent) : void
+exitSubmission(event : ActionFormEvent) : void
+uploadImage(event : ActionFormEvent) : void
+gatherGeneralInfo(event : ActionFormEvent) : void
+verifyGeneralInfo(event : ActionFormEvent) : void
+updateSubmission(event : ActionFormEvent) : void
+highLevelInfo(event : ActionFormEvent) : void
+irbInstructions(event : ActionFormEvent) : void
+irbUpload(event : ActionFormEvent) : void
+uploader(event : ActionFormEvent) : void
+dbInstructions(event : ActionFormEvent) : void
+dbSelection(event : ActionFormEvent) : void
+dbUpload(event : RenderFormEvent) : void
+dbUpload(event : ActionFormEvent) : void
+generalInfo(event : RenderFormEvent) : void
+verifyGeneralInfo(event : RenderFormEvent) : void
+accessionNumberConfirmation(event : RenderFormEvent) : void
+highLevelInfo(event : RenderFormEvent) : void
+highLevelData(event : RenderFormEvent) : void
+nonHumanDataConfirm(event : RenderFormEvent) : void
+thankYou(event : RenderFormEvent) : void
+irbInstructions(event : RenderFormEvent) : void
+irbUpload(event : RenderFormEvent) : void
+dbInstructions(event : RenderFormEvent) : void
+dbSelection(event : RenderFormEvent) : void
+curation(event : RenderFormEvent) : void
+finalReview(event : RenderFormEvent) : void
+published(event : RenderFormEvent) : void
-storeSubmission(event : FormEvent, sub : OldSubmission) : OldSubmission
-updateNonHumanData(newSub : OldSubmission, val : String) : void
-createAccessionNumber(length : Integer) : String
-setProjectListBox(event : RenderFormEvent, listBoxBeanID : String, s : OldSubmission) : void
-createProject(request : ActionRequest, newSubmission : OldSubmission) : String
-writeLogToProject(sub : OldSubmission, logMsg : List<String>) : void

```

BDRHelp
+init(config : PortletConfig) : void
+bdrHelp(event : ActionFormEvent) : void
+bdrHelp(event : RenderFormEvent) : void

Welcome
+init(config : PortletConfig) : void
+welcome(event : ActionFormEvent) : void
+welcome(event : RenderFormEvent) : void

Figure 3.5. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 2)



Figure 3.6. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 3)

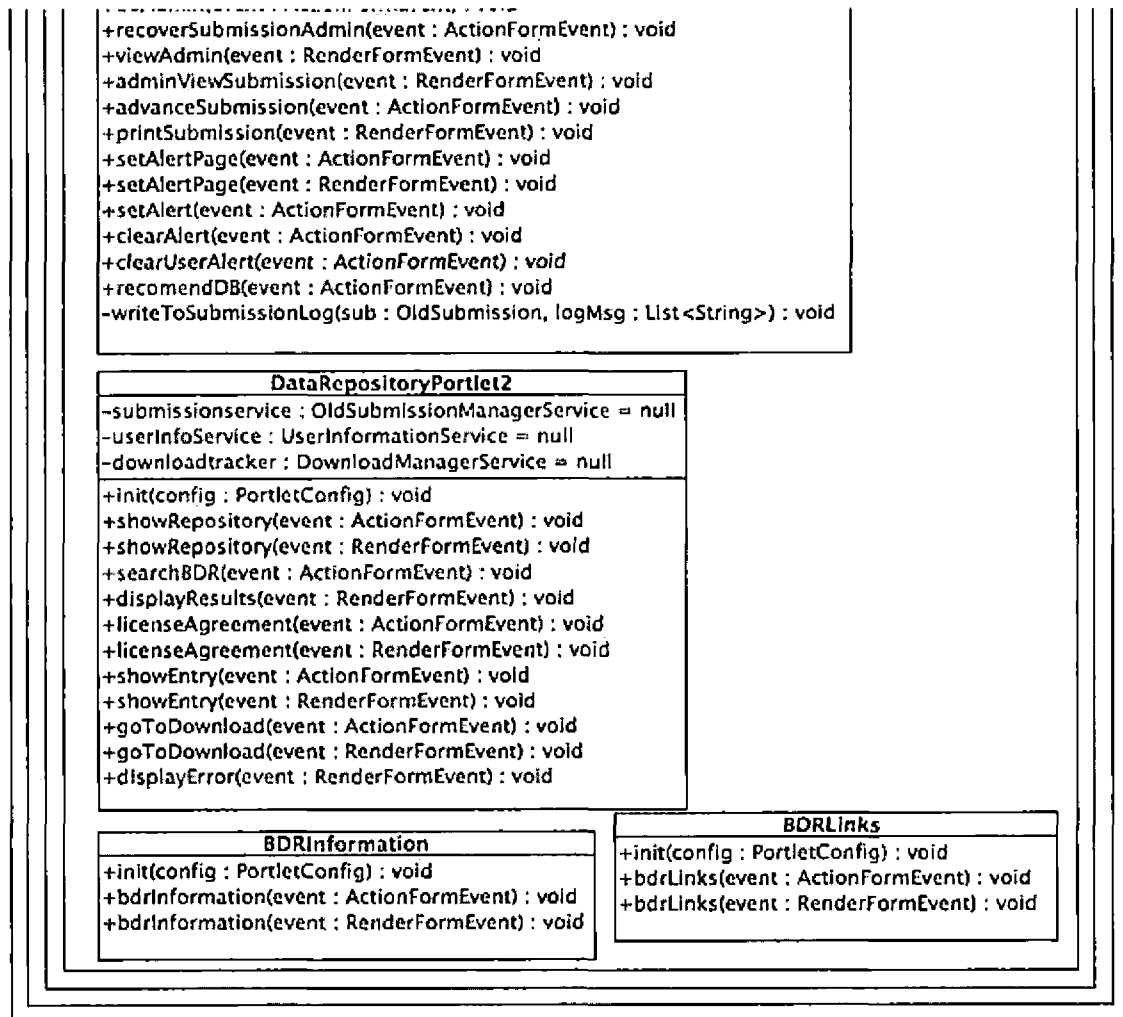


Figure 3.7. Class Diagram of the Package net.birn.portlets.resources.datarepository (Part 4)

3.1.7.1.2 BDRHelp and BDRLinks Class. BDRHelp and

BDRLinks are portlet classes that display a simple JSP file that contains multiple links for help documentation and important PDF files.

3.1.7.1.3 DataRepository Class. DataRepository

provides the search capabilities for the BDR. The browse

feature displays a simple flat file, which is included in the portlet. For the search concept name feature, the data repository creates an axis Web service call to submit the search token to the Mediator Web service. Once the action is performed it is returned to the portal as a simplified XML node structure. These results are then marshaled into the node object and set as a session variable so that the portlet can render the search results in the corresponding ISP page.

3.1.7.1.4 NewSubmission Class. The NewSubmission class is the main portlet for this initial release of the BDR since one of the most important features is the publication of new datasets. With the use of the submission manager service to persist the provided inputs, the new submission portlet guides the user through the steps of the submission process from general information gathering, and collaboration project creation.

3.1.7.1.5 SubmissionOverview and SubmissionViewer Classes. The SubmissionOverview and SubmissionViewer is a portlet class that uses the submission manager service to show a general overview of all submissions that belong to that user, displays its progression, and the current state that it is in. The submission overview portlet allows the

user to view a detail view of the submission, this includes: a view of the submission progression, submission logs, user functions like resuming the submission from the last point the user left the workflow off at.

3.1.7.1.6 Welcome Class. Welcome is a simple portlet class that only displays a flat text file to the user at login.

3.1.7.2 Package net.birn.services.resources
.datarepository. The datarepository services package contains all the main portlet services and base java objects that are used by the system. See Figure 3.8 through Figure 3.15.

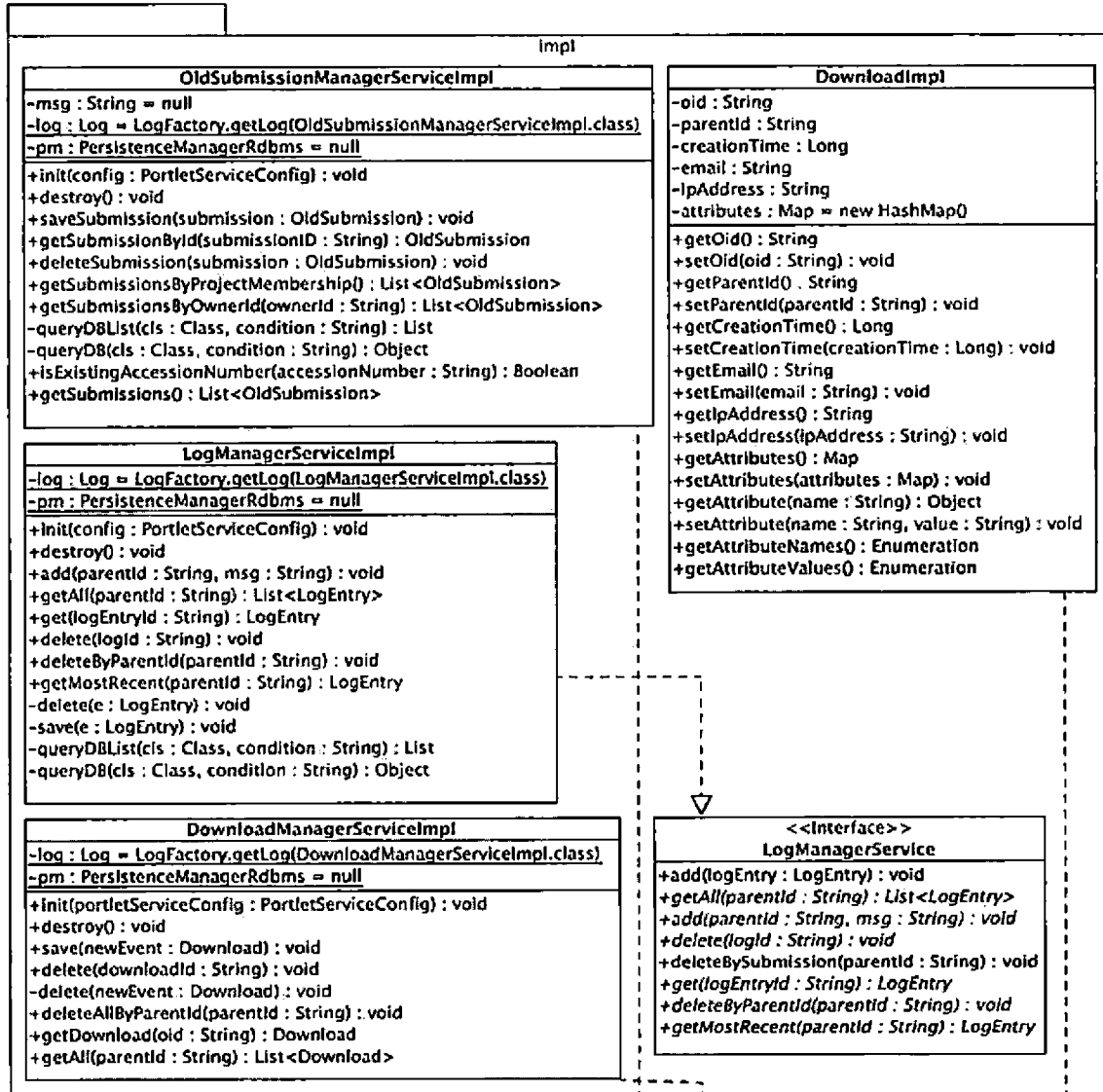


Figure 3.8. Class Diagram of the Package net.birn.services.resources.datarepository.impl

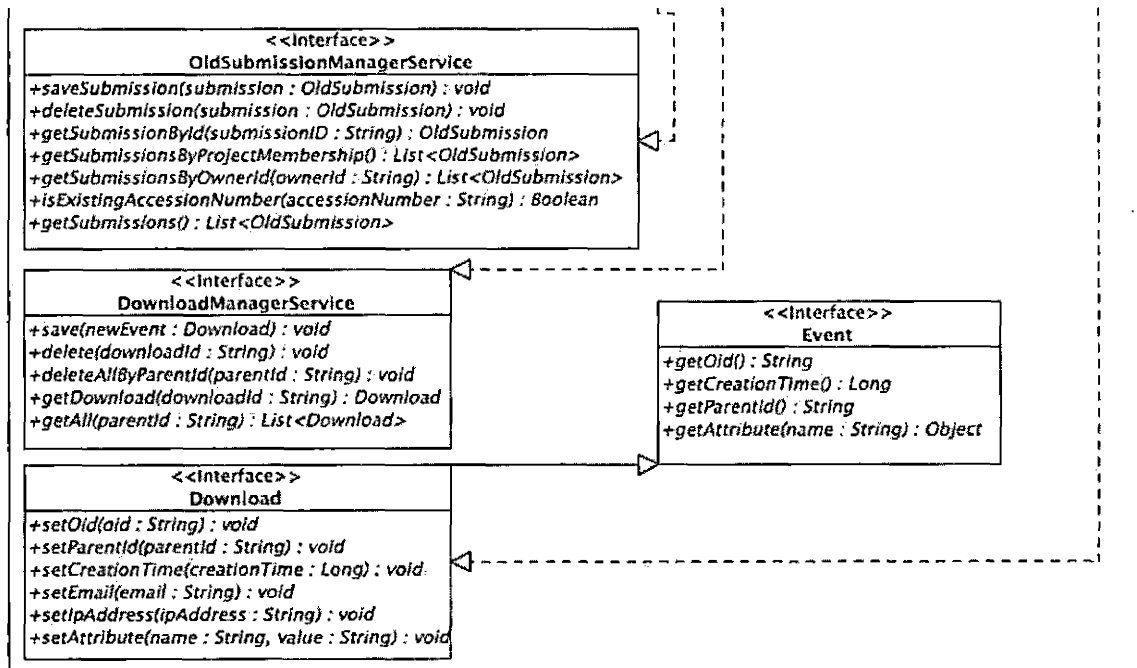


Figure 3.9. Class Diagram of the Package `net.birn.services.resources.datarepository.impl` (Continued)

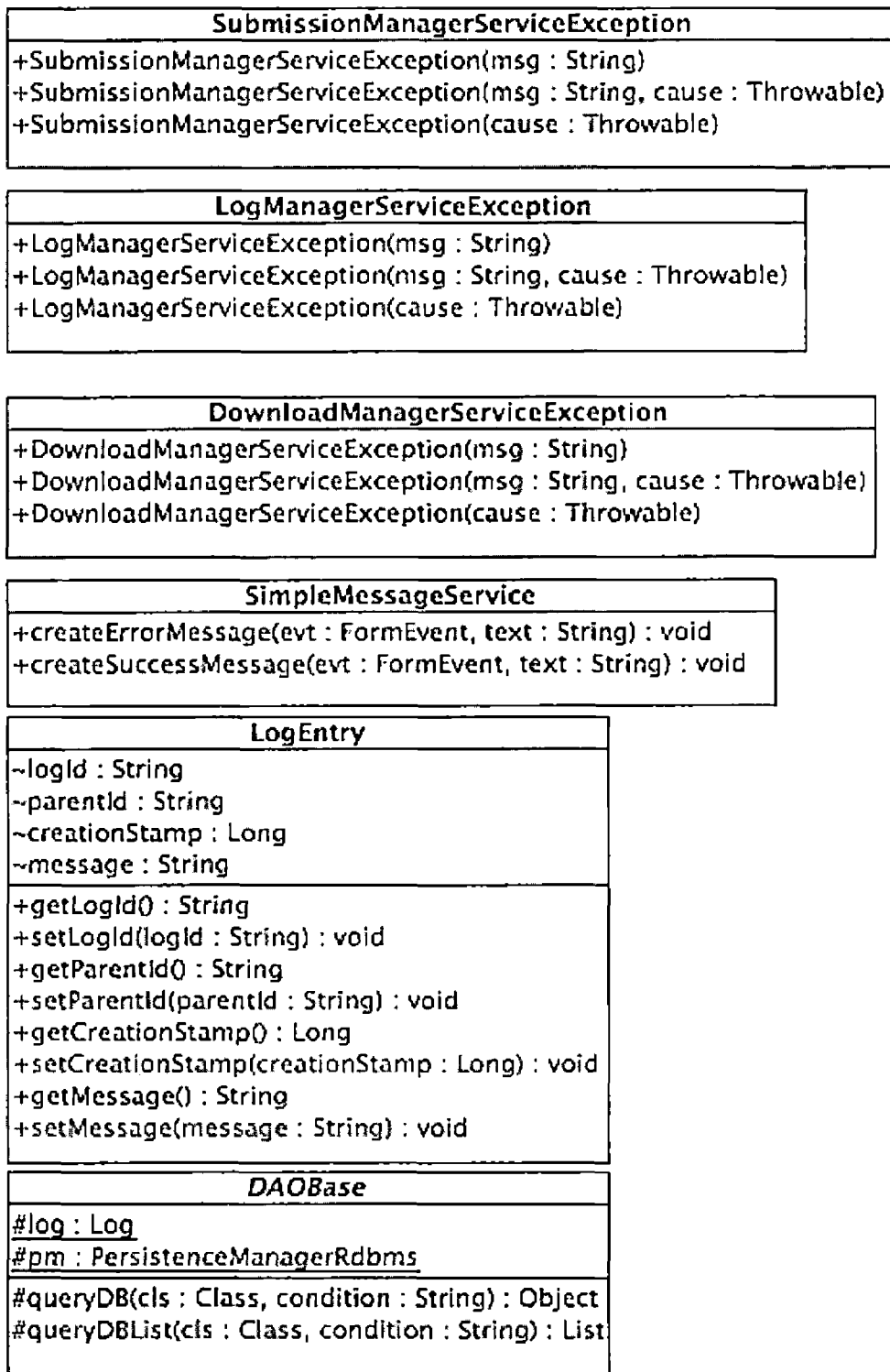


Figure 3.10. Class Diagram of the Package net.birn.services.resources.datarepository.impl (Concluded)

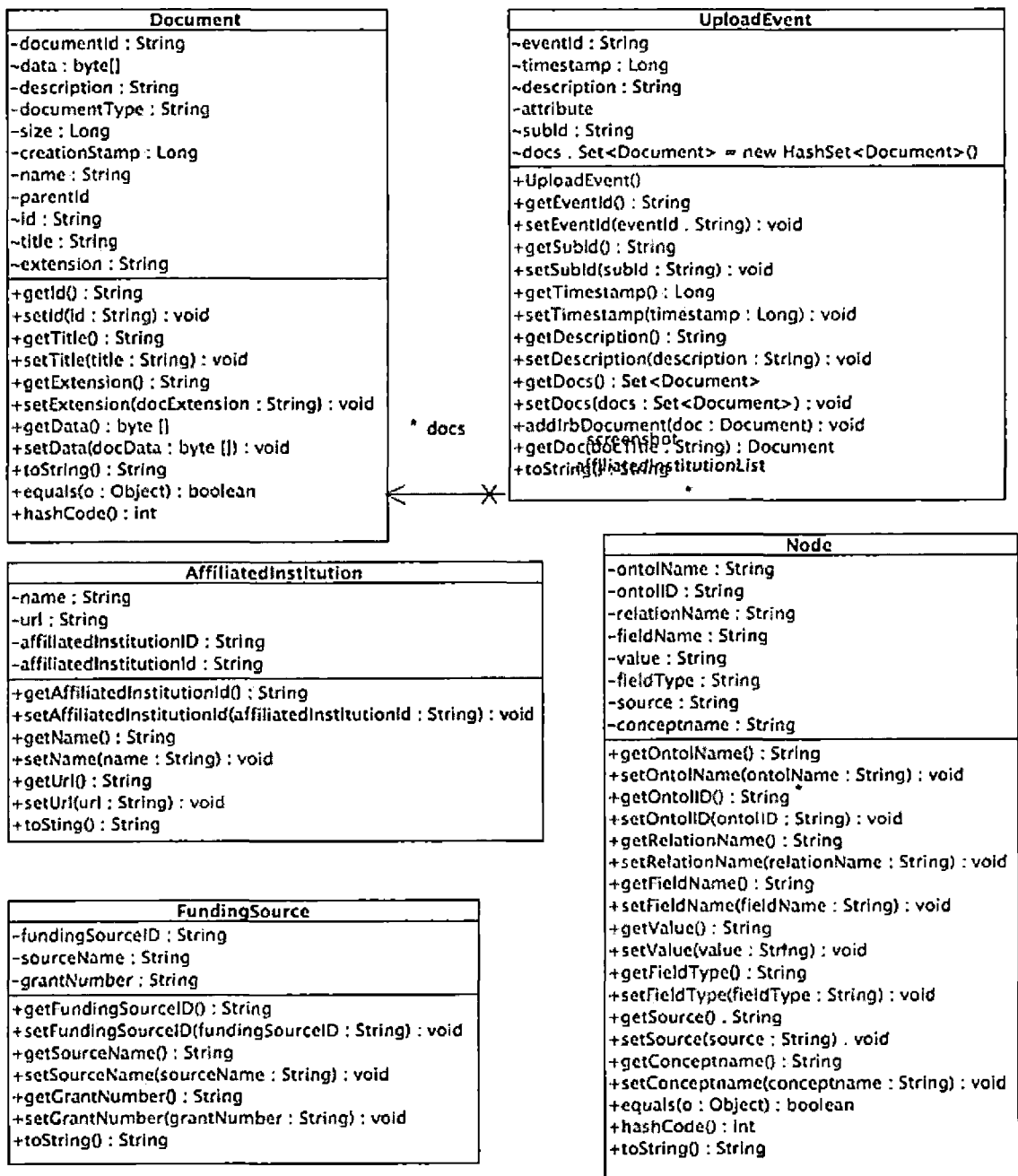


Figure 3.11. Class Diagram Of The Package net.birn.services.resources.datarepository

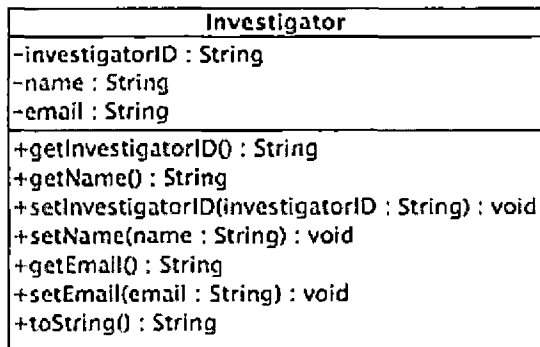
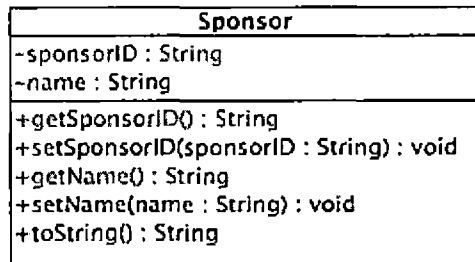
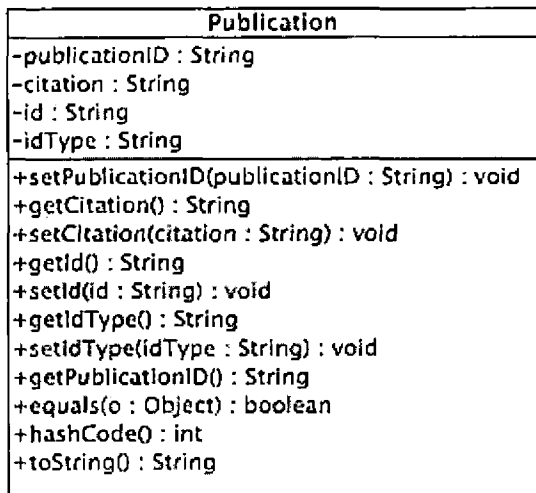
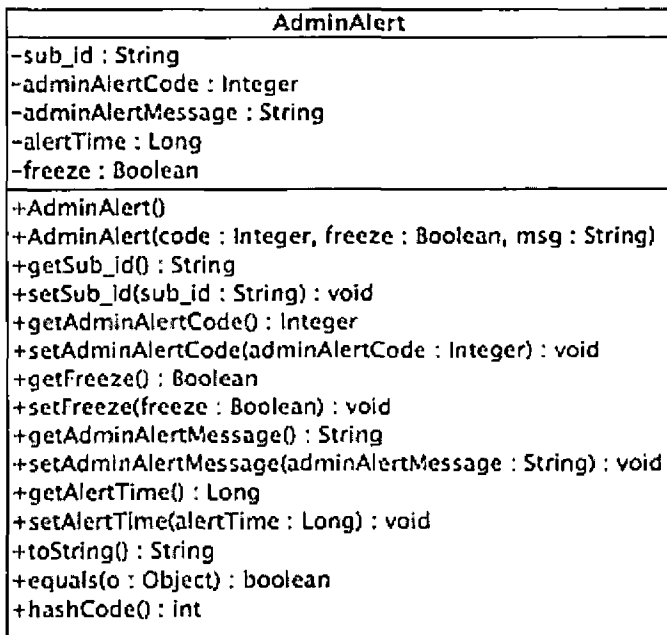


Figure 3.12. Class Diagram Of The Package net.birn.services.resources.datarepository (Continued)



Figure 3.13. Class Diagram of the Class net.birn.services.resources.datarepository.OldSubmission (Part 1)


```

+getStage() : Integer
+setStage(stage : Integer) : void
+increaseStage() : void
+decreaseStage() : void
+getPriority() : Integer
+setPriority(priority : Integer) : void
+getOwnerId() : String
+setOwnerId(ownerId : String) : void
+getTitle() : String
+setTitle(title : String) : void
+getSummary() : String
+setSummary(summary : String) : void
+getExistingProject() : Boolean
+setExistingProject(existingProject : Boolean) : void
+getProjectID() : String
+setProjectID(projectID : String) : void
+getPrimaryContact() : String
+setPrimaryContact(primaryContact : String) : void
+getPrimaryContactEmail() : String
+setPrimaryContactEmail(primaryContactEmail : String) : void
+getAccessionNumber() : String
+setAccessionNumber(accessionNumber : String) : void
+getPrincipalInvestigator() : String
+setPrincipalInvestigator(principalInvestigator : String) : void
+getPrincipalInvestigatorEmail() : String
+setPrincipalInvestigatorEmail(principalInvestigatorEmail : String) : void
+getOrganization() : String
+setOrganization(organization : String) : void
+getPhoneNumber() : String
+setPhoneNumber(phoneNumber : String) : void
+getSampleSize() : Integer
+setSampleSize(sampleSize : Integer) : void
+getGender() : String
+setGender(gender : String) : void
+getAgeRangeLow() : Integer
+setAgeRangeLow(ageRangeLow : Integer) : void
+getAgeRangeHigh() : Integer
+setAgeRangeHigh(ageRangeHigh : Integer) : void
+setAgeRangeMetric(ageRangeMetric : String) : void
+getSelectedDatabase() : String
+setSelectedDatabase(selectedDatabase : String) : void
+getAlert() : Boolean
+setAlert(alert : Boolean) : void
+getAlertMessage() : String
+setAlertMessage(alertMessage : String) : void
+getScreenshot() : Set<Document>
+setScreenshot(screenshot : Set<Document>) : void

```

Figure 3.14. Class Diagram of the Class net.birn.services.resources.datarepository.OldSubmission (Part 2)

```

+setScreenshot(screenshot : Set<Document>) : void
+getRecomendedDatabase() : String
+setRecomendedDatabase(recomendedDatabase : String) : void
+getAdminAlerts() : Set<AdminAlert>
+setAdminAlerts(adminAlerts : Set<AdminAlert>) : void
+addAdminAlert(adminAlert : AdminAlert) : void
+removeAdminAlert(timestamp : String) : void
+isFrozen() : Boolean
+hasAlert(code : Integer) : Boolean
+getRepresentedDiagnosticGroups() : Set<String>
+setRepresentedDiagnosticGroups(representedDiagnosticGroups : Set<String>) : void
+addToRepresentedDiagnosticGroups(dg : String) : void
+removeFromRepresentedDiagnosticGroups(dg : String) : void
+getDataTypes() : Set<String>
+setDataTypes(dataTypes : Set<String>) : void
+addToDataTypes(dt : String) : void
+removeFromDataTypes(dt : String) : void
+getImageTypes() : Set<String>
+setImageTypes(imageTypes : Set<String>) : void
+addToImageTypes(it : String) : void
+removeFromImageTypes(it : String) : void
+getKeywordList() : Set<String>
+setKeywordList(keywordList : Set<String>) : void
+clearKeywordList() : void
+removeFromKeywordList(oldword : String) : void
+addToKeywordList(keyword : String) : void
+getUsage() : String
+setUsage(usage : String) : void
+getAttributes() : Map
+setAttributes(attributes : Map) : void
+getAttribute(name : String) : Object
+setAttribute(name : String, value : String) : void
+getAttributeNames() : Enumeration
+getAttributeValues() : Enumeration
+getInvestigatorList() : Set<Investigator>
+setInvestigatorList(investigatorList : Set<Investigator>) : void
+addInvestigator(Inv : Investigator) : void
+deleteInvestigator(name : String, email : String) : void
+getFundingSourceList() : Set<FundingSource>
+setFundingSourceList(fundingSourceList : Set<FundingSource>) : void
+addFundingSource(fs : FundingSource) : void
+deleteFundingSource(name : String, number : String) : void
+getAffiliatedInstitutionList() : Set<AffiliatedInstitution>
+setAffiliatedInstitutionList(affiliatedInstitutionList : Set<AffiliatedInstitution>) : void
+addAffiliatedInstitution(ai : AffiliatedInstitution) : void

```

Figure 3.15. Class Diagram of the Class net.birn.services.resources.datarepository.OldSubmission (Part 3)

```

+getAffiliatedInstitutionList() : Set<AffiliatedInstitution>
+setAffiliatedInstitutionList(affiliatedInstitutionList : Set<AffiliatedInstitution>) : void
+addAffiliatedInstitution(ai : AffiliatedInstitution) : void
+deleteAffiliatedInstitution(name : String, url : String) : void
+getSponsorList() : Set<Sponsor>
+setSponsorList(sponsorList : Set<Sponsor>) : void
+addSponsor(s : Sponsor) : void
+deleteSponsor(name : String) : void
+getAgeRangeMetric() : String
+getPublicationList() : Set<Publication>
+setPublicationList(publicationList : Set<Publication>) : void
+addPublication(pub : Publication) : void
+deletePublication(citation : String, id : String, type : String) : void
+getAdditionalPublicationList() : Set<Publication>
+setAdditionalPublicationList(additionalPublicationList : Set<Publication>) : void
+addAdditionalPublicationList(pub : Publication) : void
+deleteAdditionalPublicationList(citation : String, id : String, type : String) : void
+getIrbUploads() : Set<UploadEvent>
+setIrbUploads(irbUploads : Set<UploadEvent>) : void
+addUploadEvent(ue : UploadEvent) : void
+getUploadEvent(uploadEventTimeStamp : String) : UploadEvent
+getMostRecentUploadEvent() : UploadEvent
+equals(o : Object) : boolean
+hashCode() : int
+toString() : String

```

Figure 3.16. Class Diagram of the Class net.birn.services.resources.datarepository.OldSubmission (Part 4)

3.1.7.2.1 SubmissionManagerService Interface.

SubmissionManagerService is the interface to the net.birn.services.resources.datarepository.impl.SubmissionManagerServiceImpl implementation of the submission data access object, which utilized the Gridsphere persistence manager to access the BDR database.

3.1.7.2.2 OldSubmission Class. This is the main object class which is used by the submission workflow and

administration to represent a current submission in progress. `AffiliatedInstitution`, `Publication`, `Sponsor`, `Investigator`, `FundingSource` are all sub-component objects, which are mainly kept as a list in the submission object. The `UploadEvent` class is used to categorize each individual IRB upload event other than containing some basic descriptive information. It also contains a list of the `Document` type. Each of these entries represents an individual document and its data.

3.1.7.2.3 Node Class. The `Node` object is a generic class that was created to hold the query results from the axis call to the Mediator Web services. When search capabilities were extracted from the system into another project, this code became inactive and is currently legacy code.

3.1.7.3 Package `net.birn.servlets.bdr`. `BDRFileServlet` is a servlet, which enables the user or administrator to download a file directly from the database just as if it were on the servers file system.

CHAPTER FOUR
QUALITY ASSURANCE AND TESTING

4.1 Introduction

Software quality assurance encompasses the monitoring and improvement of the quality of and adherence to software product standards, processes, and procedures. This particular section focused on using software testing to evaluate the fulfillment of requirements of the system.

4.2 Unit Testing

Unit tests are the first step of the software testing process. It is a procedure used to validate that the individual units of the source code are working properly.

Table 4.1. Report of Biomedical Informatics Research Network Data Repository Unit Test

Category	Test objective	Source Code	Result
Hibernate configuration	XML validation based on corresponding DTD.	Hibernate.properties	Pass
Java bean	Properties	net.birn.services.resources.	Pass

Classes	getter and setters	datarepository package	
Validation Classes	Performs correct validation tasks and throws exception on failures.	net.birn.services.resources. utilities.validation package	Pass
Portlet configuration	XML validation, correct installation and rendering of the portlets when deployed to Gridshpere	net.binr.protlets.resources. datarepository package and portlet.xml	Pass
JSP page	Page content is displayed correctly Links working as expected	All jsp pages included under webapp/jsp directory and all of its sub-directories.	Pass
Servlet	Sends a file to the browser via http	BDRFileServlet.java	Pass

4.3 Integration Test

Most of the functionalities in BDR are provided by integration of some of the individual components. Due to abundance of input from the client on the user interface of the BIRN Data Repository a black-box top-down method was adopted for the integration testing of features. The test report is shown in the table below. The integrated components with the tested functions are listed. All the integration tests were completed successfully.

Table 4.2. Report of Biomedical Informatics Research Network Data Repository Integration Test

Test Components	Test Objective	Result
Submission Manager Service	Handles all data accesses to the database for a particular submission object.	Pass
Collaboration functionality	Create a new collaboration project space, access existing list of BIRN projects; modify attributes and properties of the project.	Pass
Submission workflow rendering	Displays the correct jsp file in the required specific order of the workflow process.	Pass
BDR File Servlet	Sends a buffered byte stream to the browser as a file	Pass

4.4 System Test

This section tests the whole system that is created by integrating all the components to fulfill all the functionalities of the BDR. These functions include: All the functions are tested and passed successfully, as shown in the following table.

Table 4.3. Report of Biomedical Informatics Research Network Data Repository System Test for User Functions

System test for Data Contributor and Data Consumer functions		
Function	Test objective	Result
Creating User	Correct configuration for integration with existing GAMA web services	Pass
Create Submission	Correct configuration and integration with existing SRB, correct workflow, and error handling	Pass
Search by Concept Name	Correct implementation, configuration and integration with mediator web services.	Pass
Upload IRB Documentation	Fulfill requirements of uploading corresponding IRB documentation and information when human derived data is present	Pass
Leave Feedback	Communicate user experiences to the BDR	Pass

	administrators for suggestions on improved functionality	
Obtain Help	View Help topics that have been compiled together	Pass
Accept the "Terms of Use" agreement	When a guest user clicks on a download button for a entry they are presented a simple form to accept the terms of use to continue to the download	Pass
Browse BDR Entries	View a comprehensive listing of the data sets that have been published to the BIRN Data Repository	Pass

Table 4.4. Report for Systems Test for Administrative Functions

System tests for Administrative functions and shared Data contributor functions		
Function	Test Objective	Result
Create Message Correspondence	Create simple notification for the data contributor	Pass
Control Submission Status	Clear system notifications of data contributor actions; remove locking to allow movement of submission	Pass
Control submission	Advance a submission to next stage	Pass

progression	or back to the proceeding stage of submission workflow	
View BDR Submission	View a comprehensive report of all the information provided by the data contributor	
View Detailed Information from a stage	View detailed information from a completed stage	Pass
View Overall Submission Progress	View a graphical representation in the Admin - submission specific overview page of the progress of the submission through the workflow	Pass
View Uploaded IRB Documentation	View a summary of the documents uploaded by the data contributor in the Admin submission specific overview and download these documents to view the content	Pass
View message Correspondence	View messages that have been created by the data contributor	Pass
Publish Data Submission	Make the submission available to the public for viewing and download	Pass

4.5 Test Samples

In this section, samples of test cases are shown individually.

4.5.1 Creation of Biomedical Informatics Research Network Data Repository Submission

1. Click the 'Create Submission' sub tab
2. In the Create Submission portlet click the 'Begin New Submission' button (Figure 4.1).

The screenshot shows the BIRN Data Repository website interface. At the top, there is a navigation bar with links for 'BIRN Portal', 'BDR - Currently Available Data', and 'BIRN Portal'. Below this is a header section with the BIRN logo and the text 'BIRN DATA REPOSITORY' and 'BIOMEDICAL INFORMATICS RESEARCH NETWORK'. A user is logged in as 'Laurence Bonhaman'. The main content area is divided into two columns. The left column contains a 'Detailed Overview of My Submissions' section with 'In Progress' and 'Completed' tabs, and a 'Submission Instructions' section with links for 'Overview', 'Detailed Instructions', 'Associated Metadata', 'Data Contributor's Agreement', 'Data Use Agreement', 'IRB Template for Sharing Human Data', and 'Glossary'. The right column contains a text box with instructions for the submission process, followed by a 'Begin New Submission' button. At the bottom of the page, there are logos for the National Institutes of Health, the National Library of Medicine, and the Biomedical Research and Informatics Network. The date 'May 7, 2008' is displayed in the bottom left corner.

Figure 4.1. Launch Page For Beginning A New Submission

3. Enter submission title, description, and primary contact information. If the data is related to an existing project please select the 'yes' radio button the project from the drop down menu (Figure 4.2). Check the checkbox to indicate that the terms and conditions have been reviewed, and then click the Next button.

Launch New Submission

Instructions: Page 1 of 7

Please provide submission title, study description, keywords that describe your data, and agree to the terms and requirements outlined in the **BDR Data Contributor's Agreement**. You may also designate someone else as a **Primary Contact**.

* Required Fields

- Submission Title: _____

- Study Description / Abstract (Maximum of 5000 characters):

Keywords:

* Select at least 3 keywords that best describe your data:
 Which dictionary would you like to use? **BIRNLex**

Search
 Keywords: _____

Note: Please wait while the ontology loads

Unselected		Selected
(R)N6-phenylisopropyladenosine	<input type="button" value="Add >>"/>	
11-dehydro-thromboxane B2	<input type="button" value="<< Remove"/>	
129		
129P1/Rej		
129P3/J		
129P3/JEmsJ		
129P4/RrRkJ		
129S1/SvImJ		
129S6/SvEvTac		
129T2/SvEms		

Don't see all of the keywords you like? Add more of your own.

Contact Information:

Name of Contributing Investigator: Laurence Bohannon
 E-mail of Contributing Investigator: bohannon@ncmir.ucsd.edu
 Institution/Organization: UCSD BIRN-CC
 Phone Number: 858 822-0718

Name of Primary Contact
 (if different from Contributing Investigator): _____

Primary Contact e-mail: _____

Project Information:

* Is this study related to an existing submission or BIRN project? (If 'Yes' please select the project from the list below)
 No / Yes

Figure 4.2. Launch New Submission

4. The system verifies that the information you have provided is correct (Figure 4.3).

Launch New Submission

Instructions: Page 2 of 7

Please verify the information you have provided for accuracy. To make changes, click **Edit Information**. Once you have verified the information, click **Launch Submission** to proceed.

* Required Fields

New Submission Information Verification:

Submission Title: This is the example submission title

Study Description / Abstract:
This is the example Description it can be very long in length.

Keywords: Insula
C3HeB/FeJ
Anterior ascending limb of lateral sulcus
Alveus
Accessory medullary lamina
Taxon
Pericentral nucleus of inferior colliculus

Contributing Investigator: Laurence Bohannon
Contributing Investigator e-mail: boihan01@yahoo.com
Organization: Bim CC
Phone Number:
Primary Contact: John Doe
Primary Contact e-mail: john@test.com.
Related to an existing BIRN project: No

[Edit Information](#) [Launch Submission](#)

Figure 4.3. High Level Information Verification

5. The new submission portlet creates a collaborative workspace and displays to the user a uniquely created accession number. The new BDR submission has been launched to proceed click 'Continue' (Figure 4.4).

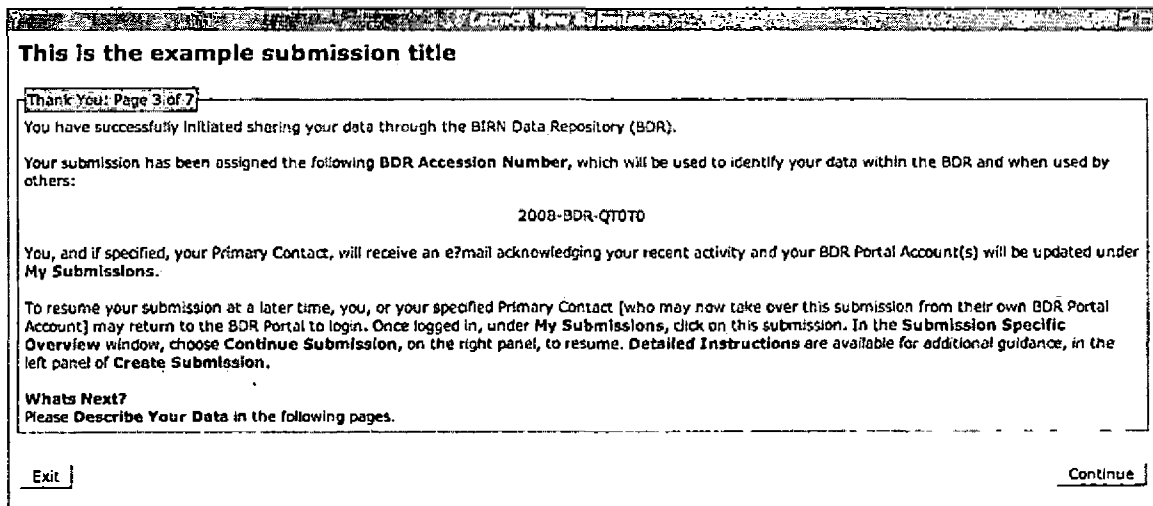


Figure 4.4. Accession Number Confirmation

4.5.2 Upload of Internal Review Board Documents

1. View IRB Instructions and then move to next page.
2. Enter the description.
3. Enter multiple files to be uploaded to the BDR.

Select the file location in the Browse button, the extension from the option menu, and the title for each entry (Figure 4.5).

4. Click the submit button (Figure 4.6).

Submitting Human Subjects Data

This is the example submission title

Instructions:

1. Under **Documentation Upload**, please Provide: a description of the file(s) to be uploaded, title, file type, and file. Click on **Add More Files** to load additional files under the same description; then click on **Submit** to send them to the BDR.
2. Repeat step 1, for additional file(s) under a different description, as needed.
3. When **ALL** the necessary files for IRB approval have been submitted and your documentation is complete, click on **Notify BDR of Completion**.

Documentation Upload:

Current Date: 05/08/2008

Description:

Title	File Type	File
<input type="text"/>	<input type="text"/>	<input type="text"/> Browse...
<input type="text"/>	<input type="text"/>	<input type="text"/> Browse...
<input type="text"/>	<input type="text"/>	<input type="text"/> Browse...
<input type="text"/>	<input type="text"/>	<input type="text"/> Browse...

Add More Files - Remove

Submitted Documentation:

Time: May 8, 2008 12:37:14 AM

Description:

This is the description

Title	File Type	File	View
This is file one	pdf	<input type="text"/>	<input type="text"/>
This is file four	tiff	<input type="text"/>	<input type="text"/>
This is file three	jpeg	<input type="text"/>	<input type="text"/>

Figure 4.6. After Completion of the Internal Review Board Document Upload, this is Displayed

4.5.3 Error Handling

There are three main types of errors that can occur in the BIRN Data Repository; uncompleted required field, invalid input, and improperly formatted inputs. All user-supplied fields may undergo any sort of combination of

error handling for the three previous cases; all error handling is performed on the server. For simplicity the three examples, one for each case of error handling, are shown below.

4.5.3.1 Required Field Error Handling - Launch New Submission. When a field that is required does not contain

SUBMISSION TITLE IS A REQUIRED FIELD.

Instructions: Page 1 of 7

Please provide submission title, study description, keywords that describe your data, and read the **Data Contributor's Agreement**. You may also designate someone else as a **Primary Contact**.

* Required Fields

* Submission Title: _____

* Study Description / Abstract (Maximum of 5000 characters): _____

Keywords:

* Select at least 3 keywords that best describe your data:
Which dictionary would you like to use? **BIRNLex**

Search Keywords:

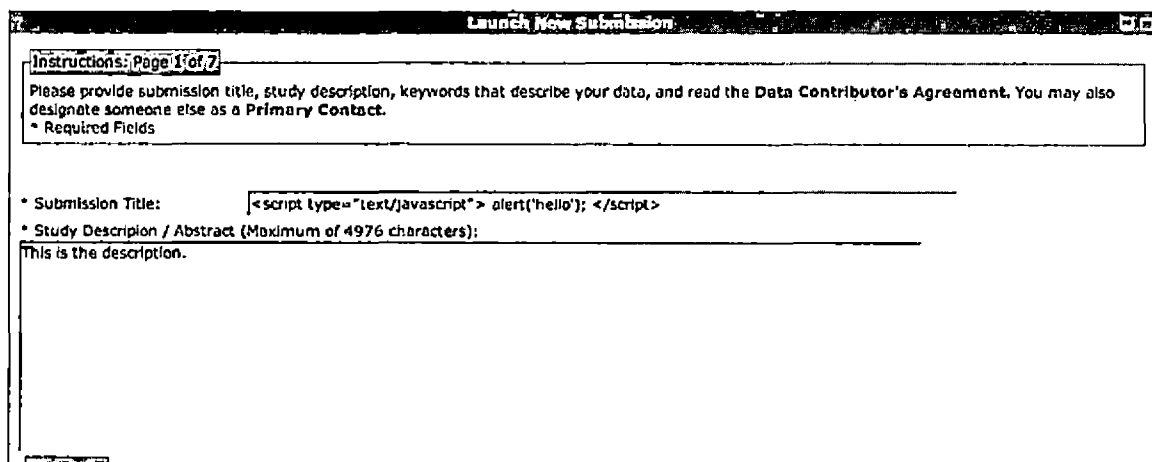
Unselected		Selected
Abducens nerve fibers		
Abducens nucleus		
Accessory medullary lamina		
Accessory nerve fiber bundle		
Alveus		
Amniota		
Amygdala		
Angular gyrus		
Annexant gyrus		
Anterior ascending limb of lateral sulcus		

Don't see all of the keywords you like? Add more of your own. _____ **Add**

Figure 4.7. Error Handling on a Required Field, Submission Title

any value, an error message is set for the next page rendering See Figure 4.7.

4.5.3.2 Invalid Input Error Handling - Launch New Submission. When an illegal character is entered as the value of a text field, an error is not reported to the user, instead a cleaning method strips out the illegal characters and enters it in the equivalent HTML code for that corresponding character. This improves the security of the system without compromising the usability; in the following example a simple JavaScript insertion demonstrates this feature (Figure 4.8). The string is displayed correctly on the following information



The screenshot shows a browser window titled "Launch New Submission". At the top, there is a tab labeled "Instructions: Page 1 of 7". Below this, a text box contains the following instructions: "Please provide submission title, study description, keywords that describe your data, and read the Data Contributor's Agreement. You may also designate someone else as a Primary Contact." Below the instructions, there is a section for "Required Fields". The first field is "Submission Title:", which has a text input field containing the JavaScript code: `<script type="text/javascript"> alert('hello'); </script>`. The second field is "Study Description / Abstract (Maximum of 4976 characters):", which has a text input field containing the text: "This is the description."

Figure 4.8. Invalid Input - Launch New Submission

verification page (Figure 4.9), however the string is encoded to display correctly the HTML characters (Figure 4.10).

Launch New Submission

Instructions: Page 2 of 7

Please verify the information you have provided for accuracy. To make changes, click [Edit Information](#). Once you have verified the information, click [Launch Submission](#) to proceed.

* Required Fields

New Submission Information Verification:

Submission Title: `<script type="text/javascript"> alert('hello'); </script>`

Study Description / Abstract:
This is the description.

Keywords: Alveus
Amniota
Amygdala

Contributing Investigator: Laurence Bohannon
Contributing Investigator e-mail: bohannon@ncmir.ucsd.edu
Organization: UCSD BIRN-CC
Phone Number: 858 822-0718
Primary Contact:
Primary Contact e-mail:
Related to an existing BIRN project: No

[Edit Information](#) [Launch Submission](#)

Figure 4.9. Invalid Input - New Submission Information Verification

```

<fieldset><legend>New Submission Information Verification:</legend>
  <table class="portlet-frame">
    <tr class="portlet-section-alternate">
      <td style="width:20%" align="right">Submission Title</td>
      <td><span dir="ltr" class="portlet-msg-info">&lt;script type="text/javascript"&gt;
        alert('hello'); &lt;/script&gt;</span></td>
    </tr>
  </table>

```

Figure 4.10. Source Code for the Submission Title Section of the Table Figure 4.9

4.5.3.3 Improperly Formatted Error Handling - Launch Submission. When a string value is entered as a value but does not conform to a specific pattern, an error message will be displayed. For instance, the in email field, the string value must strictly adhere to a set pattern (Figure 4.11 and Figure 4.12).

Contact Information	
Name of Contributing Investigator:	Laurence Bohannon
E-mail of Contributing Investigator:	bohannon@ncmir.ucsd.edu
Organization:	UCSD BIRN-CC
Phone Number:	858 822-0718
Name of Primary Contact (if different than Contributing Investigator):	<input type="text"/>
Primary Contact e-mail:	<input type="text" value="test@test.com"/>

Figure 4.11. Improperly Formatted - Launch Submission with Malformed Email Value

Launch New Submission

PLEASE PROVIDE A PROPER EMAIL ADDRESS.

Instructions: Page 1 of 7

Please provide submission title, study description, keywords that describe your data, and read the **Data Contributor's Agreement**. You may also designate someone else as a **Primary Contact**.

- * Required Fields

Figure 4.12. Improperly Formatted - Launch Submission with Error Message

CHAPTER FIVE
MAINTENANCE AND USER MANUAL

5.1 Maintenance Manual

This section contains all the structures and directories of files including source code, object codes as well as documentations. It also contains instructions on how to build and reinstall the software.

5.1.1 Directory Organization

Under the BDR directory, the file system structure adheres to the typical Gridsphere project model. There are two build files (build.properties and build.xml) and three subdirectories (lib, src, and webapp):

1. /src: This directory contains all the Java source code that generates the portlets, servlets and other Java classes that are unique to BDR.
2. /lib: This is the general package library that contains all nonstandard required jars that are not contained in any other packages. There is usually only four files in this directory all of which are required for axis Web service calls.
3. /webapps: This directory contains several sub directories, most of these directories (html,

javascript, and pdf) contain just flat files, the /jsp sub directory contains all the jsp files of the package. The WEB-INF sub directory contains all the necessary files to configure the portlet to run on the Gridsphere framework and the Gridsphere persistence manager service, which utilizes hibernate.

5.1.2 Installing and Re-Compiling of the Biomedical Informatics Research Network Data Repository Portal on a Tomcat Server

The BDR portal does not adhere to the typical preferred method of installation for portlets, deployable WAR files, due to the extensive dependencies of the portal on other portlets and middleware. These installation instructions assume the user has already installed a functional GAMA Web service, SRB server, and Postgres database. The installation of these systems goes beyond the scope of this document. In order to install the BDR portal it must be deployed along with the Gridsphere 2.1.X portal framework with several other portlet packages and run within a servlet container. This section will describe using Tomcat 5.5 as the execution environment. Deploying to other servlet containers will be specific to each container. Please refer to the user manual for your particular servlet container of choice.

1. Download the most recent version of Tomcat 5.5, for this installation we are using 5.5.26. Also download the corresponding JDK 1.4 compatibility pack for installation.
2. Unpack both of these packages to your Tomcat home directory, and set environmental variable for CATALINA_HOME to the correct path.
3. Checkout the most recent release of GS 2.1.X from the Gridsphere SVN server to a separate directory, since this is the source it doesn't have to be a Tomcat subdirectory. This will be referred to as GS_HOME in the rest of the instructions
(https://svn.gridsphere.org/gridsphere/branches/GRIDSPHERE_2_1_X/)
4. Create a subdirectory under GS_HOME named "projects".
5. Checkout the most recent release of Gridportlets 1.2.X from the Gridsphere SVN server to a newly created sub directory of the GS_HOME/projects directory.
(https://svn.gridsphere.org/gridportlets/branches/GRIDPORTLETS_1_2/)
6. Repeat the previous step of downloading the package and placing it in a newly created sub directory under GS_HOME/projects. Perform this step for each portlet

package: gama, birnsiteinfo, srbportlets, bluesquid, bdr, and bdrportal.

7. If using anything other than the default hibernate HSQL rational database system (ie. Postgres) go ahead and create a user and an empty database. Make sure the user has appropriate rights to access the database. Edit the corresponding hibernate.properties files in Gridsphere and each project to use the appropriate database JDBC jar file combination.
8. Once all packages have been placed under the GS_HOME/projects directory in their corresponding sub-directories we can begin to build the project. This is done by going to the GS_HOME directory and issuing the command "ant install."
9. Repeat this step for each portlet package in the projects directory in the following order:
Gridportlets, BIRNsite Info, GAMA, SRB portlets, bluesquid, and BDR. The last package to be installed is that of the BDR portal. To do this, type the command "ant deploy-bdrportal."
10. Startup your servlet container and point your browser to `http://<IPaddress>:<port>/gridsphere`
`/gridsphere`, there you will see the initial account

creation page for the administrator to create an admin account if the database had no entries in it.

5.2 User Manual

This section contains the user manual to guide the user to use the service properly.

5.2.1 Browsing Biomedical Informatics Research Network Data Repository Entries

The browse function of the BDR is exposed to the public; this is available through the 'Browse BDR' sub tab.

BIRN DATA REPOSITORY
 BIRN CALIFORNIA BIOMEDICAL INFORMATICS RESEARCH NETWORK

Logout
 Welcome, Laurence Bohannon

BDR | Collaboration | Layout | Administration | Accounts | Help

My Account | My Submissions Overview | Create Submission | My Profile | Help | Admin

Detailed Overview of My Submissions

In Progress

Function	Title	Status Alerts
BIRN Neuroimaging Calibration Study Phase II		<input type="radio"/>
ifad		<input type="radio"/>
This is a test upgrade submission		<input type="radio"/>
This is the title		<input type="radio"/>
BDR Demo Submission		<input type="radio"/>

Completed

Title	Downloads
You have no completed Submissions.	

Currently Available Data

The BDR is a sustainable archive of data generated by the biomedical research community. Data housed in the BDR are available for sharing and exchange. Biomedical imaging data from humans and animal models are the primary data type available with a minimal set of descriptive metadata, and in some cases, associated clinical, genetic, or other biomedical data. Data are available through BDR resources including a variety of searchable databases along with archived downloads. Future BDR releases will allow data discovery via a concept-based query interface.

All data are freely available to the public with the provision of an active e-mail address and acceptance of the standard BDR Data Use Agreement; this agreement will be required prior to data download. To bypass this requirement in the future when using this device, please enable cookies, which will save your original acceptance of the Agreement; or create your own personalized BDR Portal Account, where these provisions are agreed to at account registration.

Please consider sharing your own data, or data resulting from your use of downloaded data, through the BDR to extend and strengthen this collaborative environment.

BIRN DATA REPOSITORY SOURCES:

CURRENTLY AVAILABLE DATA FROM BIRN-ENABLED COLLABORATIONS:

Figure 5.1. Default View of the Browse Biomedical Informatics Research Network Data Repository Page

The user can click on the 'plus' icon next to the bold text 'Currently Available Data From BIRN-enabled Collaborations' (Figure 5.1).

This action will expand the list of all the title and descriptions of the datasets that are currently available to the public. By clicking on the title of a dataset of interest, this directs the user to a data description page for that particular BRIN Data Repository entry (Figure 5.2).

BIRN DATA REPOSITORY
BIOMEDICAL INFORMATICS RESEARCH NETWORK

Logout
Welcome, Laurence
Bohannon

BDR | Collaboration | Layout | Administration | Accounts | Help

My Account | My Submissions Overview | Create Submission | My Profile | Help | Admin

Currently Available Data

The BDR is a sustainable archive of data generated by the biomedical research community. Data housed in the BDR are available for sharing and exchange. Biomedical imaging data from humans and animal models are the primary data type available with a minimal set of descriptive metadata, and in some cases, associated clinical, genetic, or other biomedical data. Data are available through BDR resources including a variety of searchable databases along with archived downloads. Future BDR releases will allow data discovery via a concept-based query interface.

All data are freely available to the public with the provision of an active e-mail address and acceptance of the standard BDR Data Use Agreement; this agreement will be required prior to data download. To bypass this requirement in the future when using this device, please enable cookies, which will save your original acceptance of the Agreement; or create your own personalized BDR Portal Account, where these provisions are agreed to at account registration.

Please consider sharing your own data, or data resulting from your use of downloaded data, through the BDR to extend and strengthen this collaborative environment.

BIRN DATA REPOSITORY SOURCES:

CURRENTLY AVAILABLE DATA FROM BIRN-ENABLED COLLABORATIONS:

Function BIRN Neuroimaging Calibration Study Phase II Technological advances in functional and structural neuroimaging have revolutionized the biomedical investigation of neuropsychiatric illness. The tremendous potential that this methodology brings in advancing diagnostic and prognostic capabilities and in treatment of neuropsychiatric illness has a... More

MRM Reeler Mouse Male and female adult WT, heterozygous, and homozygous mutant ReIn mice (Jackson Laboratory, Bar Harbor, Maine, USA) were used for this study. The animals were approximately 5 months of age at the time of investigation. In preparation for morphology studies, mice were anesthetized with 100 mg/kg... More

Human Functional MRI data: Traveling subjects study. The Function BIRN Phase I Traveling Subjects Dataset includes five subjects who traveled the country and were imaged twice at each of ten FBIRN site on successive

Detailed Overview of My Submissions

In Progress

Title	Status Alerts
Function BIRN Neuroimaging Calibration Study Phase II	<input type="radio"/>
ffod	<input type="radio"/>
This is a test upgrade submission	<input type="radio"/>
This is the title	<input type="radio"/>
BDR Demo Submission	<input type="radio"/>

Completed

[Title](#) [Downloads](#) [History](#)

You have no completed Submissions.

Figure 5.2. Expanded View of Browse Biomedical Informatics Research Network Data Repository Page

The user can then view the various information for that entry: description, a figure (if provided), accession number, technical contact, recommended usages, investigators and contributors, affiliated institutions, publications, and associated funding sources/grant numbers.

Figure:



Caption: BOLD activation related to working memory task performance

Accession Number: 2008-BOR-GQF87

Technical Contact: Dingying Wei (dywei *AT* ucl.edu)

Recommended Usage: Image data is provided in Nifti-1 format which can be read by most image viewing and analysis software packages. Related demographic and clinical data can be queried in a detailed manner through the fBIRN HID and/or downloaded as a comma separate variable file for import into any standard program that reads text files. Behavioral data is available as downloaded with related fMRI image files and are provided in a text format. Related documentation for all data will be available for download.

Investigators and Contributors: Christine Fennema (Contributing Investigator) [1], Ron Kikinis, M.D.[2], Gregory McCarthy, Ph.D.[null], Aysenil Belger, Ph.D.[null], Bruce Rosen, M.D., Ph.D.[4], Gary H. Glover, Ph.D.[5], Arthur W. Toga, Ph.D.[6], Gregory G. Brown, Ph.D.[7], Daniel S. OLeary, Ph.D.[8], Judy Ford, Ph.D.[11], Kelvin O. Lim, M.D.[9], Steven Potkin, M.D.[12]

Affiliated Institutions:

1. UCSD
2. Brigham and Womens Hospital
3. Duke University/University of North Carolina
4. Massachusetts General Hospital
5. Stanford University
6. University of California, Los Angeles
7. University of California, San Diego
8. University of Iowa
9. University of Minnesota
10. University of New Mexico
11. Yale University
12. University of California, Irvine

Primary Publications/Original Articles:

- Friedman L, Stern H, Brown GG, Mathalon DH, Turner J, Glover GH, Gollub RL, Laurienti J, Lim KO, Cannon T, Greve DN, Bockholt HJ, Belger A, Mueller B, Doty MJ, He J, Wells W, Smyth P, Pieper S, Kim S, Kubicki M, Vangel M, Potkin SG. Test-retest and between-site reliability in a multicenter fMRI study. Hum Brain Mapp. 2007 Jul 17.

Funding Sources / Grant Numbers: NIH / NCRR: U24-RR021992

[Download](#)

Figure 5.3. Download Button Located at the Bottom of the Detailed View of a Published Submission

5.2.2 Downloading Data Sets

Once a dataset of interest has been found the user can download the corresponding datasets that are represented by the BDR entry. At the bottom and top of each individual entry page the user will find a download button. Click the download button (Figure 5.3).

Logged in users:

- This download button will lead directly to the corresponding link that will provide the data (Figure 5.4).

The screenshot displays the BIRN Data Repository interface. At the top, the logo for the Biomedical Informatics Research Network is visible, along with a navigation menu including 'BDR', 'Collaboration', 'Layout', 'Administration', 'Accounts', and 'Help'. A user is logged in as 'Laurence Bohannon'. The main content area is divided into two sections. On the left, under 'My Submissions', there is a 'Detailed Overview of My Submissions' section with an 'In Progress' tab. It contains a table with columns for 'Title' and 'Status Alerts'. The table lists four submissions: 'Function BIRN Neuroimaging Calibration Study Phase II', 'ifad', 'This is a test upgrade submission', and 'BDR Demo Submission', each with a status indicator. Below this is a 'Completed' section with a 'Downloads' tab, stating 'You have no completed Submissions.' On the right, a 'Resource Download Information' box provides a URL for data download: 'http://bimbdcr.nih.gov:8080/BDR/'.

Privacy | Terms of Use | Report a Problem | Feedback



Figure 5.4. View of a Typical Resource Link That Would be Provided by the Biomedical Informatics Research Network Data Repository

New Guest Users:

- This download button will lead to a license agreement page, the user may then review the BIRN Data Use Agreement. The user must click the 'I Agree' checkbox and supply an email to precede any further. The user may supply their name, institution, and purpose for download, but these are only optional fields (Figure 5.5).

Provisions for BDR Data Usage

All data are freely available to the public with the provision of an active e-mail address and acceptance of the standard **BDR Data Use Agreement**. Please consider sharing your own data, or data resulting from your use of downloaded data, through the BDR to extend and strengthen this collaborative environment.

Optional If you would like to provide information that would help the BDR track the usage of these data, please provide the optional information below. The required information will only be used in accordance with the **Privacy Policy**.

First Name:

Last Name:

Institution:

Briefly describe your interest in using these data:

Required: *

E-Mail:*

BDR Data Use Agreement: (PDF Format)

**BIRN Data Repository
Data Use Agreement**

This document outlines the provisions of a non-exclusive license for use of data shared through the Biomedical Informatics Research Network (BIRN) Data Repository (BDR). In line with one of the unique aspects of BIRN, the goal of the BDR is to make data originally acquired with investigator-specific interest available to the community for further study. All data are freely available to the public with the provision of an active e-mail address and acceptance of this standard BDR Data Use Agreement. All Users are encouraged to consider sharing resultant data through the BDR to extend and strengthen this collaborative environment.

* I have read, understand, and agree to the terms and requirements outlined in the **BDR Data Use Agreement**.

Note: To bypass this requirement in the future when using this device, please enable cookies, which will save your original acceptance of the Agreement; or create your own personalized BDR Portal Account, where these provisions are agreed to at account registration.



Figure 5.5. Download Information Form Page

- Once the required fields have been completed, click the download button to be directed to the link that will provide the data.

Returning Guest Users:

- This download button will lead directly to the corresponding link that will provide the data. However, if the user has recently removed the cookies present on their computer they must complete the license agreement section of the download steps again.

5.2.3 Creating a User

In the case of when an external individual who wants to share with the BIRN Data Repository some data that they have obtained from an unaffiliated study and does not have an existing account with BIRN or on the BIRN Data Repository, they must first create a BIRN Data Repository user account prior to submitting any data. The first step is to go to the "Request an Account" tab off of the portal's main login page (Figure 5.6).

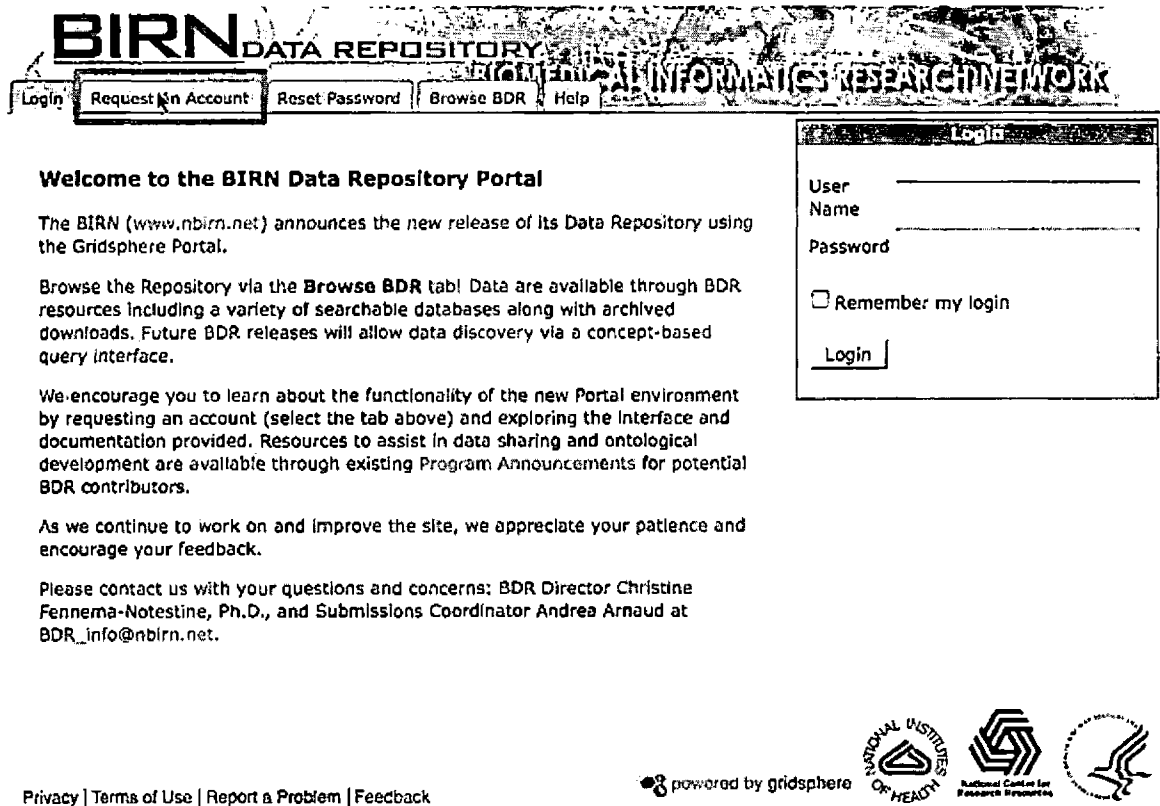


Figure 5.6. Default View of the Portal to a User Who is Not Logged In

This will lead to a form that allows the unregistered user to enter the required information that is needed by the system to create a new account (Figure 5.7).

BIRN DATA REPOSITORY **BIOMEDICAL INFORMATICS RESEARCH NETWORK**

Request an Account

Please enter the following information. Items marked with * are required.

First Name *	Laurence
Last Name *	Bohannon
Email Address *	testbdr@gmail.com
Institution/Organization *	UCSD
Desired username *	testbdr
Street Line 1 *	9500 Gilman Dr
Street Line 2	
City *	La Jolla
State *	Ca
Postal Code *	92039
Country *	USA
Work telephone *	555-5555

Briefly describe your interest in using the portal *

Briefly describe your interest in using the portal * (Maximum of 250 characters):

testing

**BIRN Data Repository
Data Use Agreement**

This document outlines the provisions of a non-exclusive license for use of data shared through the Biomedical Informatics Research Network (BIRN) Data Repository (BDR). In line with one of the unique aspects of BIRN, the goal of the BDR is to make data originally acquired with investigator-specific interest available to the community for further study. All data are freely available to the public with the provision of an active e-mail address and acceptance of this standard BDR Data Use Agreement. All Users are encouraged to consider sharing resultant data through the BDR to extend and strengthen this collaborative environment.

By your acceptance of these terms, you agree to the following provisions for yourself and any

I have read, understand, and agree to the terms and requirements outlined in the BDR Data Use Agreement.

Figure 5.7. New User Registration Form

Once the required fields, which are indicated on the form by the red stars, have been properly filled out and the user has clicked the 'Next' button, they will be directed to a verification page that redisplay the information that has been entered by the user (Figure 5.8).

BIRN DATA REPOSITORY **BIOINFORMATICS RESEARCH NETWORK**

Login | Request An Account | Reset Password | Browse BDR | Help

Request an Account

PLEASE CONFIRM THE FOLLOWING INFORMATION:

First Name	Laurence
Last Name	Bohannon
Email Address	testbdr@gmail.com
Institution/Organization	UCSD
Desired username	testbdr
Street Line 1	9500 Gilman Dr
Street Line 2	
City	La Jolla
State	Ca
Postal Code	92039
Country	USA
Work telephone	555-5555

Briefly describe your interest in using the portal

testing

Edit Request **Submit**

Privacy | Terms of Use | Report a Problem | Feedback

powered by gridsphere

Figure 5.8. Verification Page of the New User Registration Process

The user may review the information, if the user feels that they have entered the correct information, they may proceed by clicking the 'Submit' button. Subsequent to the submission a confirmation page will be displayed to the user, notifying them that the request has been submitted and instructing them to check their email (Figure 5.9). The user will receive an email that is titled "New BDR portal account activation instructions" (Figure 5.9), in the email the user is instructed to click on the link to complete the activation process (Figure 5.10).

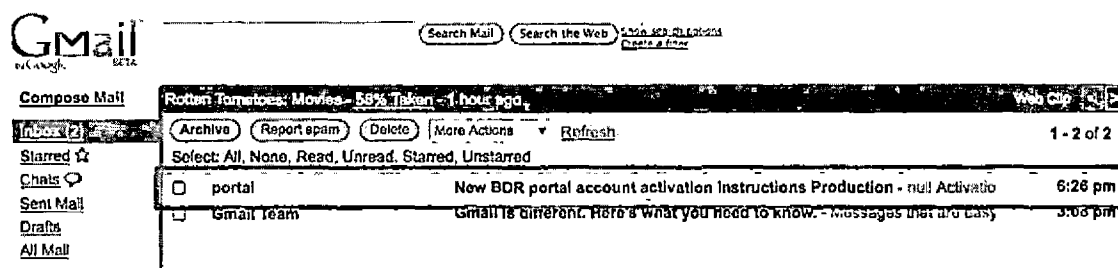


Figure 5.9. Email That Would be Received From the Portal

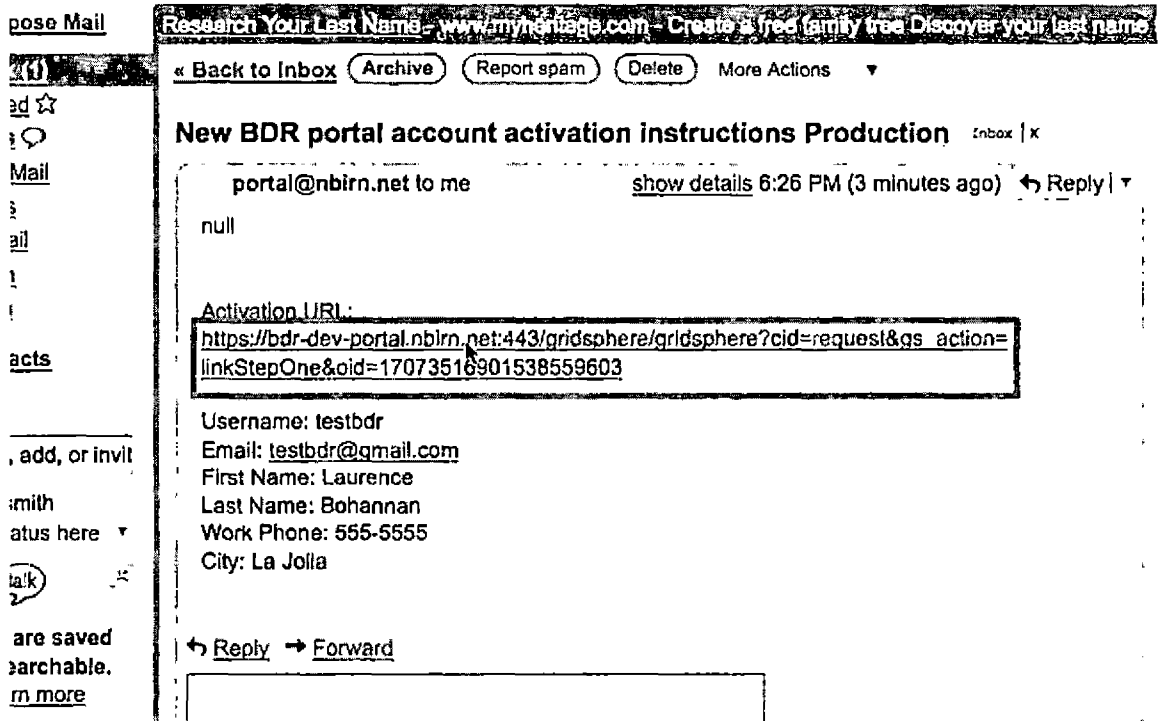


Figure 5.10. View of a Normal Account Activation Email

The link will open a browser page that will redirect the user back to the portal; the user will be prompted to verify their email address (Figure 5.11).

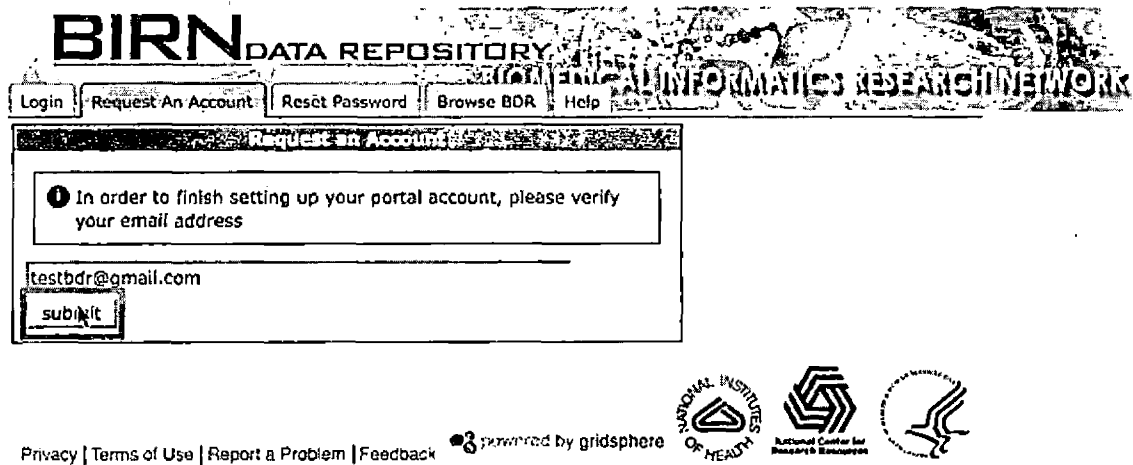


Figure 5.11. Prompt for the Users Email During Activation

After the user has submitted their email address they will need to enter their new passwords for the account (Figure 5.12).

BIRN DATA REPOSITORY BIOMEDICAL INFORMATICS RESEARCH NETWORK

Login Request An Account Reset Password Browse BDR Help

1 Please create a password and then confirm it by typing it again below. Password should be six(6) or more alpha-numeric characters and must contain at least one symbol except the single-quote (').

New Password:

Confirm Password:

submit

Privacy | Terms of Use | Report a Problem | Feedback

powered by gridsphere

NATIONAL INSTITUTES OF HEALTH

National Center for Research Resources

Figure 5.12. Prompting the User to Enter a Password

If everything has gone correctly the user will see a confirmation page (Figure 5.13). However, if there was a problem that occurred during the activation in any step the user will see an error message (Figure 5.14), in this event please contact the system administrator for assistance.

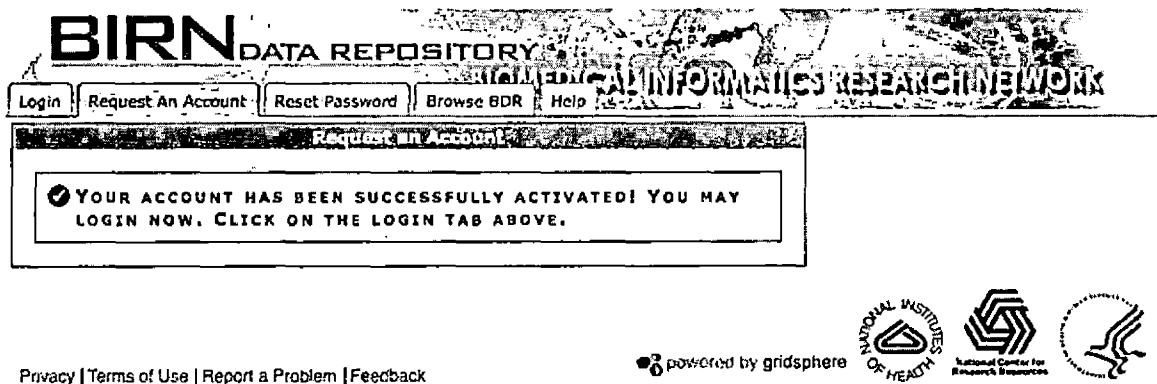


Figure 5.13. View of the Confirmation Message That Would be Generated for the Successful Activation of the Account

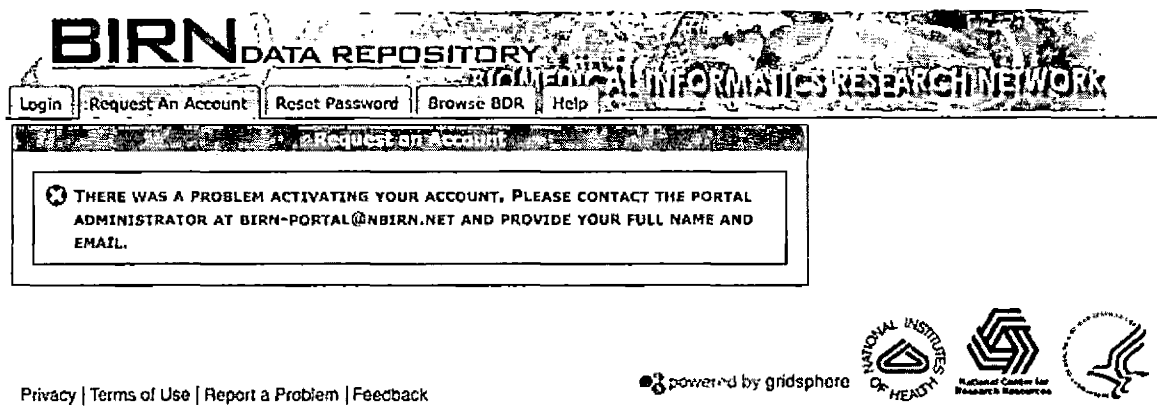


Figure 5.14. View of the Error Message That Would be Displayed if the Account Creation had an Error

5.2.4 Initiating a Biomedical Informatics Research Network Data Repository Submission

A logged in user may create a new submission by clicking on the "Create Submission" sub-tab under the main BDR tab in the portal (Figure 5.15).

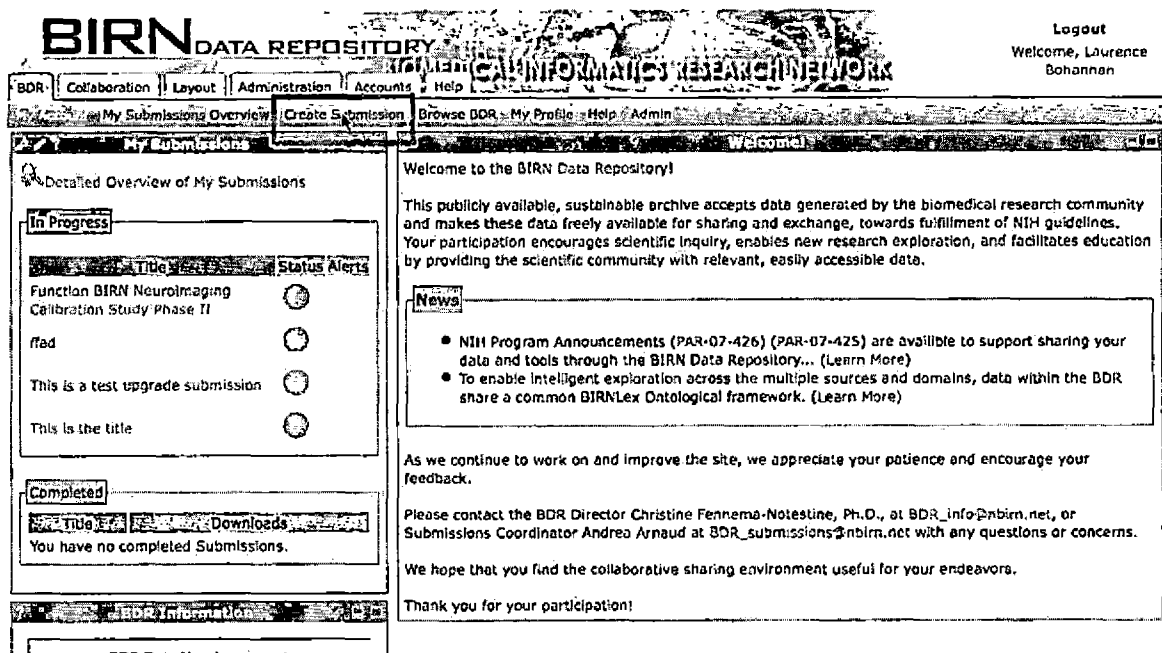


Figure 5.15. Location of the Create Submission Link From the Default View

This action will lead to a page containing the initial page for create submission portlet, the user may enter into the first page of the workflow by clicking on the "Begin

Submission" button located in the center of the portlet (Figure 5.16).

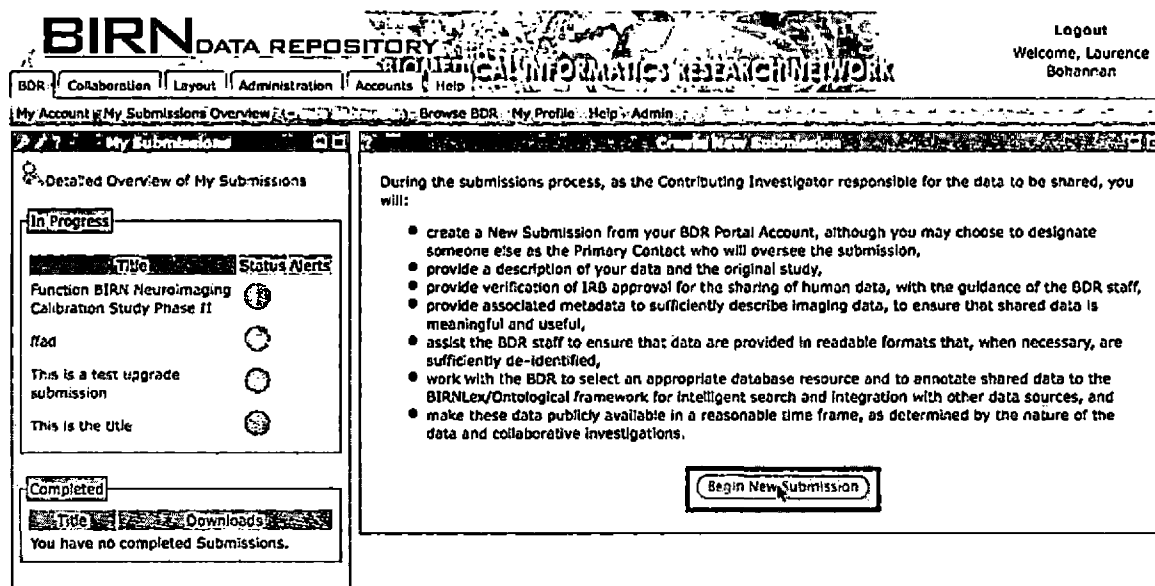


Figure 5.16. Launching a New Submission

5.2.5 Uploading Internal Review Board Documentation

In the IRB upload section of the Submission workflow, the user will be able to upload multiple documents for review by the BIRN Data Repository curators. Performing the following steps listed below, all IRB documents must be uploaded at the same moment. Please don't try to upload each file individually.

1. Enter the document title.
2. Select the extension of the document from the option list. If the extension is not present in the option list, please do not enter it as an upload, please contact the BIRN Data Repository administrators for help with this format.
3. Click the Browse button to find the location of the file on your computer.
4. If there are multiple files to upload, please repeat steps 1-3 until all are entered. Additional files may be added to the form by clicking on the 'Add More Files' link that is present under the first field (Figure 5.17).

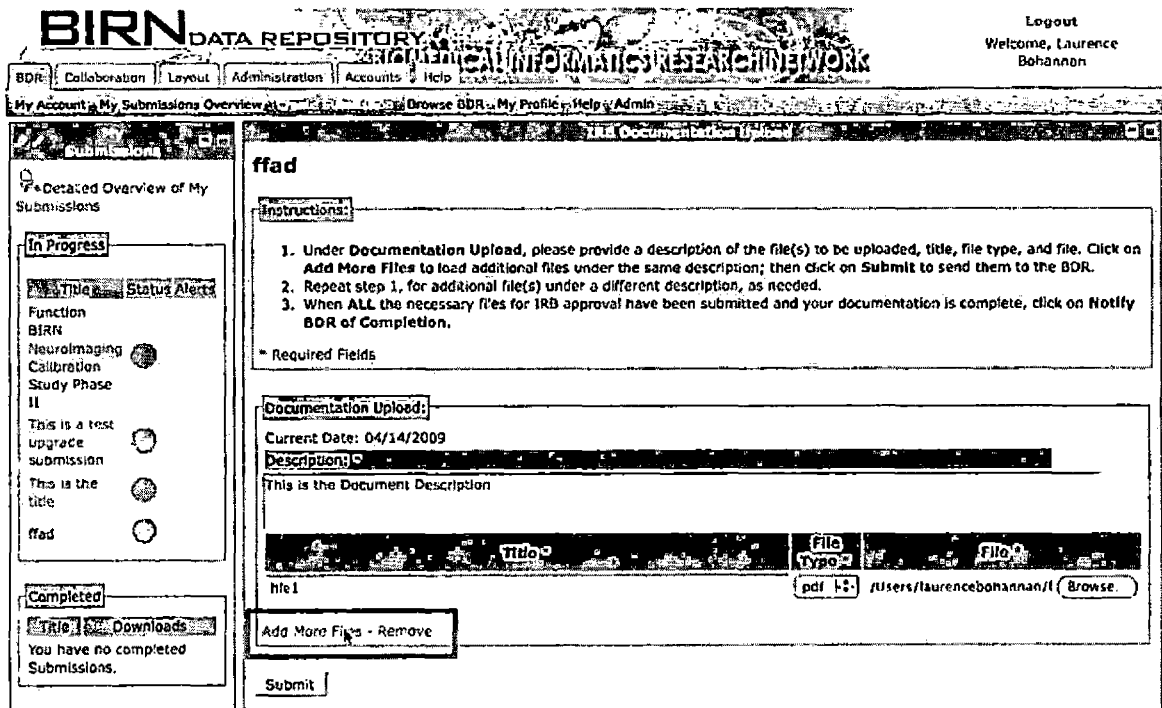


Figure 5.17. Adding Additional Files to be Uploaded

5. Once all your files are entered click the Submit button to initiate the transfer to the server (Figure 5.18). This may take several minutes depending on the size of the selected files.

Once all the necessary files have been uploaded to the BIRN Data Repository, the user is able to review the documentation, this is done through clicking on the corresponding magnifying glass associated with that file (Figure 5.19).

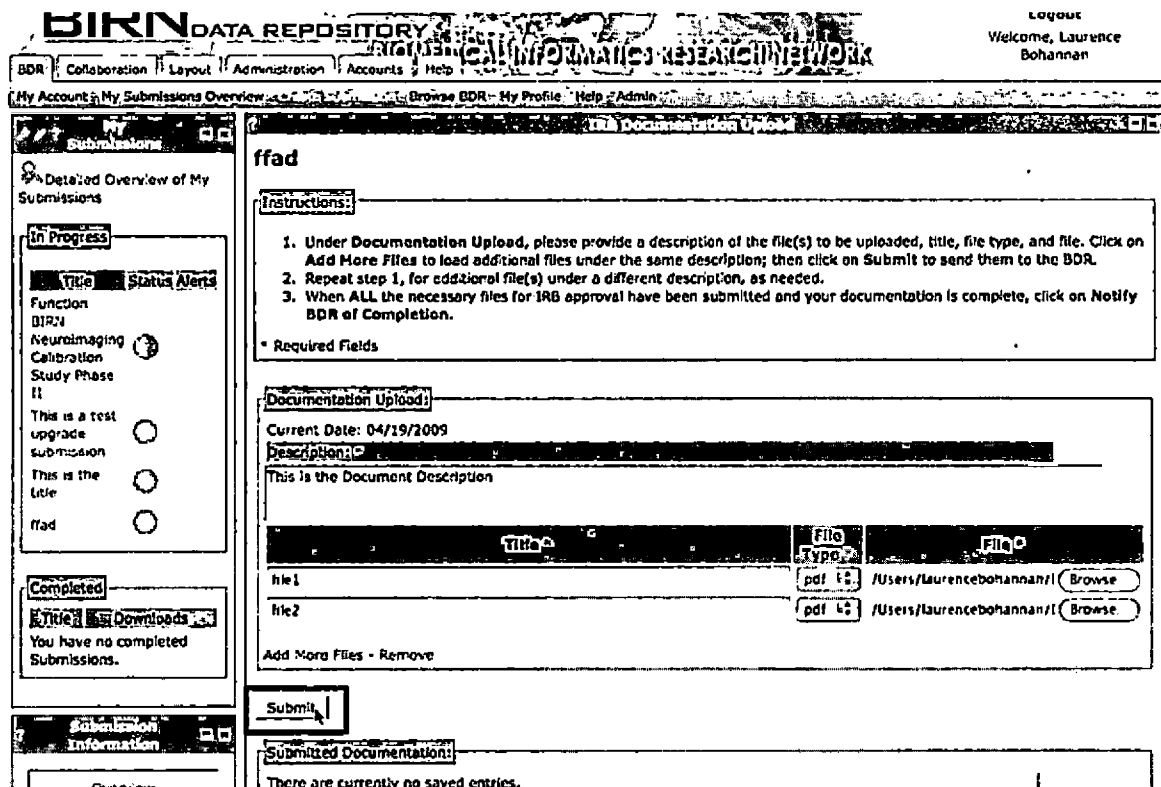


Figure 5.18. Submitting the Files to be Uploaded to the Biomedical Informatics Research Network Data Repository

If the user has decided that they are ready for the curators to review the information they have supplied they must click the button labeled 'Notify BDR of Completion' to post an administrative alert.

Neuroimaging Calibration Study Phase II

This is a test upgrade submission

This is the title

ified

Completed

Title Downloads

You have no completed Submissions.

Overview

Detailed Overview

Associated Metadata

BDR Data Contributor's Agreement

BDR Data Use Agreement

IRB Template for Sharing Human Data

Glossary

* Required Fields

Documentation Upload:

Current Date: 04/21/2009

Description:

Title	File Type	File
		Browse...
		Browse...

Add More Files - Remove

Submit

Submitted Documentation:

Time: Apr 21, 2009 7:06:38 PM

Description:

This is the upload description for the file set

Title	File Type	File
This is the title of file 2	pdf	file2.pdf
This is the title of file 1	pdf	file1.pdf

Notify BDR of Completion

Figure 5.19. Clicking the Magnifying Glass to View the Uploaded File

5.2.6 Sending Notifications to the Biomedical Informatics Research Network Data Research

In the case of a user needing particular attention in the submission workflow whether it maybe for clarification or a problem that has arisen during one of the steps, the user may set an administrative alert. This will not lock the submission, however, it will show the administrator that assistance is needed when they view the submissions overview in the administration portlet. The user can set

this alert, which includes creating a notification message, by clicking "Contact BIRN Data Repository about this submission" link in the users "Submission Specific Overview" page (Figure 5.20).

BIRN DATA REPOSITORY Logout
Welcome, Laurence Bohannon

BDR | Collaboration | Layout | Administration | Accounts | Help CAL INFORMATICS RESEARCH NETWORK

My Account | Create Submission | Browse BDR | My Profile | Help | Admin

Title: Function BIRN Neuroimaging Calibration Study Phase II

Accession Number: 2008-BDR-86PCF

Contributing Investigator: Andreas Arnaud, aarnaud1@cox.net

Primary Contact: Laurence Bohannon, bohannon@ncmir.ucsd.edu

Alerts: There are no user alerts

History of Submission:

- 08/27/2008 at 02:19:04 PM PDT: BDR Admin has moved the submission progress back to the Database Selection step of the 'Data Upload' stage
- 08/27/2008 at 02:18:52 PM PDT: BDR Admin has moved the submission progress back to the Database Upload step of the 'Data Upload' stage
- 08/27/2008 at 02:18:50 PM PDT: BDR Admin has moved the submission progress back to the 'Curation' stage
- 08/27/2008 at 02:06:14 PM PDT: BDR Admin has advanced the submission to the 'Final Review' stage
- 08/27/2008 at 02:06:04 PM PDT: BDR Admin has advanced the submission to the 'Data Curation and Annotation' stage
- 08/27/2008 at 02:06:01 PM PDT: Admin has received your notification message

Continue Submission

Contact BDR about this submission

Please provide feedback to the BDR

Collaborative Workspace

Figure 5.20. Clicking the Link to Lead to the Contact Biomedical Informatics Research Network Data Repository Page

This will lead them to a simple form that the user can complete. Once the form has been completed the user must

click the Submit button to create the administrative alert
(Figure 5.21).

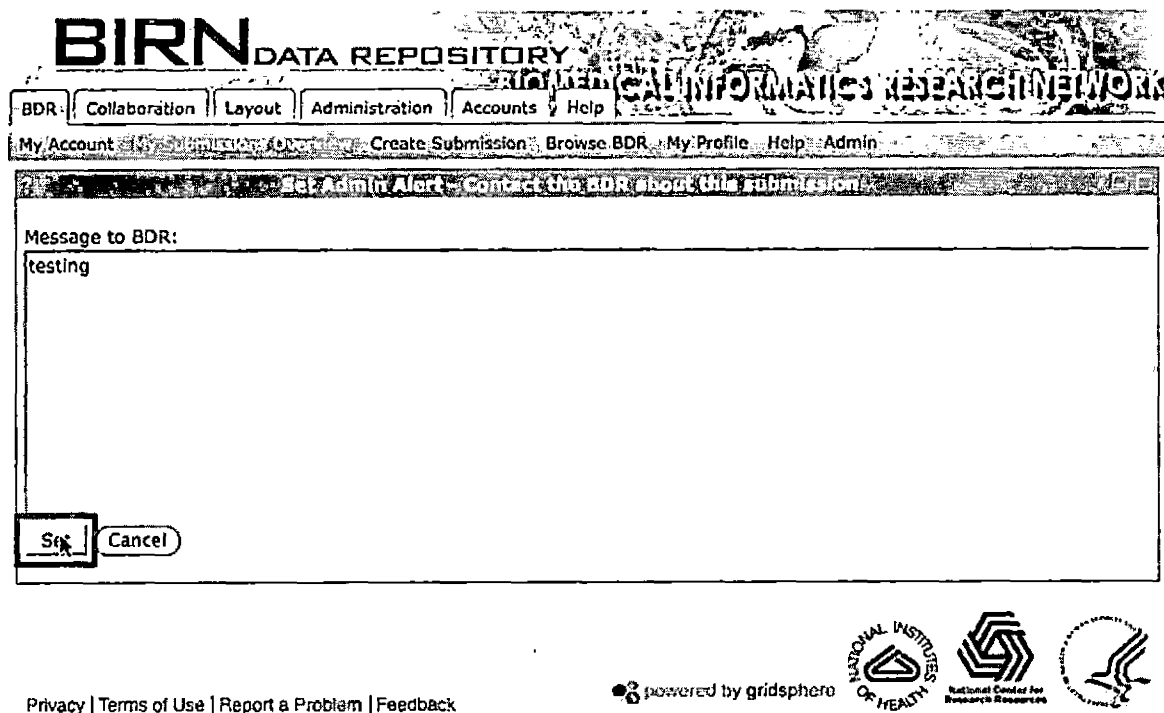


Figure 5.21. Sending the Notification to the Biomedical Informatics Research Network Data Repository

5.2.7 Viewing Submission Progress

Data contributor can view the progress of the BIRN Data Repository submissions of which they have the ownership permissions. Clicking on its corresponding link from the 'My Submission' portlet (figure 5.22) or the

'Submissions Overview' portlet will present a 'Submission Specific Overview'.

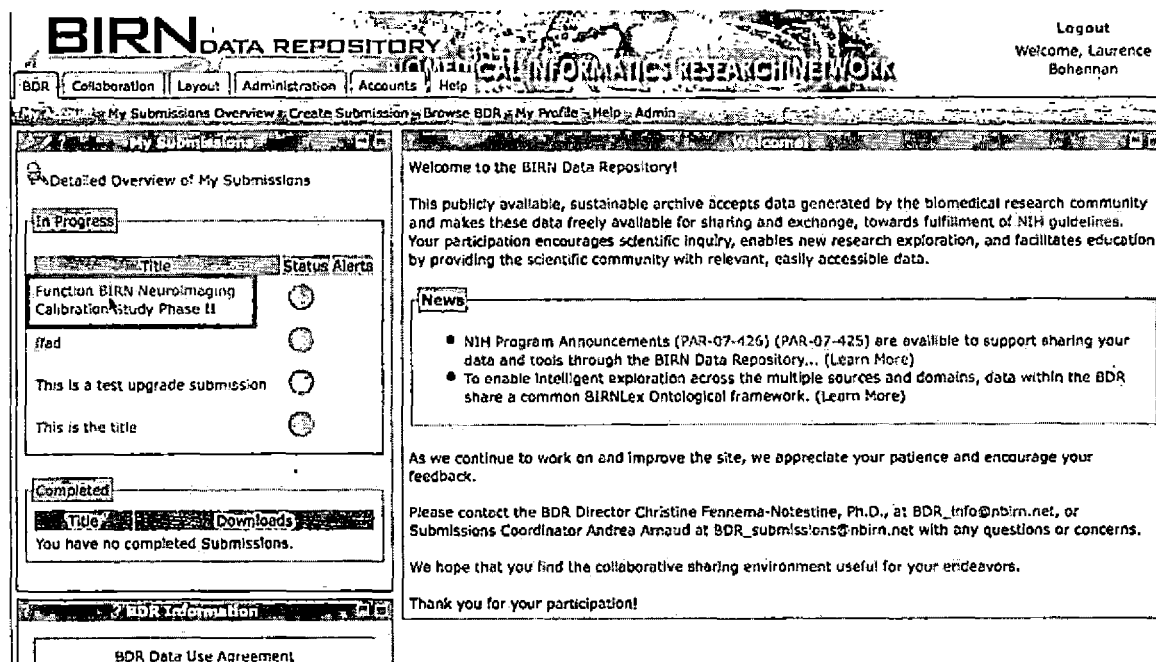


Figure 5.22. Click on a Submission to See a More Detailed View of the Information

The Submission Specific Overview will contain recent alert messages from the administrators, a historical log of actions, and an overview of the overall workflow. The current position of the submission will be shown as a yellow color in the corresponding stage. Completed submission stages will be shown in green and stages that

have not been reached yet will be shown in gray (Figure 5.23).

The screenshot shows the BIRN Data Repository interface. At the top, there is a navigation bar with tabs for BDR, Collaboration, Layout, Administration, Accounts, and Help. A user is logged in as Laurence Bohannon. The main content area is titled 'Submission Specific Overview' and displays the submission title 'Function DIRN Neuroimaging Calibration Study Phase II'. A progress bar shows the submission stages: Data Storage, Curation and Annotation, Final Review, and Published. Below the progress bar, there are fields for Accession Number (2008-BDR-68PCF), Contributing Investigator (Andrea Arnaud), and Primary Contact (Laurence Bohannon). An Alerts section indicates 'There are no user alerts'. A History of Submission section lists several events from 08/27/2008, including moves back to previous stages and advancement to the 'Final Review' stage. On the right side, there are four buttons: 'Continue Submission', 'Contact BDR about this submission', 'Please provide feedback to the BDR', and 'Collaborative workspace'.

Figure 5.23. Detailed View of a Submission

5.2.8 Administration of a Biomedical Informatics Research Network Data Repository Submission

The BDR provides an administration feature to privileged users. The administration portlet is available through Administration tab. This will provide the

administrator with an overview list of all the submissions that are in the submission workflow. This view displays the accession number, title, progress, and alert status for each submission to the administrators. From this view the administrator can then inspect closer a individual submission by clicking on the corresponding list entry. This will lead the administrator to a specific submission overview page. This displays a detailed overview of the submission progress, submission logs, availability to download the most recent IRB documentation, and the ability for the admin to advance or move a submission back to a previous section.

5.2.9 Controlling Submission Progress

When a user has completed some of the latter sections in the submission workflow, the system notifies the administrator when the user has completed the intended action. This notification, which is usually initiated by the user, places the submission into a frozen state that is pending an administrative action. This administrative action is usually intended to be a review of the user-supplied documentation. The administrator has the ability to control the state of the submission as it moves from one point to another. If the administrator deems it necessary

to keep the submission in the current state they may simply clear the hold on the submission and request that the user repeat the previous action in a different manner. However, if the user has completed all the requirements the administrator can advance the submission to the next stage by clicking the "Advance" button from the "Admin - Submission Specific Overview" (Figure 5.24).

Accession Number	Contributing Investigator	Primary Contact
2008-BDR-HNX89	Andrea Arnaut aarnaud@ucsd.edu	Fiona Ho ofspodf@ucsd.edu

Alerts For Admin:	
Sep 5, 2008 10:25:38 AM	User has completed the data section of submission workflow. Clear
Alerts Message to User:	There are no user alerts
History of Submission:	
09/05/2008 at 10:25:38 AM PDT:	Admin notified of stage completion.
09/05/2008 at 10:25:38 AM PDT:	Submission has completed the 'Describe Your Data' stage
09/05/2008 at 10:25:38 AM PDT:	Submission has completed the 'Launch New Submission' stage
09/05/2008 at 10:25:38 AM PDT:	The submission has been created
09/05/2008 at 10:25:38 AM PDT:	Collaborators project space created.
09/05/2008 at 10:25:38 AM PDT:	Admin is notified IRB Documentation is complete.
09/05/2008 at 10:25:38 AM PDT:	Submission has uploaded a set of IRB Documents
09/04/2008 at 05:19:58 PM PDT:	Submission has completed the 'Launch New Submission' stage
09/04/2008 at 05:19:58 PM PDT:	The submission has been created
09/04/2008 at 05:19:58 PM PDT:	Collaborators project space created.
09/04/2008 at 05:19:58 PM PDT:	Admin is notified IRB Documentation is complete.
IRB Documents	There are currently no saved entries.

Recommend a database for the contributor to use:
Set
Print Submission
Set Alert Message
Clear User Alert
Advance Submission
Move Submission Back A Step

Figure 5.24. Links That Correspond to Controlling the Progress of the Submission

Similarly if the administrator would like to move the submission to a previous section on the workflow they must

click the "Move Submission Back" button. The administrator must remember to remove any holds from the submission so that the contributor can proceed with work. The Administrator can only change the state of a submission when it is on a particular portion of the workflow. This eliminates accidental movements of submissions to other pages.

5.2.10 Setting a User Alert

If the administrator needs to notify the user of something pertaining to their submission, but does not

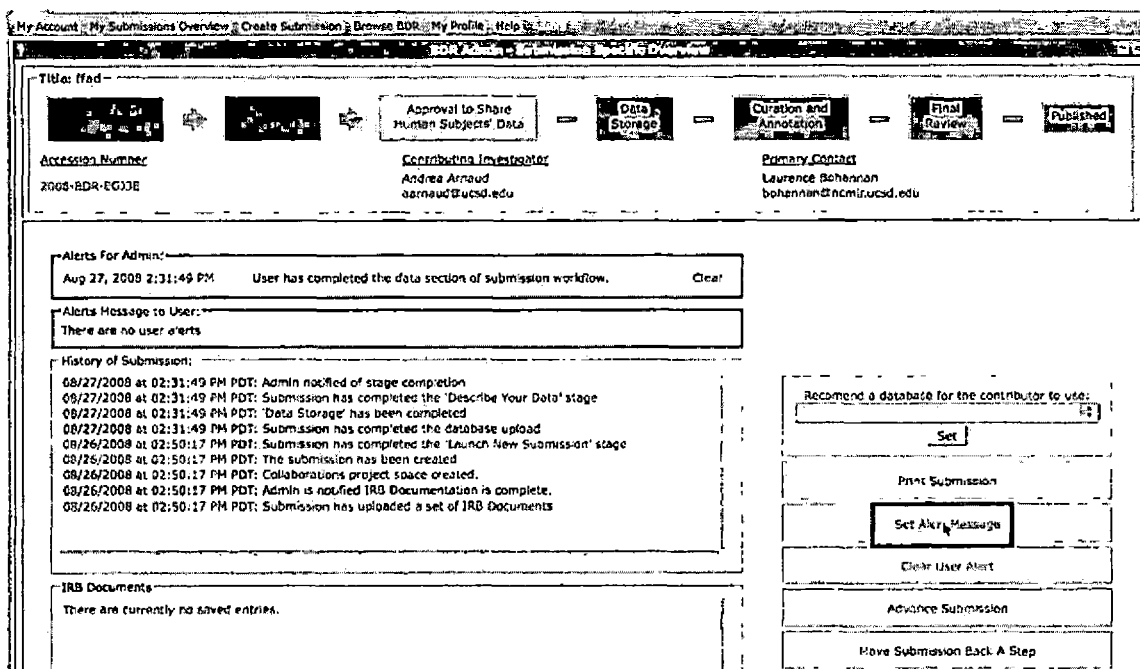


Figure 5.25. Link to Create an Administrative Alert for a Submission

necessarily want to place a lock on the submission, they can create a User alert message. This will appear as an exclamation point icon next to the users submission entry in the "My Submissions" portlet. Once the user logs in they would be able to see this notification. The user alert notification can be set from the "BDR Admin" portlet detailed overview page by clicking on the "Set Alert Message" link (Figure 5.25). This will lead the administrator to a simple form that can be filled out and submitted to create the user alert.

5.2.11 Clearing an Administrative Alert

When a user has a question or has completed a particular task there is an administrative alert that is set. This administrative alert will remain active until it is cleared by the administrator. Depending on the action that triggered the alert, this may place the submission into a locked state. This will prevent the user from advancing and moving it back to a previous step, this will appear as a red highlighted notification message above the admin alerts (Figure 5.26).

BDR Admin Submissions Overview			
Accession #	Title	Status	Alerts
2008-BDR-M3M97	Mouse Model of Parkinson's Disease		
2008-BDR-67AFE	Dopamine transporter knockout mouse MRI		
2008-BDR-6QF87	Function BIRN Neuroimaging Calibration Study Phase II		
2008-BDR-YBN1H	jldjfaljdffaf		
2008-BDR-XQGYS	MRM Reeler Mouse		
2008-BDR-WM64G	MS Mouse Atlas		
2008-BDR-MZGZJ	MRM Reeler Mouse		
2008-BDR-6P5PF	dgsdgsfdgs		
2008-BDR-88PCF	Function BIRN Neuroimaging Calibration Study Phase II		
2008-BDR-G3ZA9	MRM Reeler Mouse		
2008-BDR-EGJ3E	ffa		
2008-BDR-3777G	This is a test upgrade submission		
2008-BDR-R7UGE	MS Mouse Atlas		
2008-BDR-MMBNQ	khgjghkjgjk		
2008-BDR-HNX89	fds		
2008-BDR-8A4Y7	sfsdfgs		
2008-BDR-ERXA3	Multi-site calibration		

Figure 5.26. Default View of the Admin Portlet

An administrative alert can be cleared from the "BDR Admin" portlet detailed overview page by clicking on the "Clear" link accompanying the notification message (Figure 5.27).

If there is a lock on the submission for that particular notification, it will be removed.

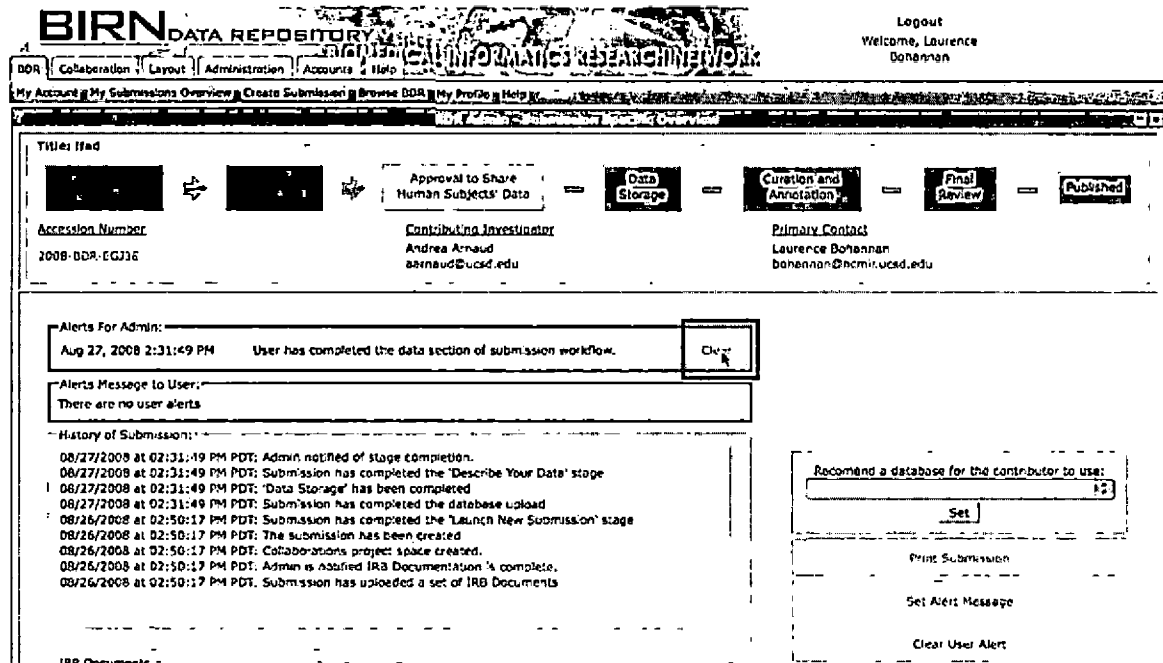


Figure 5.27. Clearing an Alert From the User to the Admin

5.2.12 Printing a Submission

An additional function that is available to a user with administrative privileges is the ability to print all the information associated with a submission. This is done by clicking on the print submission link located on the

"Submission Specific Overview" view of the BDR Admin Portlets (Figure 5.28).

The Portlet then displays information associated with the submission in a formatted list view (Figure 5.29). All information is present except for any associated documents that were uploaded during the IRB Information section of the workflow, these documents can be downloaded and printed out from the IRB Document pane of the "Submission Specific Overview."

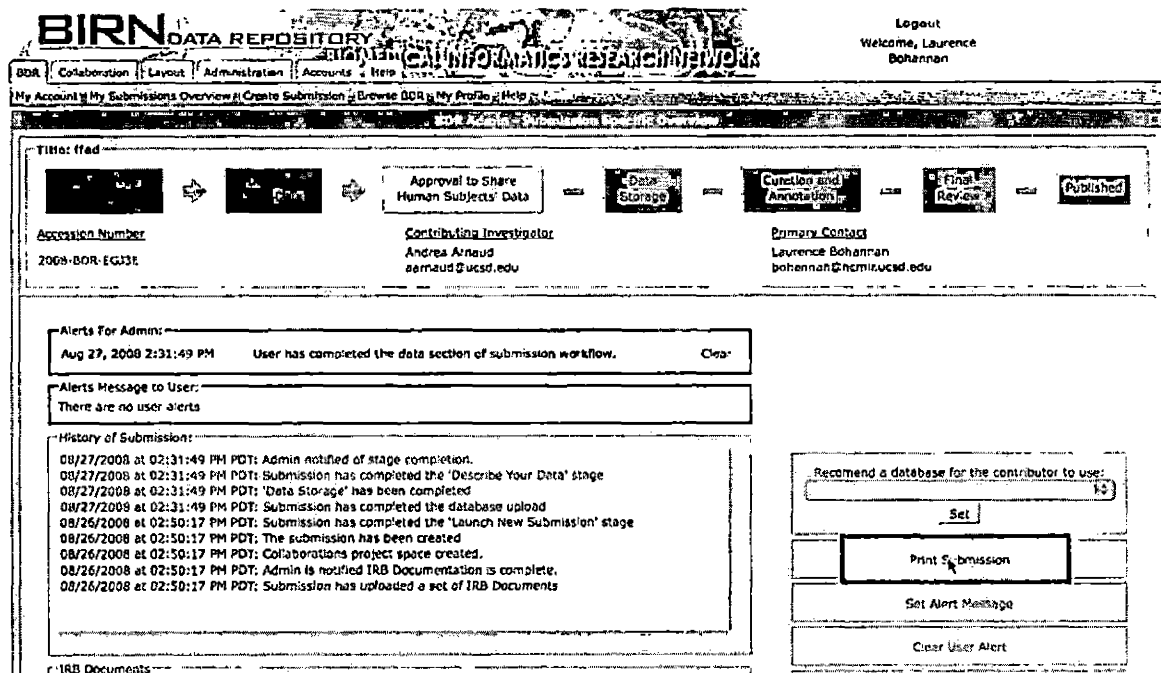
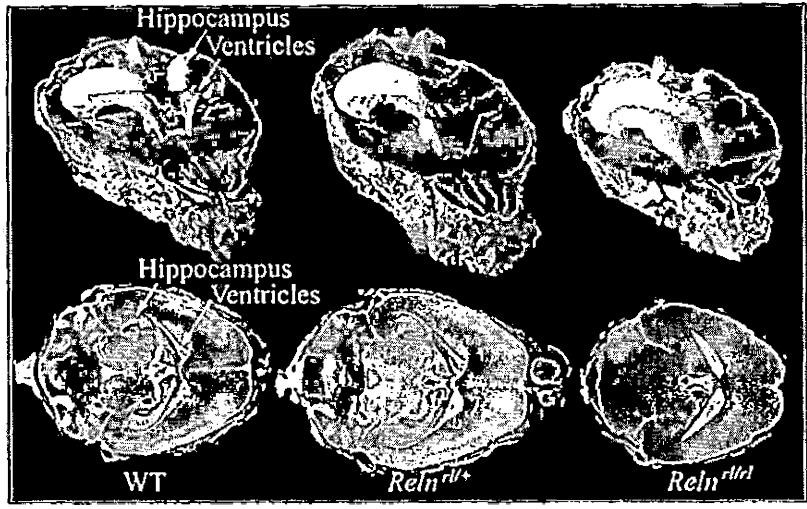


Figure 5.28. The Link That Leads to the Print Feature in the Admin Overview Page

Submission Title: ffad
 Study Description / Abstract: ffad
 Keywords: B6;D2-Tg(PDGFB-SNCA)NNEma
 AttributeDateofSacrifice
 B6;129-/Stc6a3tm2Mca
 Contributing Investigator: Andrea Arnaud
 Contributing Investigator e-mail: aarnaud@ucsd.edu
 Organization: UCSD-BIRN Coordinating Center
 Phone Number: 858.534.6844
 Primary Contact: Laurence Bohannon
 Primary Contact e-mail: bohannon@ncmir.ucsd.edu
 Related to an existing BIRN project:
 Affiliated Institution(s) & URL(s): UCSD-BIRN Coordinating Center
 Investigator(s) & Contributor(s): Andrea Arnaud (Contributing Investigator)
 Sponsor(s):
 Funding Source(s)/Grant Number(s):

Figure:



Caption: Magnetic resonance microscopy at 21.5 mm used to study the neuroanatomical phenotype for the reeler mouse (Rein). The ventricles (cyan) and hippocampus (yellow) were segmented in wild type (WT), heterozygous, and homozygous Rein mutants.

Primary Publication(s)/Original Article(s):
 Additional Related Publication(s):
 Data Types: Clinical measures
 Microarray
 Behavioral/cognitive measures
 Image Types: Functional MR
 Diffusion
 Number of individual cases in sample: 3

Figure 5.29. View of the Print Submission Page

5.3 Improvements

The BIRN Data Repository was developed in a relatively short time span with multiple revisions to the structure of the workflow throughout the development process. Now that the majority of the steps of the workflow have been firmly established for the version 1.0 release, the code base can be refactored to optimize the database accesses for the data objects. In general most of the attributes of the submission data object are currently persisted to the same table, due to the nature of the hibernate-based persistence manager these queries return very 'wide' rows. These large result sets can easily add up when querying the database for multiple entries. For the first version of the software, the BIRN Data Repository utilization will be low, so it is very unlikely that a un-optimized design structure like this will make a significant impact on the performance of the system. Additionally, due to a hibernate class loading constraints with the portlet services that utilize the GridSphere persistence manager, a custom hibernate persistence framework would improve the performance by enabling the lazy loading feature, which is usually enabled by default in applications that utilize the Hibernate framework.

Another aspect that can improve the overall maintainability of the software would be to refactor the names of some of the objects, services, and JSP pages that are used in the system. It becomes apparent once a person that is tasked with updating or adding features to the BIRN Data Repository will change significantly the design and aim of the system from the original inception of the project. Hence in some locations of the source tree, the file names may not accurately represent the function or contents of the file. Similarly, there are several objects that may represent a non-fully implemented feature or likewise, a feature that was removed due to time constraints.

A feature that would greatly help the tidiness of the portal would be to implement a soft delete function for a submission. It was originally thought that this feature would not be needed as urgently since BIRN would not want to delete the actual data from the tables of the database. Also, the rate for additions of new submissions would have allowed more than enough time to include this feature in a later release. However, if BIRN would have been able to implement this feature, they would have opted for a flag method to simply ignore the submission that have been

marked for deletion and if really needed, the entries could be removed through an administrative system dump of the database. Improvements like these coupled along with user feed back from the BIRN Data Repository portal on the overall process of the workflow will help improve the usability of the features implemented in the submission process and maintainability of the system as a whole.

CHAPTER SIX

CONCLUSION AND FUTURE DIRECTIONS

6.1 Conclusions

Computers have become a vital part of scientific research, more so than just the vast amounts of calculations that they enable researchers to perform in a fraction of the time, but also for the new possibilities for correspondence and collaboration with a new media. In an interview with Jeffery Grethe, the scientific coordinator for the BIRN coordinating center at UCSD, on the usefulness of the system, he reiterated how the creation of a broad-based, federated data repository that enables intelligent search across multiple resources, which are comprised of a variety of data collected from different species, will drive development of new hypotheses and provide an abundance of collections for training and research.

Additionally he didn't fail to mention how a data sharing infrastructure like the BDR promotes many goals of the NIH research endeavor through the public exchange of data which enables new research exploration, as investigators from different disciplines view the richness

in data acquired for different purposes. The BIRN Data Repository encourages scientific inquiry by enabling investigators with diverse backgrounds to examine data related to their own area of specialty, to pool data across multiple sites, and to study rare data made available to a wider audience. Investigators who may not have specialized resources, such as large-scale computational resources for intensive analyses, novel imaging techniques, or access to clinical populations, are able to address broader-reaching questions relevant to their scientific hypotheses.

The power of the BIRN Data Repository is demonstrated by increased sensitivity in investigations where data may be combined, and by providing a common resource for test and training data to enhance development methods. The accessibility of resources, in the BIRN Data Repository and advancements derived from these resources, will go on to promote education and help train multi- and interdisciplinary investigators, students, and research teams, providing a resource complementary to clinical inquiry, data collection, data management, and ethics training (NIH Data Sharing Policy).

When the system went public in late August of 2008 the main functions of the BIRN Data Repository was to assist in

the fulfillment of the NIH guidelines for the publication of stored data and to further increase the availability of data to researchers, not just existing users of the network but to public researchers that may have no affiliation with BIRN. Initial usage came from the researchers that currently share their datasets with BIRN on the static pages. The BIRN Data Repository portal provides researchers with a public location to post their newly obtained datasets to the public in a streamlined and tailored system. This publicly available data must be cleaned and prepared prior to its release on the Internet especially if the data is derived from human subjects. The system will aid in the submission process for the publication of datasets on the Internet. By streamlining the submission workflow, the researchers will be guided through the formalities and the steps, which are necessary to prepare the data and receive the appropriate approvals for the publication.

Additionally, the BIRN Data Repository will provide an interface in which researchers can easily search for information of interest, with capabilities like searching for keywords. The system is intended to aid researchers who are interested in contributing and sharing their obtained

data to the public community, as would researchers interested in obtaining data on a specific topic would find the site very useful. The system was developed on the Gridsphere portal framework backed by a Postgres database and incorporated several BIRN package libraries to aid in utilizing some of the preexisting functionalities; mainly these include the Storage Resource Broker (SRB) portlets and the collaborations (Bluesquid) portlets.

6.2 Future Direction

The BIRN Data Repository portal is an initial version to provide the basic functionalities of submission creation to begin the dataset submission and proof of concept for the BDR search capabilities.

There is a wide variety of ways for a project as large as the BIRN Data Repository to be improved upon. First of all, the data that has been submitted to the submission workflow can be more tightly integrated into the existing features of the Bluesquid collaborative portlets. Some of the information that was gathered through the workflow into lists such as publications, citations, and affiliated institutions can be published into searchable RSS feeds. This would increase the likely hood of external search

source finding pertinent information and increasing the overall traffic to the system. Due to limited time for the development cycle of the BDR some features had to be delayed for a following release, of these they include: detailed usage or statistic tracking for dataset downloads and submission creation. Additionally, the implementation of a system favorites portlet, these are features like suggested datasets based on recent releases and persisted topics of interest for repetitive searches for logged in users are capabilities that would be very useful for existing users.

One other task is to improve the user interface of the submission workflow. Not all pages may necessarily present the most intuitive layout or implementation of a particular function. This will be a continuing task throughout the whole development process. In the workflow the IRB upload is currently separated from the data upload, there are some scenarios that arise in which it is not the most optimal placement process wise for the documents to be associated to the data separately. This may lead to situations of excessive confusion, suboptimal communication between the data contributor and the IRB or possibly inadequate obscurification of human data. The efficient implementation for a fast search method for individual keywords from the

entire UMLS dictionary will need to be addressed relatively soon since it is a popular database for concept names which are used in the identification of biomedical artifacts. This may prove to be difficult since the current size of the list is approximately 5 million terms. However this does not account for duplicate terms, a possible solution would be to implement an Ajax based dynamic search as the user types characters. It would significantly reduce the amount of information that would potentially be displayed at one moment.

There are several security issues that need to be addressed in the subsequent versions. First, is that of fictitious emails when it comes to guests who would like to register for the first time a confirmation email is sent out to the corresponding email addresses that was provided in the registration form. This could be a source of malicious spam emails if someone where to create a script to exploit this feature by constantly hitting the page to fill out the needed parameters and provide a target email to spam. One possible solution would be the incorporation of a captcha into the registration form to try to ensure that the submitter is human and not an automated script. Another possible problem that may arise is from that of

inappropriate material. Even though the submission process accounts for this through the curation step of the workflow, it is only a quick review for non-human data submissions. This review is at no lengths nearly as extensive of a scan as compared to that which is preformed on the human data submission. Taking into account the wide variety of the data formats, volume and the fact that this is largely an error prone visual inspection which is administered by one of the curators, it would only be expected before an overlooked submission would be able to make its way through.

Additionally, the current implementation of the term search of published datasets in the BIRN Data Repository incorporates a fairly naive pattern search to return the results. However, with further research into the development efforts to transition to a more semantic Web environment along with the expansion of the BIRN Lex terms, the need for implementing an integrated ontological based search method has become a very important next step in the process of returning more relevant published datasets. A greater amount of time must be spent on the implementation of a semantic search feature. Despite how naive it may be in the initial stages of the

implementation, the eventual incorporation of such a feature would have profound results on the effectiveness of the searches performed by users. The BIRN scientific community continues to develop its ontology for neurological artifacts, however this process can be difficult due to the fact that a consensus must be reached before any conclusion can be drawn on what features or methods may constitute a term. Possibly one aspect of development may incorporate a method to discuss these possible terms in a more collaborative environment. Nonetheless the eventual development of the ontology will benefit the BIRN Data Repository, in the accuracy of the results returned and the number of degrees to which a user may search for a term.

BIBLIOGRAPHY

- [B1] Asleson, Ryan, and Nathaniel T. Schutta. *Foundations of Ajax (Foundation)*. New York: Apress, 2005.
- [B2] Bloch, Joshua. *Effective Java Programming Language Guide*. Boston: Addison-Wesley, 2001.
- [B3] Cederholm, Dan. *Web Standards Solutions The Markup and Style Handbook (Pioneering Series)*. Richmond: Friends of ED, 2004.
- [B4] Christian., Bauer,. *Hibernate In Action*. Greenwich, CT: Manning, 2005.
- [B5] Fowler, Martin. *UML Distilled A Brief Guide To The Standard Object Modeling Language*. Reading, Mass: Addison Wesley, 2000.
- [B6] "Gridsphere Documentation." *www.gridsphere.org*. 19 Mar. 2008 <<http://www.gridsphere.org/gridsphere/gridsphere/guest/docsTab/r/>>.
- [B7] *IEEE Recommend Practice of Software Requirements Specifications*. Tech. no. IEEE Std. 830-1998. 1998.
- [B9] Linwood, Jeff, and Dave Minter. *Building Portals with the Java Portlet API (Expert's Voice)*. New York: Apress, 2004.
- [B10] Novotny, Jason, Michael Russel, and Oliver Wehrens. "GridSphere: An Advanced Portal Framework." *IEEE*

Proceedings of the 30th EUROMICRO Conference, Rennes, 2004.

- [B11] Peak, Patrick. *Hibernate Quickly*. Greenwich, CT: Manning, 2005.
- [B12] "PostgreSQL: Documentation: Manuals: PostgreSQL 7.3: PostgreSQL 7.3.21 Reference Manual." *PostgreSQL: The World's Most Advanced Open Source Database*. 19 Mar. 2008
<<http://www.postgresql.org/docs/7.3/static/reference.html>>.
- [B13] "What Is A Portlet?" *What Is A Portlet?* 19 Mar. 2008
<<http://www-3.ibm.com/software/webservers/portal/portlet.html>>.