

# CS 4320 Final Project

## OCDX Engine

Requirements Analysis, System Design, and Sprint Documentation

[GitHub Repository](#)

[Deployment Environment](#)

### Team 4 Members:

Holt Skinner

Justin Hofer

Ali Raza

Bo Zhang

Bradley Rogers

Pramod Pudotha

## Table of Contents

Requirements Analysis	3
User Descriptions	3
Use Cases	3
Functional Requirements	4
Non-Functional Requirements	4
User Non-functional	4
System Requirements	4
User Requirements	5
Initial Software Design	5
Sketches	5
Data	8
Database Design	9
Sprint 1	10
Kick-off Meeting	10
Deployment Environment	10
Organize Github repo	10
Database	11
Testing	11
User Acceptance Test (UAT) Scenarios	11
Unit Test Scenarios	12
Regression Testing	13
Integration Testing	13
User Uploads Manifest	13
User Reviews Manifest	13
System Admin Bans an Illegal Data Scientist	13
System Admin Deletes an Illegal Manifest	13
Failure Case	13
Testing Edge Cases	13
Testing Summary	14
Sprint 2	14
Meeting Notes	14
Database	15
Python code for inserts, updates, searches, and deletes.	16
Unit Tests	17
User Interface	22
User Accounts	22
Testing Updates	22
Sprint 3	23
User Interface Temporal Logic	23
Testing Updates	24
User Documentation	24
Meeting Notes	24
11-15-2016	24
11-16-16	25

11-17-16	25
Sprint 4	26
Meeting Notes	26
11-27-16	26
Database Changes	27
Deployment Instructions	27
User Documentation	27
Automated Deployment	27
Unfinished Business	28
Past Sprint 4	28
Wrap-Up Meeting	28
Front-end Updates	28
Business Logic Updates	29
Database Updates	29
Deployment Updates	29
Hindsight is 20/20	29
Change Log	30
Glossary	30

## Requirements Analysis

### User Descriptions

*Rogers*

- **Data Scientist**
  - A Data Scientist uploads and downloads data.
  - A Data Scientist conducts sophisticated and systematic analyses of data.
  - A Data Scientist Extracts insights from large data sets.
- **System Administrator**
  - A System Administrator reviews the uploaded data.
  - A System Administrator can manage users.

### Use Cases

*Skinner/Hofer*

- A Data Scientist can browse manifests.
- A Data Scientist can search for a manifest by title, author, or date added.
- A Data Scientist can contribute to existing dataset.
- A Data Scientist can download information from uploaded manifests.

- A Data Scientist can upload a new manifest.
- A Data Scientist can save manifests.
- A Data Scientist can upload a data set.

## Functional Requirements

### *Skinner*

- The system will take an inputted manifest and place it into storage.
- The system will retrieve a manifest from storage and present it to a data scientist.
- The user interface can accept and serve search, upload, and update requests to and from data scientists.
- The system will process search, upload, and update requests, both into and out of database layer.

## Non-Functional Requirements

### *Hofer/Raza*

#### User Non-functional

- Data Scientists will be able to perform efficient searches based on keywords.
  - Users should not wait more than about two seconds for a query to resolve.
- Data Scientists will be able to upload and update manifests.
- Data Scientists will be able to upload and download data files.

## System Requirements

### *Raza/Hofer*

#### Space Requirements

- The system must have enough physical space to handle extremely large data-sets.
- The system may need to be a distributed system using clusters for efficiency.

#### Reliability Requirements

- If the system crashes, the data and any manipulations/analytics must be preserved.

#### Privacy Requirements

- The data access must be limited to authorized users only.
- The system must enable secure data transfer over the internet.

## Web Server

- Web Server must provide reliable service for the appropriate amount of traffic that will be sent and received from the system.

## Database Backend

- Database must be able to convert data into easily storable format, and return in original format.

## Storage Array

- System must create intermediate backups and update logging to revert to earlier states if needed.

# User Requirements

## *Zhang/Rogers*

- Data scientists will be able to upload Manifest.
- Manifest can be reviewed by other data scientists.
- Data scientists can edit or delete their Manifest.
- Data scientists can search for Manifest they wish to view or test.
- Data scientists will be able to include special comments or suggestions on manifest.
- Data scientists should be able to notify the changes or suggestions that improves manifest to other users.
- System admins can ban an illegal data scientist.
- System admins can delete an illegal manifest.

# Initial Software Design

## Sketches

## *Skinner/Rogers*

OCDX

Logout

BROWSE UPLOAD

Upload

Title:

Description:

Availability: ☐ Public ☐ Private

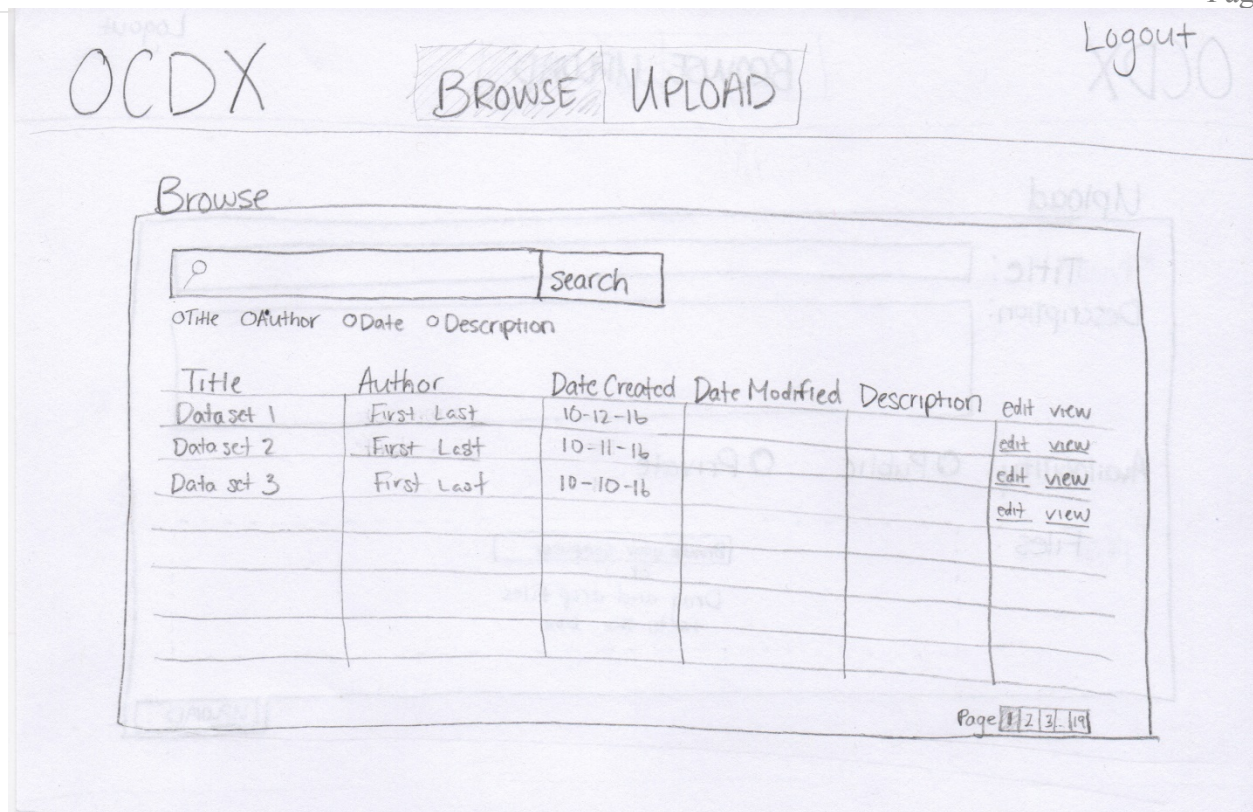
Files: 

Browse your computer  
or  
Drag and drop files  
into the box

UPLOAD

### UPLOAD NEW MANIFEST(S)

FROM THIS PAGE, A DATA SCIENTIST CAN UPLOAD A NEW MANIFEST AND SCRIPTS, SET A TITLE FOR THE MANIFEST AND SET A DESCRIPTION.

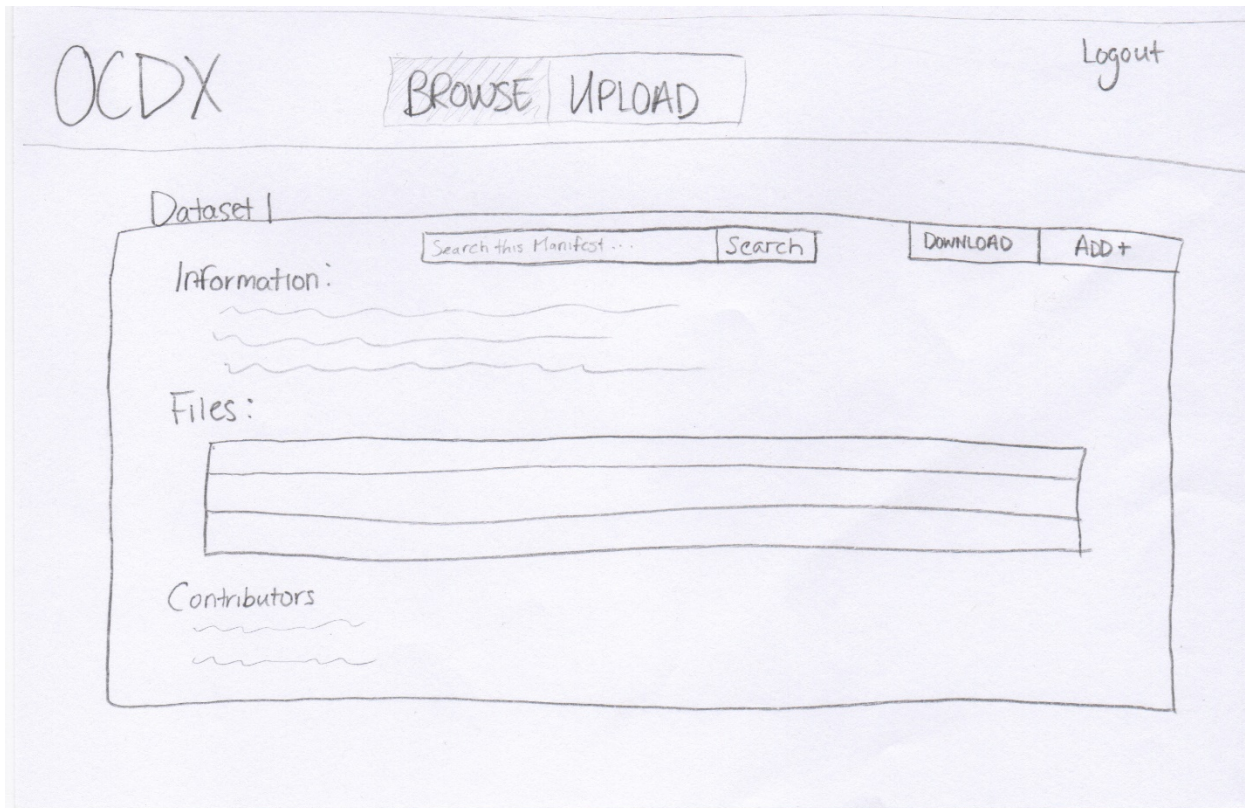


## BROWSE UPLOADED MANIFESTS

A DATA SCIENTIST CAN SEARCH AND BROWSE UPLOADED MANIFESTS.

DATA SCIENTISTS CAN ORGANIZE THE SEARCH RESULTS BY TITLE, AUTHOR, DATE OR DESCRIPTION.





## MANIFEST DESCRIPTION/EDIT MANIFESTS

THIS PAGE PROVIDES THE FULL INFORMATION FOR A GIVEN MANIFEST.

A DATA SCIENTIST CAN VIEW THE IMPORTANT INFORMATION ABOUT A MANIFEST INCLUDING THE TITLE, AUTHOR(S), AND DESCRIPTION.

A DATA SCIENTIST CAN CONTRIBUTE TO A MANIFEST IF THEY HAVE PERMISSION TO DO SO.

A DATA SCIENTIST CAN DOWNLOAD UPLOADED FILES FROM A GIVEN MANIFEST.

## Data

*Zhang/Hofer*

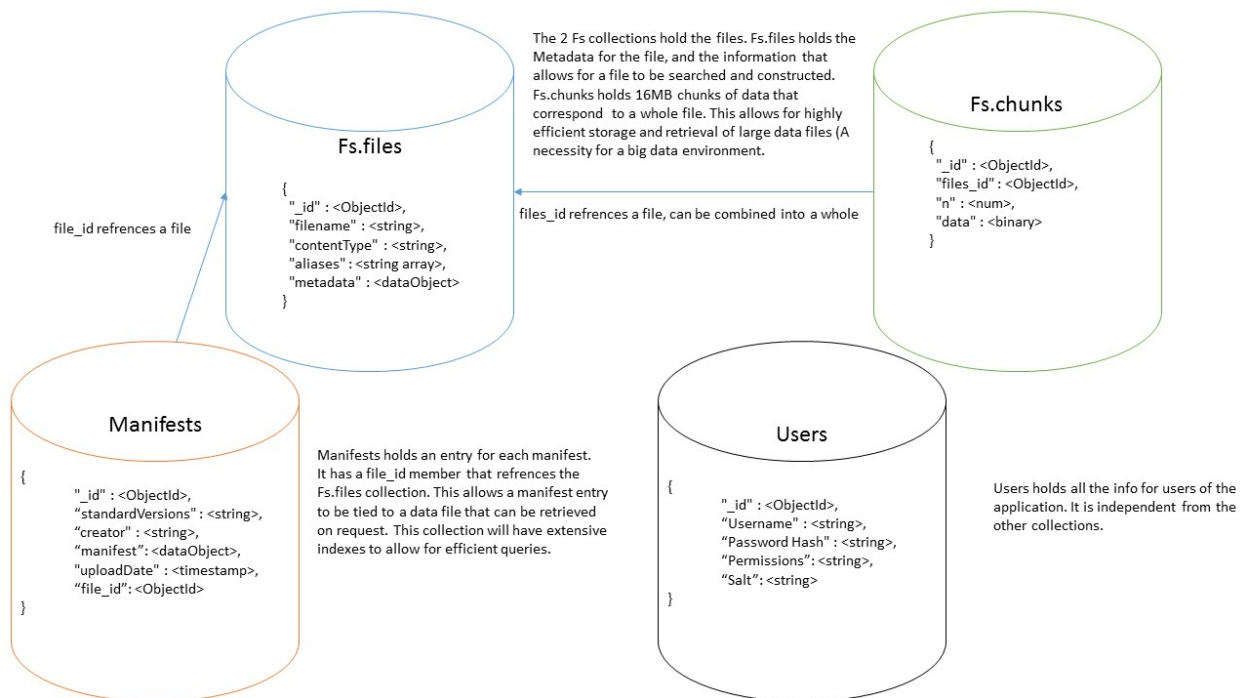
- **Members**
  - ID
  - name
- **OCDX\_manifest**
  - author
  - date



- size
- category
- version
- modify\_request
- **Users**
  - ID
  - name
  - email
  - phone
- **Login\_attempts**
  - userID
  - date
  - attempts

## Database Design

*Hofer*



## MONGODB PSEUDO-ERD

# Sprint 1

## Kick-off Meeting

*Time: 10/22/2016*

*Location: Lafferre Hall*

### Team Members:

Holt Skinner, Justin Hofer, Ali Raza, Bo Zhang, Bradley Rogers, Pramod Pudotha.

### Decisions Made

Materialize framework by Google will be used for UI.

MongoDB is selected to be the database.

PHP will be used for business logic.

The website will be deployed on Amazon Web Services.

### Task Lists

1. Complete setup of deployment environment. (Ali)
2. Organize Github repo and Slack Channel (Holt)
3. Design of user interface. (Holt, Bradley)
4. Finalize ERD. (Justin)
5. Build user acceptance scenarios, unit test scenarios, and other testing plans. (Bo)

### Remaining Tasks

1. User interface implementation.
2. Database implementation.
3. Tests implementation.

## Deployment Environment

- Hosting Platform: [Amazon Web Services](#)
- Operating System: [Ubuntu 16.04.1](#)
- Web Server: [Apache HTTP Server 2.4](#)
- Database: [MongoDB](#)

## Organize Github repo

- Documentation

- Sample Manifests
- Source code

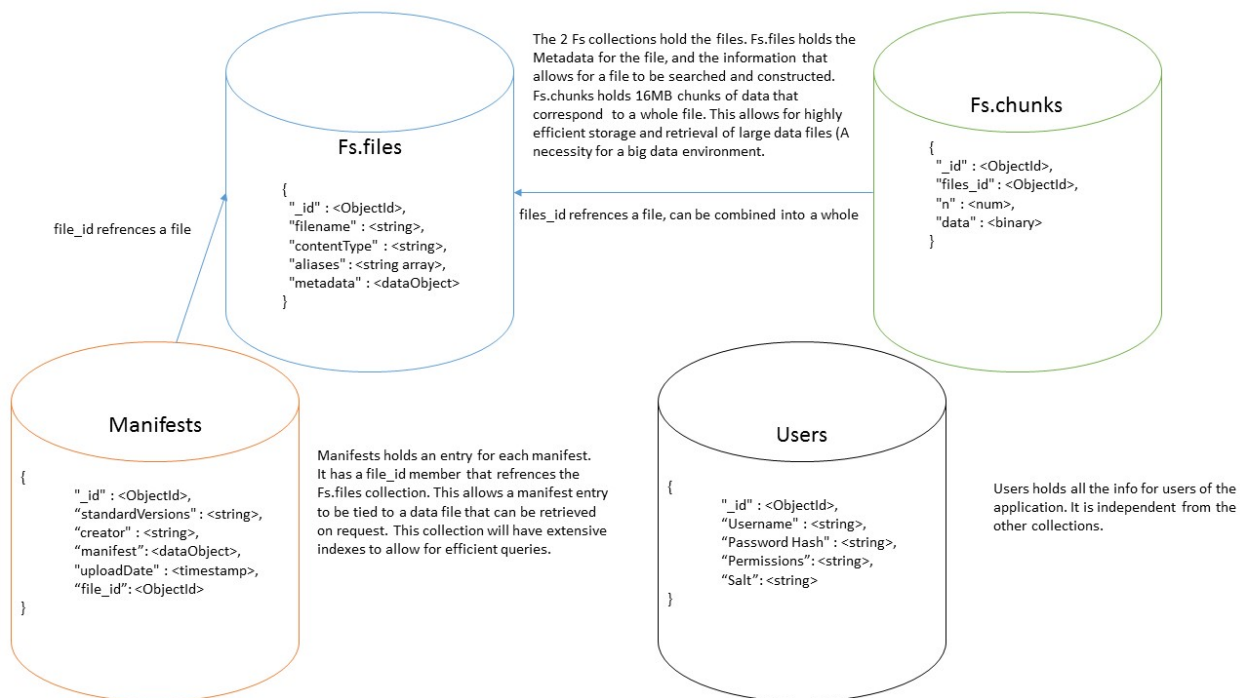
## [CS4320-Final Project Repository](#)

## Database

### Creation and Seeding

- Create a collection named “Manifests”: [Database Creation Commands](#)
- Create sample manifest: [Sample Manifest](#)

### Design Architecture



## Testing

1. Build user acceptance test scenarios for documented requirements on separate Wiki page, linked to all sprints.
2. Build unit test scenarios.
3. Describe regression testing and your regression testing plan.
4. Describe how your team will perform integration testing. What needs to be integrated? When?
5. Describe which tests are for verification and which tests are for validation.

## User Acceptance Test (UAT) Scenarios

### Data Scientist Uploads Manifest

The user enters the upload page, selects file from local directory and uploads manifest to the server.

### Data Scientists Review Manifest

The user selects a manifest and review it. The required manifest is sent in response to query.

### **Data Scientists Search for Manifest**

The user searches a manifest by keywords, the results can also be limited by time and author filters.

### **Data Scientists Notify Changes to Other Users**

The data scientist sends a notification about changes to the author.

### **System Admin Bans an Illegal Data Scientist**

The data scientist account is transferred to the "banned" group.

A notification is sent to the user as well as the reason for the ban.

### **System Admin Deletes an Illegal Manifest**

The manifest is deleted (moved to "trash" group), and a notification and reason are sent to its author.

### **User Experience**

The stability of the system is acceptable.

The reaction time is short.

## **Unit Test Scenarios**

### **Login Function**

If the user exists, the action continues.

If the user does not exist, further action is denied and the error message is shown.

### **Search for Manifest Function**

The system finds the record from database by matching the keywords.

The system returns an error message if no record is found.

### **Upload Function**

The file is stored in the database.

### **Download Function**

The required file is pulled from database and presented to the user.

## Edit/Delete Manifest Function

The manifest is changed and database is updated.

## Failure Case

Error message is displayed. The expected value and return value are compared.

## Regression Testing

- Regression testing is a type of software testing that verifies that software previously developed and tested still performs correctly even after it was changed or interfaced with other software.
- In this system, a set of unit tests are prepared to cover all the functions of the software. The tests are run after every update or bug fixing.

## Integration Testing

- Integration testing is the phase in software testing in which individual software modules are combined and tested as a group.
- We will perform integration testing after unit tests. It tests how the website responds to a user behavior.

## User Uploads Manifest

- The user can sign in, and choose the file to be uploaded, after the system check the file, the file is uploaded and the success message shows.

## User Reviews Manifest

- The manifest is searched on the server. If it is still valid, it is extracted from database and shows to the user.

## System Admin Bans an Illegal Data Scientist

- The data scientist account is transferred to the "banned" group. A notification is sent to the user as well as the reason for the ban.

## System Admin Deletes an Illegal Manifest

- The manifest is deleted (moved to "trash" group), and a notification and reason are sent to its author.

## Failure Case

- The system returns an error page when the required page is not found, or the server is down.

## Testing Edge Cases

### Validating user inputs

When user searches for a manifest, or filling the information for a manifest

- the input cannot be empty
- the input cannot be too long or too large

Unit tests are used for these special cases

## Responsive layouts

The design layout should account for the device to be used. For this case, it is not a problem since it can be assumed medium or large device is used for data scientists. However, the application for mobile devices can be considered.

## Accessibility

Accessibility will be considered for special cases.

- User with color-blind problem may have difficulty in viewing the image. When there is an important message, the app should notify with text instead of color information.
- User with slow internet connections may wait a long time for loading the page. The design should try to minimize unnecessary effort on transmission.

This is tested by asking other people to use the web app.

## Edge cases to ignore

Not all potential cases are required to be accounted for. For example, most of people have java embed in their web browser, and most people do not use an out-of-date browser such as IE5. If problems are caused by these reasons, it can simply show an alert for the user to notify the problem. Since there are too many edge cases with low possibilities.

## Testing Summary

- The unit testing, integration testing and regression testing are for verification, the user acceptance test is for validation.

# Sprint 2

## Meeting Notes

*Time: 11/8/2016*

*Location: Lafferre Hall*

### Team Members:

Holt Skinner, Justin Hofer, Ali Raza, Bo Zhang, Bradley Rogers, Pramod Pudotha.



## Decisions Made

Complete web homepage that has links to browse, upload, edit and login.

Apply Python to replace PHP because of compatibility issues.

Update MongoDB Pseudo-ERD.

Allow Data Scientists to login with Google account.

## Task Lists

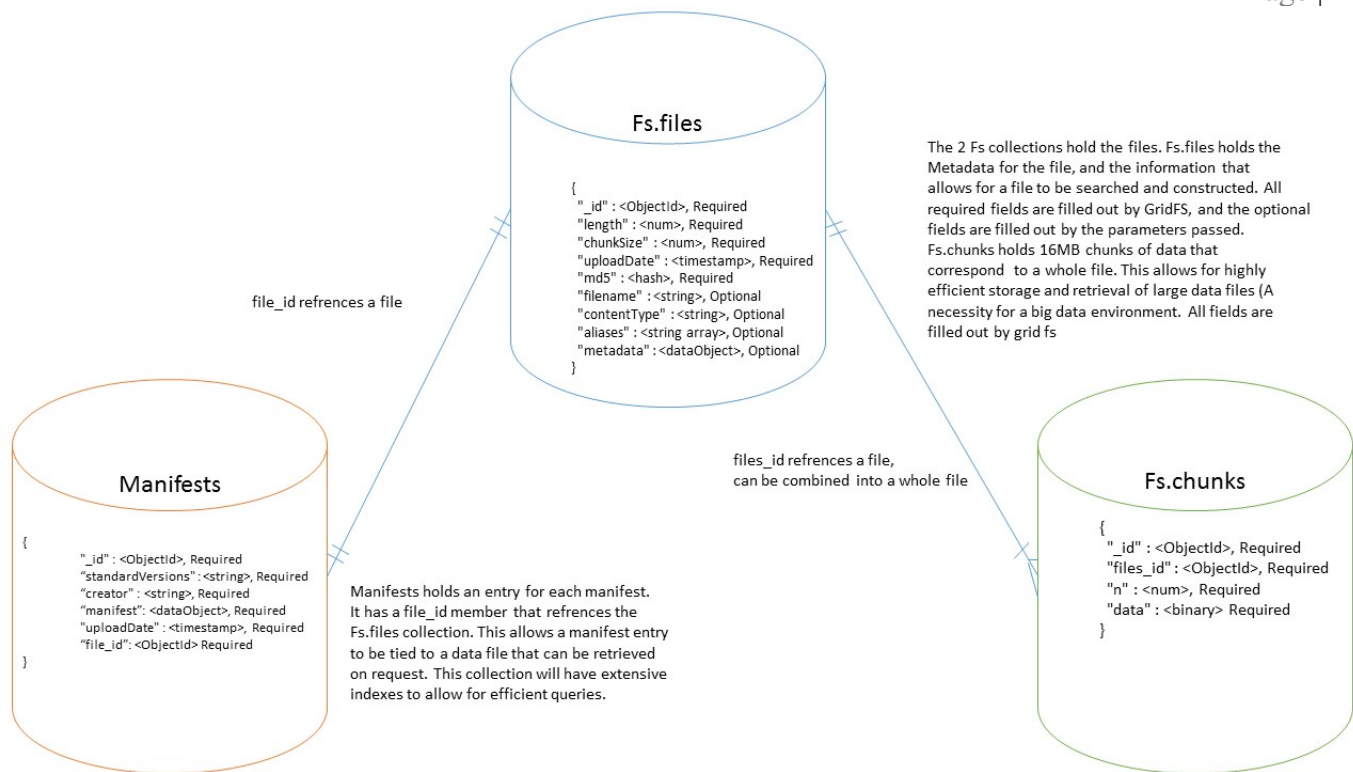
1. Complete basic functions for database interface with Python, and create unit tests for these functions. (Justin)
2. Redesign and update ERD. (Justin)
3. Configure deployment environment for Python-MongoDB. (Ali)
4. Finish homepage user interface. (Holt, Bradley)
5. Work on additional pages, and create login button linked to Google API. (Bradley, Holt)
6. Update testings. (Bo)

## Remaining Tasks

1. Search bar in homepage is not functional.
2. Edit, delete, upload functions are not implemented.
3. Database and front-end are not linked.

## Database

The database interface, written in python, allows for easy use of insert, update, search, and delete functionality for the database. Error checking is implemented to ensure that database integrity is maintained. Document Validation (Encoded by the creation statements) will check that all manifests are up to standard, and as such, do not need to be checked by the dml (Although they should be checked in the business layer). The Unit tests validate these functions, and will ensure that they are valid throughout creation of the system. The database also no longer has a user collection, as it is no longer required for login. As such, the user collection has been removed (Including from the ERD). The ERD also now uses Crows Feet notation to represent relationships, and required fields are marked.



## UPDATED MONGODB PESEUDO-ERD

### Python code for inserts, updates, searches, and deletes.

```
from pymongo import MongoClient #Mongodb functionality

#initialize to the collections that we want
client = MongoClient()
db = client.test

m_col = db.Manifests #the collection of manifests

def search_manifest(lookup):
    ''' finds all manifests that match pattern provided and returns the cursor of
    results
    Returned value is a cursor that can be iterated through with a for loop '''
    return m_col.find(lookup)

def insert_manifest(manifest):
    ''' Insert given manifest if it exists. Returns True on success,
    or False on error (Failed document validation or no manifest was passed in)
    ...

    if(manifest): #basic error checking
        post_id = m_col.insert_one(manifest)
        ''' we have the object id for the new manifest, and we could return it if
    we like
```

```

We could also return a tuple containing the Boolean value and the object
id

    This way we could later do a lookup on the manifest, which would give us
    access to its metadata. As we do not have any important metadata at the
    moment, we will just give a simple error check '''
    if(post_id):
        return True
    return False

def remove_manifest(oid):
    ''' Delete manifest specified by internal object id. Access this object id
    within the manifests metadata. returns True on succesful delete,
    or False if unable to delete (No matching oid, no oid provided) '''
    if(oid):
        ''' Delete based on oid. Only can fail if the oid is invalid
        In which case, wow, we are corrupting our own metadata somewhere
        in the business logic or the view. At least we would know that we have
        an error '''
        result = m_col.delete_one({'_id': oid})
        if(result.deleted_count == 1):
            return True

    return False

def update_manifest(oid, manifest):
    ''' Updates the manifest specified by the given internal object id.
    Changed to match the new manifest provided. Returns True
    if the document was succesfully updated, and returns False if it
    failed. '''
    if(oid and manifest):
        #We actually want to remove the old manifest, and replace it with a new
one
        old_doc = m_col.find_one_and_replace({"_id": oid}, manifest)
        #the old manifest was returned, so we can store an archive of manifests
down
        #the road. For the monent, we will just check that something was there
before
        if(old_doc[0]):
            return True
    return False

```

## Unit Tests

```

from dml import insert_manifest, remove_manifest, update_manifest, search_mani-
fest

to_insert = {
    "manifests": {

```

```

"manifest": {
  "standardVersions": "ocdxManifest schema v.1",
  "id": "https://datahub.io/dataset/iDas",
  "creator": "Ali Raza",
  "dateCreated": "2016 - 10 - 27",
  "comment": "First test manifest",
  "researchObject": {
    "title": "iDAS Manifest",
    "abstract": "Data collected at the Interdisciplinary Data Analytics and Search lab at the University of Missouri by Computer Science researchers and Data Scientists.",
    "dates": {
      "date": {
        "date": "2005 - 04 - 27",
        "label": "start"
      }
    },
    "privacyEthics": {
      "oversight": {
        "label": "No assertion"
      }
    },
    "informedConsent": "No assertion",
    "anonymizedData": {
      "label": "No assertion"
    },
    "privacyConsiderations": "No assertion"
  },
  "provenance": {
    "narrative": "The Interdisciplinary Data Analytics and Search (iDAS) lab is one of the many research labs operating out of The University of Missouri, Columbia. As the name implies, iDAS combines researcher across departments to achieve solutions to problems in academia. Founded in 2005 by Dr. Chi-Ren Shyu, iDAS researchers are primarily Computer Scientist, but the lab also works with Medical Doctors, Biologist, and Statisticians."
  },
  "publications": {
    "publication": "No assertion"
  },
  "locations": {
    "location": {
      "url": "",
      "comment": ""
    }
  },
  "files": {
    "file": {

```

```

        "name": "iDAS - data.csv"
    },
    "format": ".csv",
    "abstract": "Metadata for 5000 records collected",
    "size": "No assertion",
    "url": "No assertion",
    "checksum": "No assertion"
},
"permissions": "No assertion"
},
"dates": {
    "date": {
        "date": "2014 - 02 - 15"
    },
    "label": "Created"
},
"creators": {
    "creator": {
        "name": "Chi-Ren Shyu",
        "role": {
            "label": "Other"
        }
    },
    "type": {
        "label": "No assertion"
    },
    "contact": "cshyu@wikimedia.org"
}
}

```

```

to_replace = {
    "manifests": {
        "manifest": {
            "standardVersions": "ocdxManifest schema v.1",
            "id": "https://datahub.io/dataset/sociallyCompute",
            "creator": "Sean Goggins",
            "dateCreated": "2016 - 08 - 13",
            "comment": "Second test manifest",
            "researchObject": {
                "title": "Socially Compute Manifest",
                "abstract": "Data mined from socail networks for the purpose of
consumer trend analytics.",
                "dates": {
                    "date": {
                        "date": "1992 - 03 - 11",
                        "label": "start"
                    }
                }
            }
        }
    }
}

```

```

    }
  },
  "privacyEthics": {
    "oversight": {
      "label": "No assertion"
    }
  },
  "informedConsent": "no assertion",
  "anonymizedData": {
    "label": "No assertion"
  },
  "privacyConsiderations": "No assertion"
},
"provenance": {
  "narrative": "Socially Compute is an ongoing project aiming to analyze trends of everyday people to make meaningful connections."
},
"publications": {
  "publication": "No assertion"
},
"locations": {
  "location": {
    "url": "",
    "comment": ""
  }
},
"files": {
  "file": {
    "name": "Socially Compute - sc.csv"
  },
  "format": ".csv",
  "abstract": "Metadata for 15000 records collected over two decades",
  "size": "No assertion",
  "url": "No assertion",
  "checksum": "No assertion"
},
"permissions": "No assertion"
},
"dates": {
  "date": {
    "date": "2016 - 10 - 28"
  },
  "label": "Created"
},
"creators": {
  "creator": {
    "name": "Sean Goggins",

```



```

        "role": {
            "label": "Other"
        }
    },
    "type": {
        "label": "No assertion"
    },
    "contact": "sg@wikimedia.org"
}
}

#Insert the first test manifest
test = insert_manifest(to_insert)
if(not test):
    print("Bad insert")
else:
    print("Good insert")

#Ensure that the manifest was inserted properly
found = search_manifest({})[0]
if(found['creators']['contact'] == to_insert['creators']['contact']):
    print("Match")
else:
    print("No Match")

#Update to Second manifest
test = update_manifest(found['_id'], to_replace)
if(not test):
    print("Bad update")
else:
    print("Good update")

#Ensure good manifest update
found = search_manifest({})[0]
if(found['creators']['contact'] == to_insert['creators']['contact']):
    print("Did not replace")
elif(found['creators']['contact'] == to_replace['creators']['contact']):
    print("Good Replace")
else:
    printf("replace corruption")

#remove the manifest
test = remove_manifest(found['_id'])
if(not test):
    print("Bad remove")
else:
    print("Good remove")

```

## User Interface

- The Homepage UI is functional.
- A Data Scientist is able to search for manifests from the homepage and view details for the desired manifest through a link for each search result.
- Currently, the search results are dummy datasets until the database can be linked with the View.
- The GUI is developed using the Materialize framework for styling and JavaScript/jQuery for the front-end business logic.
- The navigation bar at the top of the page provides easy access to all main functions of the application.
- The "Login" button opens a Modal box that allows a Data Scientist to log in with a Google Account.

[Source Code For Homepage](#)

## User Accounts

- Implemented via Google's OAuth 2.0 API.
- A Data Scientist can log in with a Google account.
- A Google login page appears to the user, the user logs in with Google credentials.
- The Google API returns a unique user\_id token, which is stored in the Database.
- Due to limited use cases, and limitations of the Google API, there is not a dedicated administrative user. Instead, all logged in data scientists have full read/write access.
- [Documentation from Google](#)
- Source Code

[Import of Google API](#)

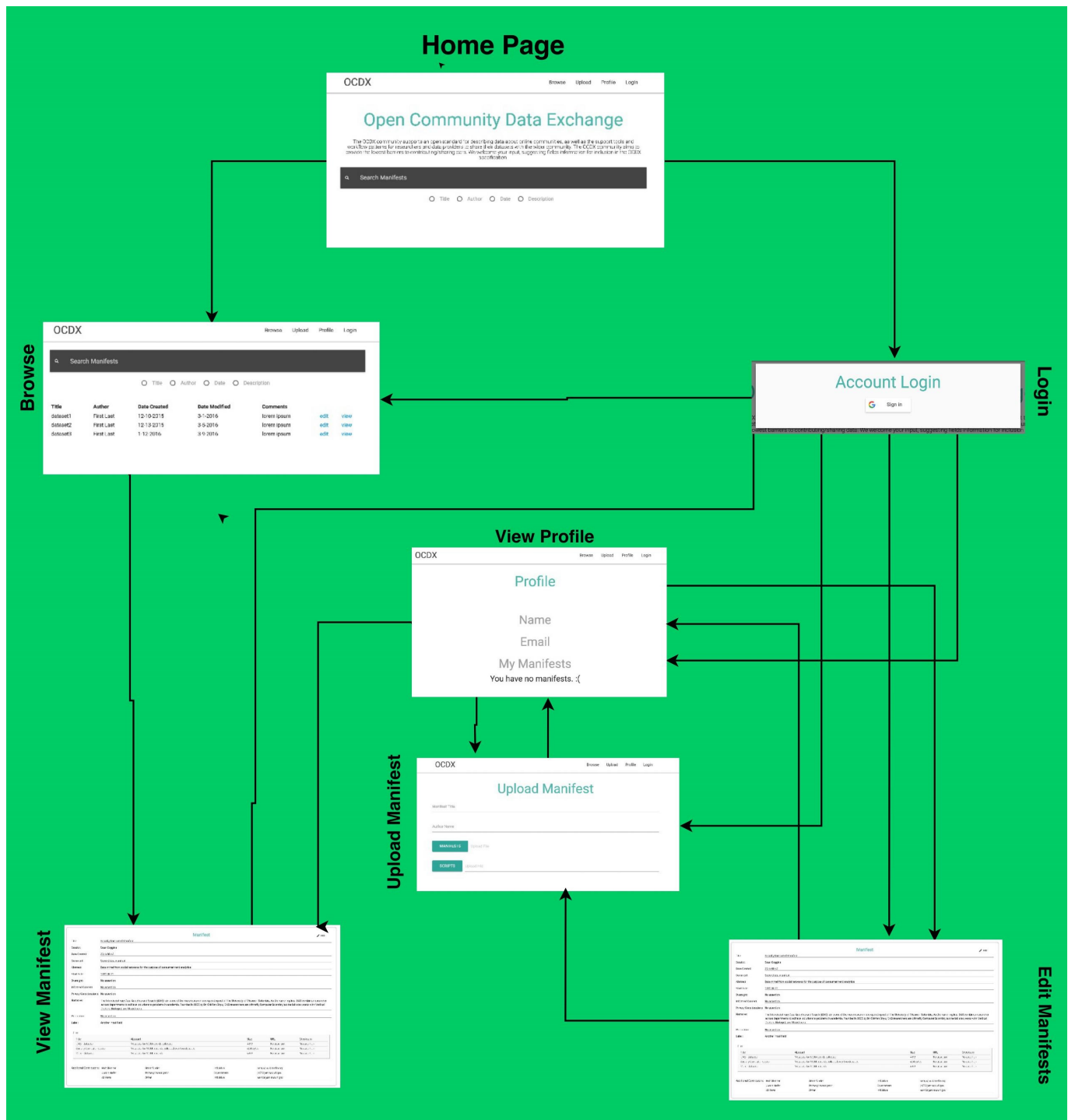
[User Profile Page](#)

## Testing Updates

- Update testing designs in Sprint 1.
- Unit tests for the Database insert, update, search, and delete functions are implemented as a python script. (See Sprint 2, Database)

# Sprint 3

## User Interface Temporal Logic



### Complete

- Home Page/Search
- Browse

- View Profile
- Account Login
- Upload Manifest

### Work in Progress

- View/Edit Manifest

### Omitted

- Data Visualizations

## Testing Updates

### Files (Justin Hofer)

- Need to ensure that a user can upload, remove, and download files. User also needs to be able to change how a file is tied to a manifest (i.e. change the file, upload a new one, remove a file).
- To test this, an automated script will be run which will first perform unit tests that will upload a test file, verify it is there, retrieve it, validate the file's data, then perform a delete and ensure its removal. Then the test will attempt to tie a test file to a manifest, and remove the tie to the manifest. Finally, manual integration tests will be performed to ensure lossless transfer.

## User Documentation

- [User Documentation Version 1](#)

## Meeting Notes

### 11-15-2016

### Members Present

Justin, Brad, Holt, Ali, Bo, Pramod

### Task List

Ali:

- Convert UI logic to python applications using flask
- Begin User documentation
- Deployment documentation (Sprint 4)

Justin:

-GridFS for file system implementation

Bradley/Holt:

-Continue front end design

Pramod:

-Update documentation

Holt:

-Documentation correction/formatting

## Decisions

Convert PHP files to python

**11-16-16**

## Members Present

Brad, Bo, Ali

## Work Done

- Ali began user documentation for sprint
- Bo is working on temporal logic for sprint 3
- Brad is working on the viewing/editing manifests

## Decisions

- Email Dr. Goggins for clarification on Manifest format
  - Email has been sent.
  - Currently waiting for response.
- Test whether PHP 5.4 works w/ Mongo
  - Ali will do this on a separate server.
  - If not successful, we will convert PHP to Python/Flask

**11-17-16**

## Members Present

Holt, Brad, Justin, Ali, Bo

## Decisions

---

Allow users to insert manifest via JSON

Give access to manifest specifications

Make individual markdown of task list

Implement Ajax calls to Python scripts

Continue Grid FS implementation

Create Temporal Logic (Ali)

## Actions

Converted PHP server-side include statements to jQuery or Python

## Sprint 4

### Meeting Notes

11-27-16

*Skinner*

### Members Present

Brad, Justin, Holt

### Decisions Made

Put out fires from Thanksgiving Break.

Clarified strategies for front-end to back-end communication.

Discussed changes for Sprint 4 Documentation.

Communicated strategies for edit functionality.

Discussed presentation strategies and organization.

Justin fixed giant merge conflict caused by Holt.

### Task List

Justin

- Finish Writing Search Tests
- Update Documentation for GridFS



Brad

- Fix Edit and Browse Pages to work with connection to python scripts.
- Linking front end and backend for edit and browse using Flask REST API

Holt

- Change .php files to .html files and manually merge Bo's changes from Sprint 3. (GitHub wasn't working for some of the changes)
- Refactor all Requirements Analysis and Sprint Documentation for re-grading
- Prepare Presentation Organization

## Database Changes

*Author: Justin Hofer*

- Search is implemented through regular expression. All fields relevant to the current search are examined, i.e. a dates search examines both the external date fields and the internal date fields within the manifest itself. Differences between spacing for dates are also accounted for. Manifests are also indexed for efficient searching.
- GridFS is implemented to allow file storage. We have no max file size, or illegal file types. This allows for extreme forward compatibility, as new file formats may appear in the future, and we will support them. As this is a big data solution, we also need to be able to handle large files, and lacking a maximum file size ensures this will not be a problem.
- Unit tests and edge cases are implemented for both of these. Tests for various types of user input (and lack thereof) allow us to ensure that the system will be able to perform as expected.
- Link to tests: <https://github.com/holtwashere/CS4320-FinalProject/blob/master/Source/db/dml.py>
- Link to implementation: [https://github.com/holtwashere/CS4320-FinalProject/blob/master/Source/db/dml\\_unit\\_tests.py](https://github.com/holtwashere/CS4320-FinalProject/blob/master/Source/db/dml_unit_tests.py)

## Deployment Instructions

*Rogers/Hofer/Zhang*

### User Documentation

- The user documentation gives a quick start for new users as well as detailed tutorials about how to use the web app.
- [User Documentation Version 2 Link](#)

### Automated Deployment

- Automated deployment is enabled by running a bash script as root in an EC2 instance.

- Follow [README.md](#)
- The deployment can be finished in one minute.

## Unfinished Business

### *Skinner*

- Change jQuery include statements to a server-side Python script.
  - The jQuery .load() causes a stutter when the pages first load.
- Clean up connection between User Interface and Python query scripts.

## Past Sprint 4

### Wrap-Up Meeting

*12-9-16 @ 2PM in W2003*

#### Members Present

Holt, Brad, Bo, Ali, Justin

#### Decisions Made/Tasks Completed:

##### Clean up Repository

Many deprecated files were deleted, hence the missing links to completed work in previous Sprint sections.

All final code and logic can be found in the GitHub Repository.

##### General Code and Documentation Cleanup

Ran Atom Beautifier to format all code following standard practices.

Finalized Search Functionality and linked with UI.

Finalized ability to edit manifest front-end validation logic.

## Front-end Updates

Allowed editing manifest based upon security and privacy restrictions.

Multiple bugs were fixed between Sprint 4 and the Final Submission pertaining to searching editing manifests.

[Link to edit.html](#)

## Business Logic Updates

404 and 500 pages were added to correspond to common server errors.

[Link to 404 markup](#)

[Link to 500 markup](#)

[Business Logic to route errors](#)

## Database Updates

The database now supports searches with any amount of improper spacing for added ease of use.

Search dates with any number of spaces

12-1-16

12 - 1 - 16

12    -1-16

etc.

Search with capital and space oddities

aliraza

Ali Raza

aL i Ra zA

a l i r a z a

## Deployment Updates

To facilitate continuous development and deployment, the team has created two EC2 instances, one for Development and experimentation, and one Product environment for active users. This paradigm began before the in-class demo to allow for continuous development on the Flask server without disturbing the live version that was to be used for demonstration purposes. The product environment can only be updated by the system administrator, Ali, to prevent merge conflicts and bugs from getting loaded into the active version.

## Hindsight is 20/20

*What the group should have done from the beginning*

*Skinner*

- Make Python the language for all business logic, rather than PHP.
- Thoroughly document all decisions made and steps taken.
- More thorough meeting documentation.

- Communicate and listen before making commits and pushing to git repo.
- Link Github Repo to EC2 instance to avoid fragmentation.
- Use a Google Doc for Sprint Documentation rather than Markdown or Microsoft Word.

## Change Log

- Version 1.0: Pre-Implementation Design *10-16-2016*
- Version 1.1: Sprint 1 *10-31-2016*
- Version 1.2: Sprint 2 *11-11-2016*
  - Made Fixes/Changes to Initial Requirements Analysis/Design and Sprint 1 Documentation
- Version 1.3: Sprint 3 *11-18-16*
  - Made updates to Sprint 2 Documentation for Regrade.
- Version 1.4: Sprint 4 *11-28-16*
  - Added Sprint 4 Section.
  - Major refactoring of full documentation formatting for regrade on all sections.
  - Added links to completed work in every section.
- Version 2.0: Final Submission *12-16-16*
  - Clean up Documentation for regrades.
  - Finished Application.

## Glossary

- GridFS: MongoDB's filesystem that allows for large files to be stored and retrieved efficiently.
- Index: A structure that allows for fast lookups within a database.
- MongoDB: An open source JSON based Big database
- PyMongo: Python api allowing access to MongoDB
- Python: Server side scripting language
- Regex (Aka. Regular Expression): A pattern matching tool allowing for text searches on specific fields in MongoDB.
- Flask: A Python Framework to allow RESTful API capabilities.