System Analysis and Design – Term Project

Term Project

|  |
| --- |
| detail of persons hands with scissors, markers, workingDATA CATALOG |

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | **Submitted by:**  Faheem Vellore  Rehnuma Seheli  Sadiksha Upadhyay  Zhu Qing Nim |  |

TABLE OF CONTENTS

[data catalog – State of arkansas 3](#_Toc531007186)

[Summary of Project 3](#_Toc531007187)

[Description of Organizations and Problems 3](#_Toc531007188)

[Project Initial Study 5](#_Toc531007189)

[1) Objectives 7](#_Toc531007190)

[2) Feasibility Study 7](#_Toc531007191)

[a. Technical 7](#_Toc531007192)

[b. Economical 8](#_Toc531007193)

[c. Operational 9](#_Toc531007194)

[PROJECT MANAGEMENT PLAN 10](#_Toc531007195)

[Work Plan 10](#_Toc531007196)

[Size, Duration, Cost Estimation: 12](#_Toc531007197)

[USER REQUIREMENT ANALYSIS 17](#_Toc531007198)

[SYSTEM ANALYSIS AND DESIGN 19](#_Toc531007199)

[Activity Diagram 19](#_Toc531007200)

[Use Case Diagram 20](#_Toc531007201)

[Class Diagram 24](#_Toc531007202)

[Sequence diagram 25](#_Toc531007203)

[Object Persistence Design 27](#_Toc531007204)

[Conclusion 28](#_Toc531007205)

[Appendix 29](#_Toc531007206)

# data catalog – State of arkansas

## Summary of Project

Arkansas State Data Catalog powers governor, agencies, the general assembly and other potential investors to make informed decisions by the data asset shared by the agencies. Data Catalog is a *storage management and distribution solution* of data to be mainly used by public institutions to share their data with the general public, interstate agencies and other potential investors to leverage data assets to make better data driven decisions.

## Description of Organizations and Problems

State of Arkansas has different agencies working on separate platforms, even though agencies have the same purpose of serving the citizens. Some of the problems of working on a separate platforms are:

1. Getting correct data takes time and effort. A common citizen has less or no knowledge of where the data is stored and what data is available to them.
2. There is no operational access control over the data through governance.
3. There are many redundant data across different agencies.
4. Information about data is not transparent. All the governing policies (HIPAA, FERPA, GDPR etc.), memorandum of understandings is not available openly which may lead to misuse of data intentionally or unintentionally.

In 2017, the State of Arkansas produced bill Act 912 which was established to create a panel on Data Transparency. Data Transparency facilitates standardized access to data within and between or outside systems or state agencies to establish a Data Catalog. Below are the listed agencies for data transparency (Data Catalog) as stated in the Act 912.

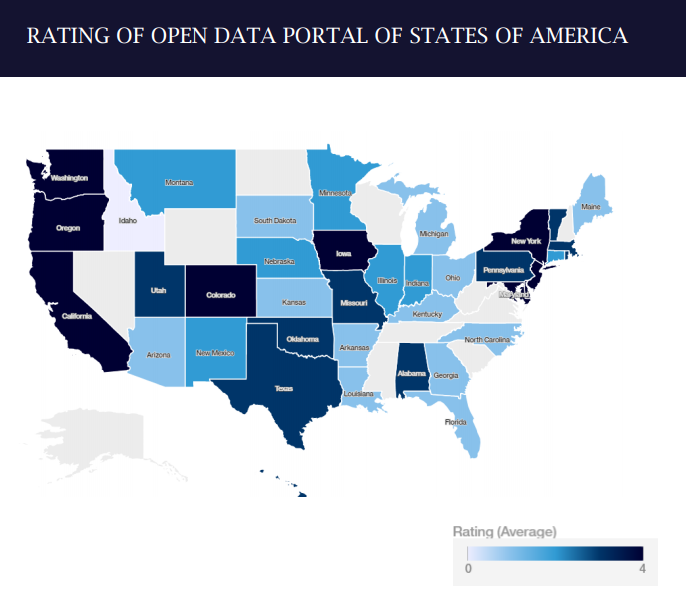
1. Administration of Courts
2. Arkansas Community Correction (ACC)
3. Arkansas Department of Correction (ADC)
4. Arkansas Crime Information Center (ACIC)
5. Arkansas Police Department
6. The Department of Finance and Administration (DFA)
7. The Department of Health (ADH)
8. The Department of Education (ADE)
9. The Department of Higher Education (ADHE)
10. Arkansas Department of Career Education (ACE)
11. The Department of Human Services (DHS)
12. The Department of Workforce Services (DWS)
13. Arkansas Department of Labor (ADL)
14. The Department of Information System (DIS)

Government data, especially open data is a tremendous resource that is yet largely untapped. Many individuals and organizations collect a broad range of different types of data in order to perform their task. Government is particularly significant in this respect, because of both the quantity and centrality of the data it collects. Also, most of that government data is public data by law, and therefore could be made open and available for public to use.

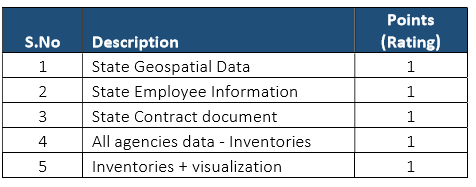
Data Catalog is also of value to government itself. For example, it can increase government efficiency by reducing redundancy of work. Once all the data is collected in one place (Data Catalog), any other agency could use the data without having to create it. This will not only save money but it can also reduce time and therefore increase efficiency.

## Project Initial Study

United States has passed an executive order to make government data open and Machine-readable. Because of this reason, many state took initiative to publish the data such as GIS data, state employee salary, contracts and other data set, which are stored and used in daily basis for different agencies. The map below shows an attempt to understand the level of work performed by each state regarding data transparency.



The table below depicts different state open data portal with an attempt rating on quality of data sharing. Note that the quality of data sharing is based on below parameters. These ratings is only for understanding of level of work done by each state to be used by Arkansas Department of Information System.



**Analysis/Design**:

The system analysis/ design phase determines the logical function of the system, such as Activity Diagram, Use Case Diagram, Use Case Description, Class Diagram, Object Persistence Diagram, Sequence Diagram etc. Activity Diagram is created to ensure the designers and users are clear on the capabilities and functionality of the system. Use Case Diagram and Use Case Description further define system detail functionality. Class Modeling then gives static aspects of the system, while Sequence Diagram describes the dynamic and interactions of the system’s components. Finally, a Relational Model (Object Persistence Diagram) is created in order to identify the logical representation of the relationships within the data.

**Implementation/ Testing/Deployment**:

Upon formation of the Data Catalog system, all 14 government agencies will upload their data in the Data Catalog and provide to the users. Once all the implementation of the system is done, we will have a requirement testing by walkthrough process which involves the team leader, programmers, admins, agencies, and quality controller. The deployment of Data Catalog system for the State of Arkansas needs approval from all the state agencies’ directors and the final approval from the governor.

## Objectives

* To have a common platform for agencies of State of Arkansas to share Data for the citizens.
* To increase transparency in the functioning of Arkansas Government and open avenues for many more innovative users of data.
* To avoid data redundancy among all the agencies.
* To have control over the restricted data for managing security and tracking.
* To prioritize the importance of most frequently used data available to public.
* To use this knowledge to inform decision making among executives, policymakers and users.

## Feasibility Study

## Technical

**Familiarity with Application (high)**

* The team leader and programmers assigned to this program have extensive experience with Java and SQL.
* The objectives of this application are well within their capabilities.
* The members involved in this project have enough knowledge regarding the program, they have built the Data Catalog system from the State of California.

**Familiarity with Technology (high)**

* The government body of other states as well as the United States as a whole has developed an application of this scope before.
* Development tools and products for application development are available in marketplace.

**Project Size**

* We estimate that the project size is low in size in terms of programming.

## Economical

**Tangible Benefits**

* Data Catalog is expected to reduce redundancy of work within state agencies of Arkansas by approximately 1% ($14,363,631 annually taken from data.transparency.org).
* Data Catalog can also attract investors/industries to invest in state of Arkansas approximately $5,000,000 annually.
* Data Catalog can also lead to creation of mobile applications which in turn may generate revenue. Approximate expected revenue is $1,000,000 annually. (Similar to state of Ilion’s last year’s revenue.)

**Intangible Costs and Benefits**

* Increase quality of service to members.
* Improve system adaptation.
* Reduce outdated data from the State Departments.

## Operational

* The proposed system will increase the transparency of data within state agencies and users and help the members to have informed decisions. The data catalog is in line with Act 912 (as attached).
* We believe that this system will help make operations much smoother because all the information can be found in the same place.
* In the era of data driven decision making, it is really important for users to have updated data with associated information within the reach of citizens.

# PROJECT MANAGEMENT PLAN

## Work Plan

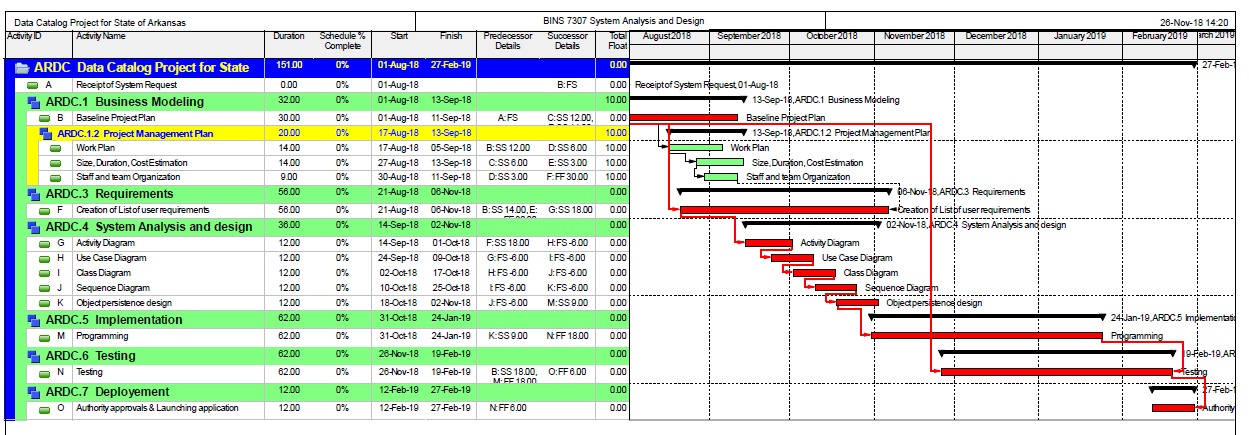


Figure 1 Gantt Chart

Activity Network Diagram

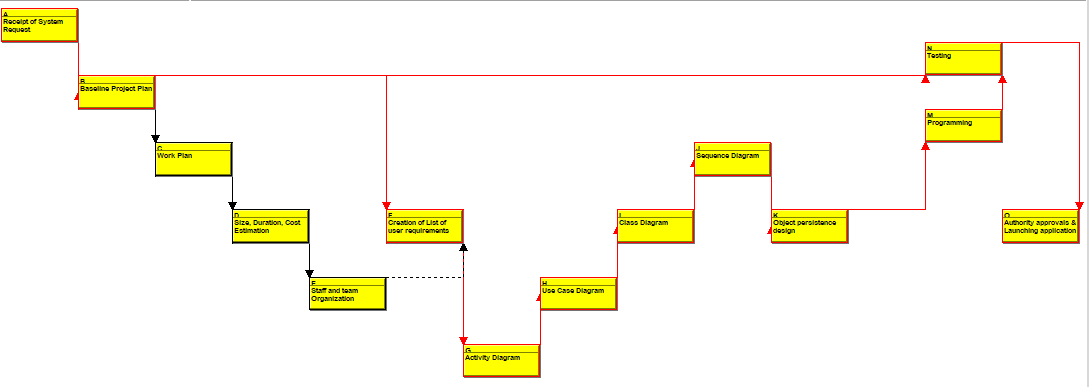
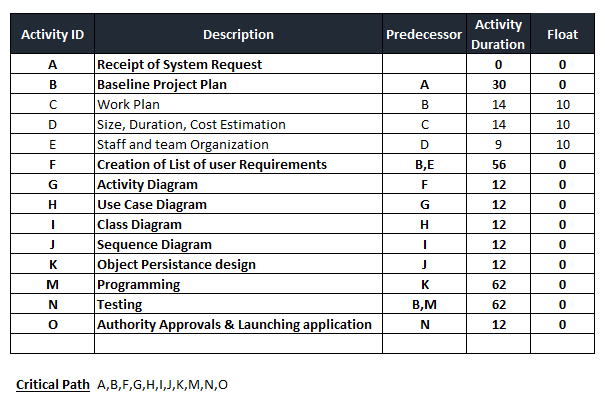


Figure 2 Activity Network Diagram



## Size, Duration, Cost Estimation:

This platform is a data catalog of 14 different agencies located in the state of Arkansas. To be efficient in gathering all the information data in the same location to avoid redundancy, increase transparency to improve important decision making.

The system will be used by any users (member or non-members) and 14 different agencies. For part of the system to be used by the users, there are 6 inputs screens needed (Create an Account, Login to Platform, Request Membership Type, Request Data Download, Search for Data, Filter Data Search); 4 outputs screens (Members Account Information, Data File Homepage, Data Searched Information, Dataset Downloaded Information); 3 main queries (Datasets Search, Dataset Downloads, Dataset Filtered), one file (Datasets Downloads); 2 program interfaces (Data Center, Members’ Profile Information). The complexities of all these elements were evaluated as low.

For the part of the system that used by the admin and agencies (14 state departments), 3 input ( Dataset Uploads, Forget Password Page, Answer Security Question), 3 outputs (Members Registration Confirmation, Result & Search Information, Print Reports), 6 queries (Check Database Storage, Valid Members Information, Valid Membership Type, View Members Information, Upload Datasets, Print Data Download List), 1 file (Upload Dataset), and no additional interface program. The complexities of all these elements were evaluated as low.

The project would use a combination of 60% function point would be Java and 40% would be SQL. The project is considered organic mode.

**Legends:**

**Users:** Any citizen who can view data.

**Non-Members:** Any user that view data but not able to download data.

**Members:** Any user that sign up for an account to download data includes authorized member and non-authorized member.

**Agencies:** The 14 agencies that holds an authority to upload datasets.

**Admins:** Administrative that hold the authority to create agencies account and authorize members’ type.

**Input: 9**

(Members): Members Account Login, New Member Register, Request Membership Type, Request Data Download, Data Search, Data Filter

(Agencies): Dataset Uploads, Forget Password Page, Answer Security Question

**Outputs: 7**

(Members): Member's’ Account Homepage, Data File Homepage, Data Searched Information, Data Downloaded Information

(Admins): Data Upload Notification, Result & Search Information, Print Reports

**Queries: 9**

(Members): Search Datasets, Download Dataset, Sort Dataset by Categories

(Admins): Check Database Storage, Valid Members Information, Valid Membership Type, View Members Information, Upload Datasets, Print Data Download List

**Files: 2**

(Members): Download Datasets

(Agencies): Upload Datasets

**Interfaces: 2** Data Storage, Account List

**Function Point Calculation:**

|  |  |  |
| --- | --- | --- |
| Description | **Complexity ( Low )** | **Total** |
| Inputs | 9 X 3 | 27 |
| Outputs | 7 X 4 | 28 |
| Queries | 9 X 3 | 27 |
| Files | 2  X 7 | 14 |
| Interfaces | 2  X 5 | 10 |

**Total Unadjusted Function Points (TUFP): 106**

**Project Complexity:**

(0: No Influence, 1: Incidental, 2: Moderate, 3: Average, 4: Significant, 5: Essential)

|  |  |
| --- | --- |
| Technical Factor | **Score** |
| Data Communication | 3 |
| Distributed Data Processing | 3 |
| Performance Criteria | 1 |
| Heavily Utilized Hardware | 0 |
| High Transaction Rates | 1 |
| Online Data Entry | 3 |
| End-user Efficiency | 1 |
| Online Updating | 3 |
| Complex Computations | 1 |
| Reusability | 1 |
| Ease of Installation | 1 |
| Ease of Operation | 1 |
| Portability | 1 |
| Maintainability | 1 |

**Project Complexity (PC):** 21

Calculating the adjusted project complexity (PCA)

**PCA** = 1.00 + (0.01 x PC)

= 1.21

**TAFP** = TUFP x PCA

= 106 x 1.21

=128.26

**60% Java**: 128.26 \* 60% \* 53 = 4,078.67

**40% SQL:** 128.26 \* 40 % \* 35 = 1,795.64

**Total Line of Code (LOC)** = 5,874.31

**COCOMO (COnstructive COst MOdel)**

(for Organic Software project below are the constants applied)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| SW Project | a | B | C | D |
| Organic | 2.4 | 1.05 | 2.5 | 0.38 |

**Man-month - MM**

MM = x EAF

= X 1

=15.39 MM

**Duration - D**

D =

=

= 7.06 Months

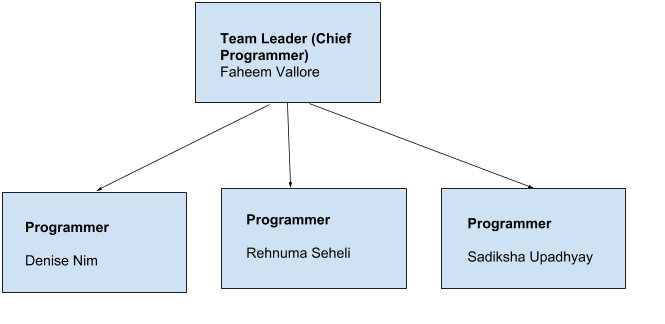
**Average Staffing** = 15.39/7.06 = **2.18**

Based on this project, we estimate the **total cost of $153,900** with the Man- Month of 15.39 and total duration of 7.06 months.

c) Staffing and Team Organization

**Chief Programmer Team Approach**

The team will have an average group of two staff throughout this project. Below is a chart of a team leader and three programmers. The three programmers will be supervised by the team leader.



Faheem Vellore

# USER REQUIREMENT ANALYSIS

1. **The system shall provide access to admin to manage the data catalog.**
   1. The system shall provide access for admin to create accounts for agencies.
   2. The system shall provide access for admin to get report of data downloaded.
   3. The system shall provide access for admin to authorize membership type.
   4. The system shall provide access for admin to retrieve password for agencies.
2. **The system shall provide access to agencies to maintain their accounts.**
   1. The system shall provide an option to the agencies to upload data from their accounts to a data storage.
   2. The system shall create a notification to the agencies after uploading data.
   3. The system shall provide access for agencies to get the list of data uploaded.
   4. The system shall be able to request a new password from admins.
   5. The system shall notify the admin if the database system is full.
3. **The system shall provide access to members to manage their accounts.**
   1. The system shall provide option to the members for login to their accounts.
   2. The system shall ask members about security questions to retrieve their accounts.
   3. The system shall provide data search option to all members from the data storage.
   4. The system shall provide data sort and filter option to all members from the data storage.
   5. The system shall provide an option to view the results to the members from the data storage.
   6. The system shall allow members to request authorized membership type.
4. **The system shall provide access to non-members to use the data storage.**
   1. The system shall provide data search option to all non-members from the data storage.
   2. The system shall provide data sort and filter option to all non-members from the data storage.
   3. The system shall provide an option to view the results to the non-members from the data storage.
   4. The system shall provide an option for non-members to create an account.
5. **The system shall maintain security in data downloading option for all members.**
   1. The system shall provide access to all authorized members to download the open and restricted data.
   2. The system shall provide access to all non-authorized members to download the open data only.
6. **The system shall detect the data trend.**
   1. The system shall generate a list of top downloaded data from the members.
   2. The system shall get the information about uploaded data (metadata).

# SYSTEM ANALYSIS AND DESIGN

## Activity Diagram

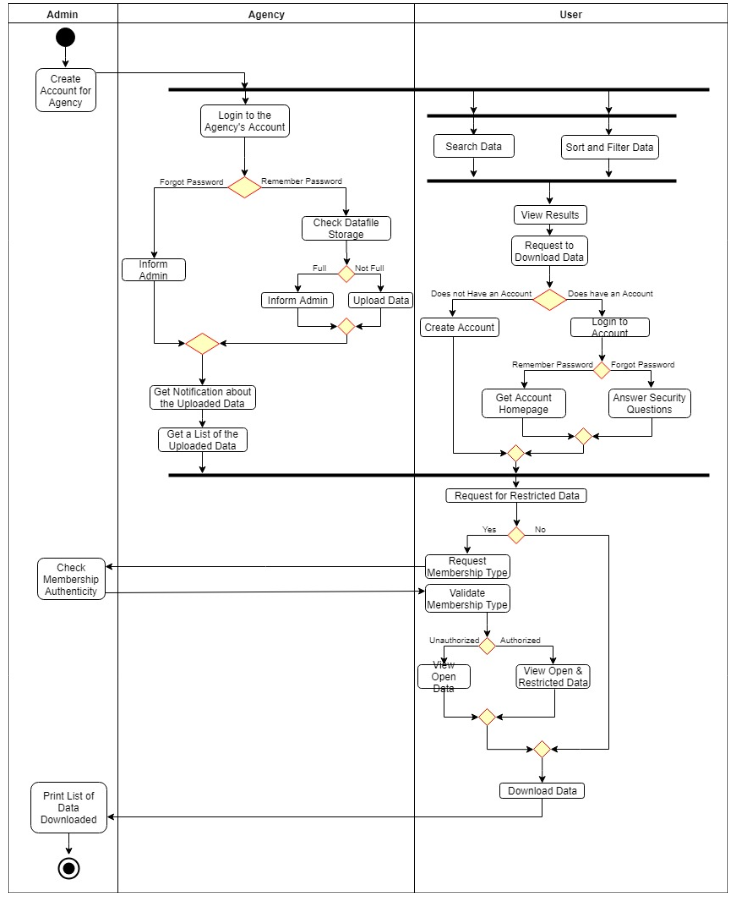


Figure 3 Activity Diagram

## Use Case Diagram

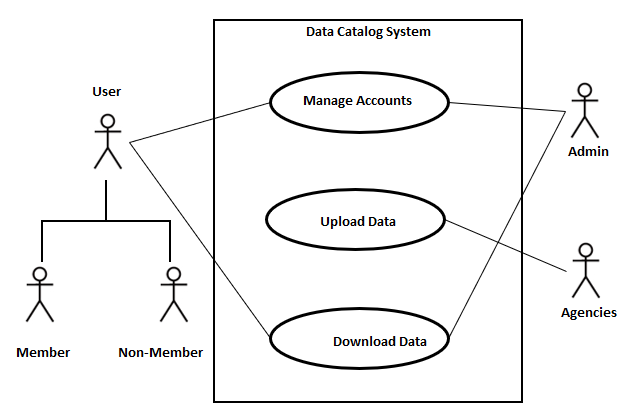


Figure 4 Use Case Diagram

**Legends:**

**Users:** Any citizen who can view data.

**Non-Members:** Any user that view data but not able to download data.

**Members:** Any user that sign up for an account to download data includes authorized member and non-authorized member.

**Agencies:** The 14 agencies that holds an authority to upload datasets.

**Admins:** Administrative that hold the authority to create agencies account and authorize members’ type.

|  |  |
| --- | --- |
| Use Case ID | 1 |
| Use Case Name | Manage Accounts |
| Actors | Admin, User |
| Brief Description | This use case describes the process of a creating and managing accounts for agency and users. |
| Basic Course | 1. Admin creates an account for selected individual in the agencies.      1. A user can login to their account. 2. The system maintains all the login information and associated login rights. 3. The system should display open and restricted data for authorized membership type. 4. The system should display open data only for non-authorized membership type. |
| Alternative Courses | 2.1 If the user is not a member, the system provides an option for creating a membership before downloading data.  2.2 If the users forgets password, the system asks for associated security questions before proceeding to new password.  2.3 If a user wants to view sensitive and closed information, a special request needs to be made to admin for associated login rights.  4. If the user is non-authorized, member can request admin for authority to view sensitive data. |
| Others |  |

`

|  |  |
| --- | --- |
| **Use Case ID** | 2 |
| Use Case Name | Upload Data |
| Actors | Agencies |
| Brief Description | This use case describes the process of uploading data by agency members. The data is stored with related information about the data called “metadata” in the data storage. |
| Basic Course | 1. Agencies login with credentials in the data catalog system. 2. The system validates the agency membership with credentials. 3. Agencies upload the data with associated metadata. 4. A notification is sent to the agency email address regarding the data upload. 5. Agencies will get a list of uploaded data from the system. |
| Alternative Courses | 1. If the agencies try to login with wrong credentials, system should ask for a new password by verifying security questions.  3. If the data storage is full, system creates a notification to inform admin. |
| Others |  |

|  |  |
| --- | --- |
| Use Case ID | 3 |
| Use Case Name | Download Data |
| Actors | Users, Admin |
| Brief Description | This use case describes the process of a downloading data by the members. Users can sort and filter, search, and view data. |
| Basic Course | 1. Users login with credentials in the data catalog system. 2. The system validates the membership with the credentials. 3. Users can search about data in the data catalog system. 4. Users can sort and filter the search results. 5. Users can view open data. 6. Users can download open data. 7. Admin can view and print the history report of viewing and downloading data assets. |
| Alternative Courses | 2. If the users forget password, the system asks for associated security questions before proceeding to new password.  6.  If user is authorized, the system allows them to download restricted and open data. |
| Others |  |

## Class Diagram

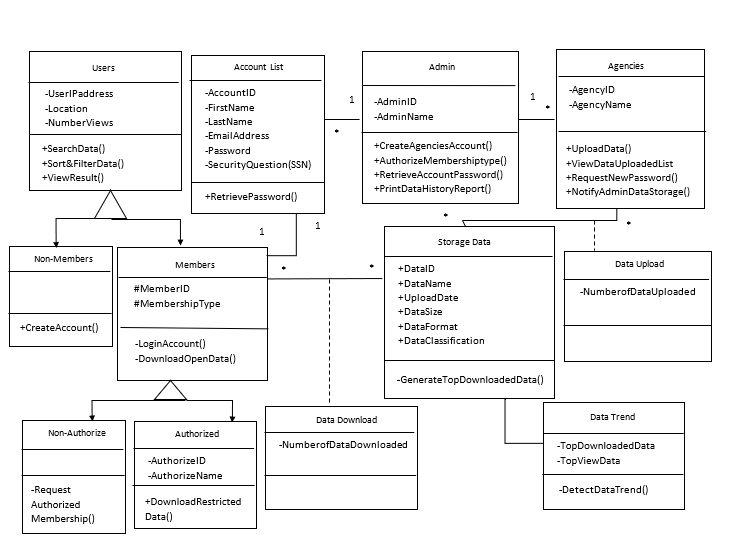


Figure 5 Class Diagram

## 

## Sequence diagram

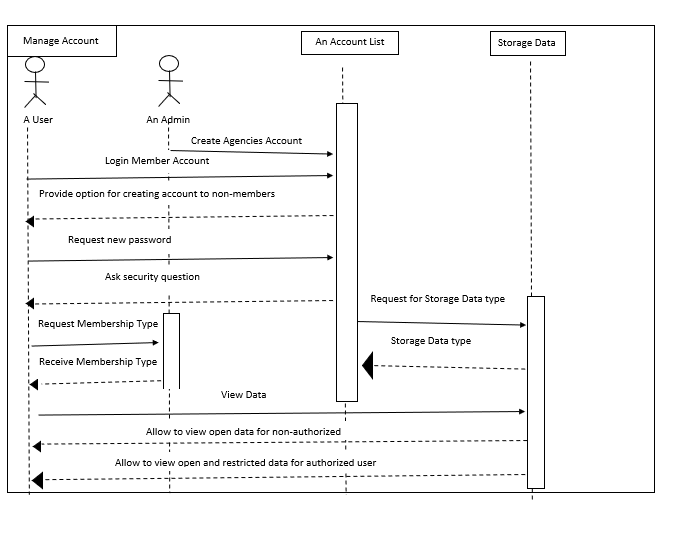


Figure 6 Sequence Diagram: Manage Account

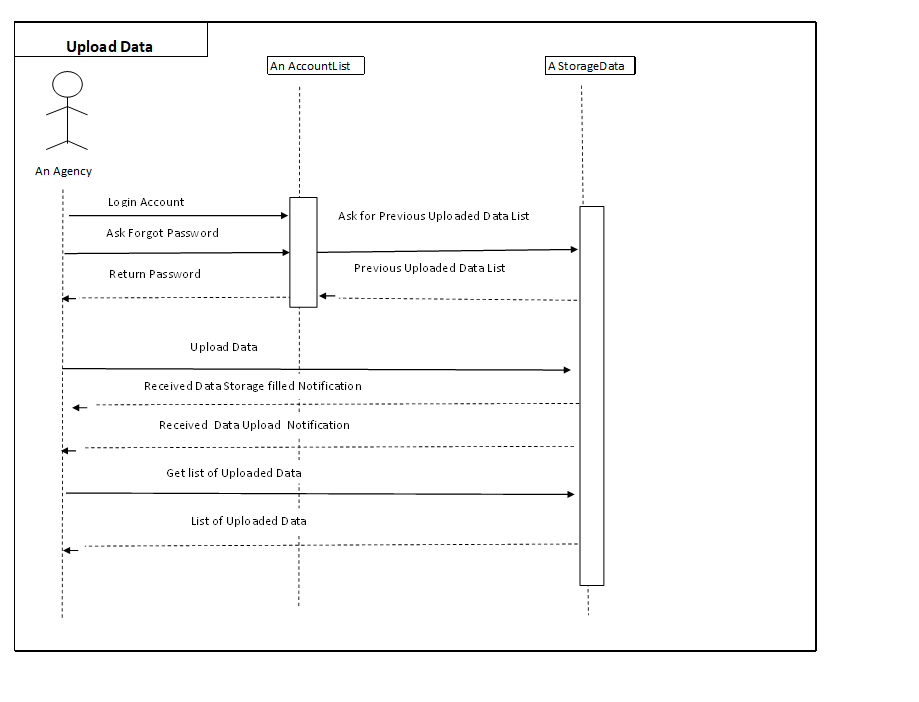


Figure 7 Sequence Diagram: Upload Data

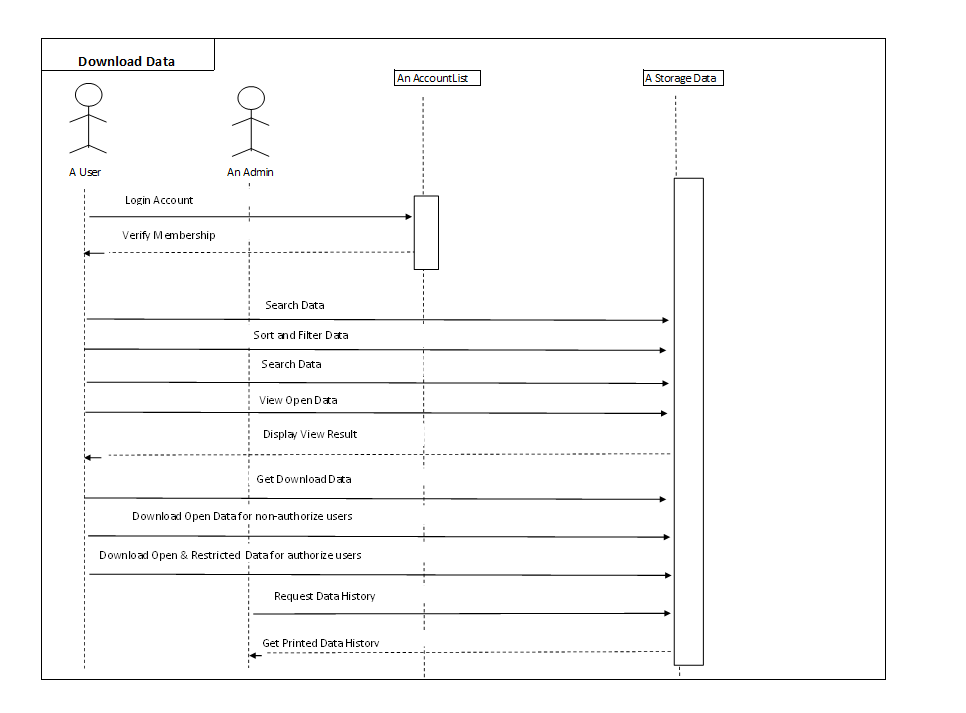


Figure 8 Sequence Diagram: Download Data

## Object Persistence Design

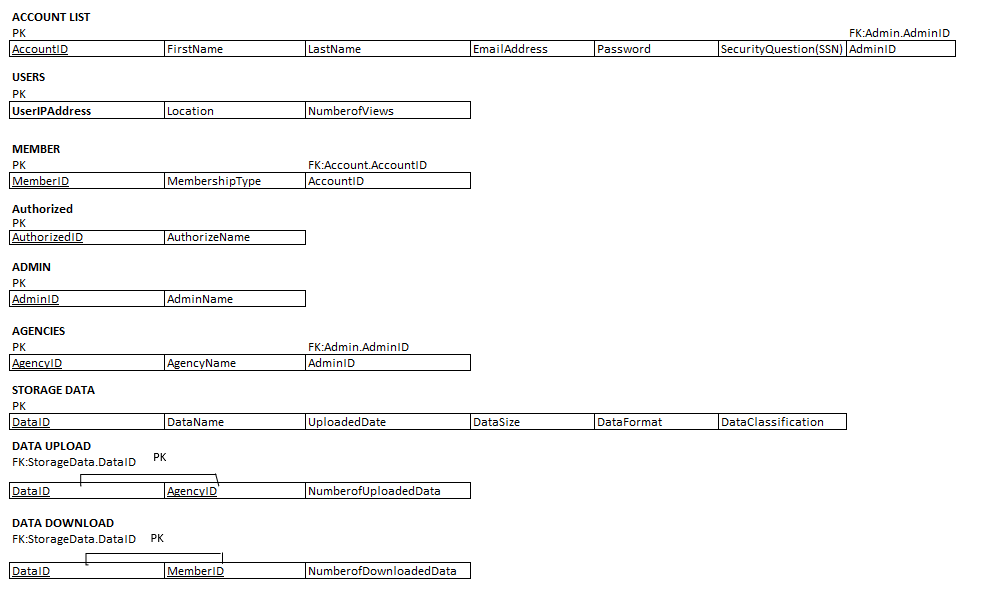


Figure 9 Relational Model

# Conclusion

The proposed Data Catalog system will have a common platform for agencies to upload data for the citizens. It will also increase the transparency in the functioning of Arkansas government and open avenues for many more innovative users of data. This Data Catalog system can also avoid data redundancy among all the agencies and have control over the restricted data for managing security and tracking. This can prioritize the importance of most frequently used data available to public and could generate the report of the downloaded data trend within agencies and the states. Data Catalog can also help to make an inform decision among executives, policymakers and users.

After implementation of the system, the second stage is proposed to have further addition to the system, such as Data Visualization, Interactive Web-browser, Automated Customer Assistances, subsequently attract investors/industries to invest in state of Arkansas approximately $10,000,000 annually and increase the potential users.

# Appendix

List of Figures

[Figure 1 Gantt Chart 10](#_Toc531006607)

[Figure 2 Activity Network Diagram 11](#_Toc531006608)

[Figure 3 Activity Diagram 19](#_Toc531006609)

[Figure 4 Use Case Diagram 20](#_Toc531006610)

[Figure 5 Class Diagram 24](#_Toc531006611)

[Figure 6 Sequence Diagram: Manage Account 25](#_Toc531006612)

[Figure 7 Sequence Diagram: Upload Data 26](#_Toc531006613)

[Figure 8 Sequence Diagram: Download Data 27](#_Toc531006614)

[Figure 9 Relational Model 27](#_Toc531006615)

[Figure 10 : Act 912 30](#_Toc531006616)

# 

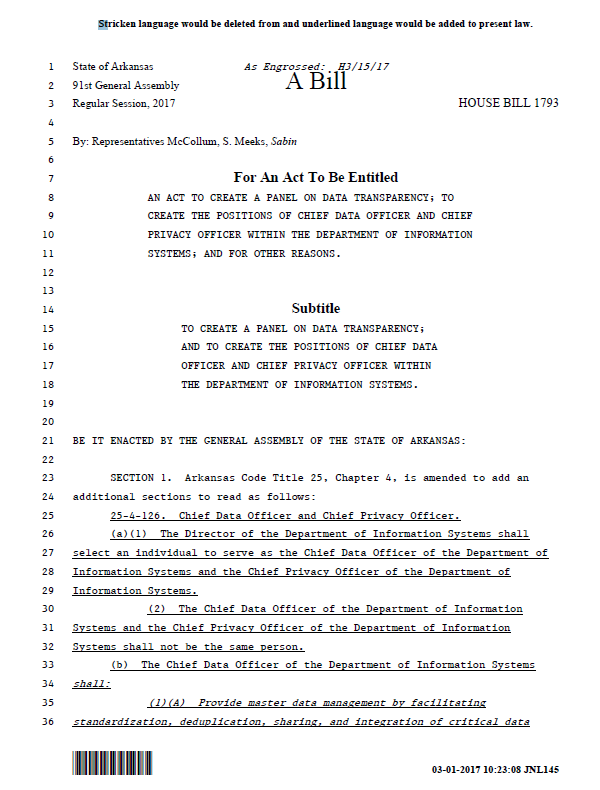


Figure 10 : Act 912

