
CSCE 606 Bush Content Management Project

Iteration 2

1. Team Roles

- a. **Product Owner, Scrum Master:** Alexandre Robic
- b. **Team members:** Krishna Kushal, Hema Deva Sagar Potala, Youngki Kim, Samuel Jeon

2. Product name: Mille Fetch

3. Client information: Robert Holzweiss, Deputy Director of George H.W. Bush Presidential Library and Museum

4. Customer Meeting

a. Meeting 1

- i. **Date:** Feb 8, 2022
- ii. **Time:** 4:30 PM - 6:30 PM
- iii. **Location:** George H.W. Bush Presidential Library and Museum
- iv. **Contents:** Introduce features for searching engine required

b. Meeting 2

- i. **Date:** March 21st
- ii. **Time:** 10:00 AM -11:00 AM Monday
- iii. **Client:** Chris Pembelton(Archivist)
- iv. **Location:** Zoom
- v. **Contents**
 - Went over iteration 0 with Chris Pembelton
 - Showed sample data taken from FOIA

- Going over user story for search : natural language query
- Showed database schematic
- Went over Yubico Key, Dual authentication

vi. **Points/Remarks made by the client**

- Need bookmark feature for user to mark the documents before downloading
- Need to be able to download over 2000 documents
- Two ways of searching 1. FOIA number 2. Folder title
- Search results on a folder level where topics in the folder are related to that folder can have irrelevant data in the folder,
No document-level control.
- Can have % or – characters while searching
- Parameter unique to the folder located in folder title and can be used to search, e.g., Date
- How do we link metadata with the document that the researcher pulls? Track metadata of each document (using ids)

vii. **Requested for**

- Metadata of folder titles and description in excel(**Received**)
- Will get around 100 files on different topics (will come in a few days)

c. **Meeting 3 (TBD)**

5. User stories

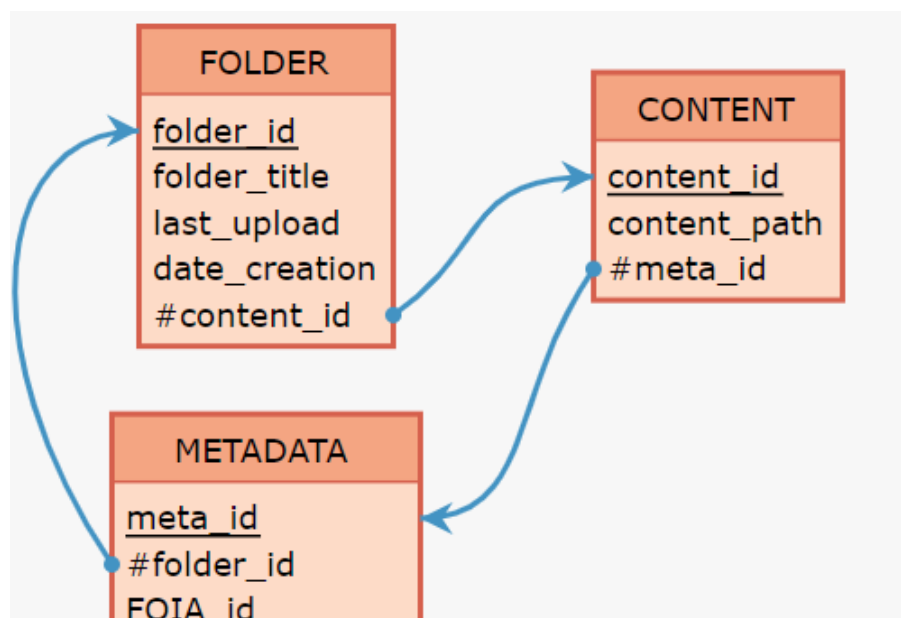
- a. **Implemented:** we picked up 3 new chores, 1 new feature, 1 spill over chore and 1 spill over feature for this iteration 2. In total we completed all the 4 chores and 1 feature (the spill over feature from last iteration)

Task type	Task Description	Status	Note
Chore	Design database model architecture	Complete	This could potentially get an

	draft for the application		update based on client's feedback. Already received an update based on feedback collected at the last meeting.
Chore	Create a 2-factor login system using OTP on the yubikey	Complete	
Chore	to obfuscate authentication details for more security	Complete	Hiding these details adds more security
chore	To mock a small set of data that will help implementation of search engine	Complete	To facilitate the search engine feature
Feature	A user should be able to see relevant search results when giving a natural language query.	Finished - requested for the review	This is the spillover feature from the last iteration.

Design Diagram: We put together a first draft of the database models during the last iteration for our application. This iteration we made a few modifications and the updated design is attached below. PS. It was created in mocodo.

Update: based on our client's feedback on monday, we updated our database schematic to the new model shown underneath



- b. **User stories Not implemented:** As mentioned above 2 chores and a feature was chosen throughout the course of this last iteration (2). Below are the ones that are in progress

Task type	Task Description	Status	Note
feature	The only authorized administrator should be able to login into the administrator account	In-progress	

It shows what the basic logic of a searching engine looks like, however, we need to go over the feature that the searching engine only shows the related data we want

George W. Bush Library: Research SiteAbout UsSearchMore

Search inside the SiteSearch

MillieFetch Search

The title that you want to find

Trump

Search

George W. Bush Library: Research SiteAbout UsSearchMore

Search inside the SiteSearch

Search Results

Folder Id	Title of the Folder
1	Donald Trump's Foreign Policy
2	President Donald J. Trump's Foreign Policy Puts America First
3	The Hill: The Memo: Tillerson Flap Puts Spotlight ON TRUMP's FOREIGN POLICY moves
4	G20 Subsidies TOK oil, gas and coal production china
5	IMI: UZBEKISTAN'S second oil and gas show held
6	Empty Promises G20 subsidies tol oil, gas and coal production
7	The Enbridge pipeline tar sands oil spill into the kalamazoo river in michigan
8	9/15 TRANSCRIPT OF THE ENBRIDGE PIPELINE OIL SPILL HEARING

c. **User stories updates/changes:**

User story	Description	Type
Obfuscate the 2FA data being transmitted	Prevent a malicious tech beginner individual from being able to see what data is being exchanged regarding the yubikey otp system	Chore
Implement a 2FA system for logging in	Implement a yubikey otp 2fa system	Chore

d. **User Stories need to be updated along with the project**

User story	Description
Feature: Search by folder name As a professional researcher I want to look up one or more files in a specific folder.	Given that I am on the main search page (first image) When I follow "Advanced search." Then I should be on the "advanced search" page (second image) When I type "Gulf War" in the title bar Then I type {ISBN NUMBER} in the ISBN bar Then I click "Search." I should be on the results page (image 3) and see all the results of folders with the term "Gulf War" in the folder or folders that contain a document that has the words "Gulf War" in the title or content or folders that are parts or folders that have the ISBN number {ISBN NUMBER} NB: "or" is inclusive in this context.

Note: It will be added with updated behavior, after getting proper feedback from the client.

6. Authentication system:

The authentication system will consist of 2 factors: a classic username/password login page and a second authentication using an OTP (One Time Password) generated using a Yubikey.

Admin page

This page will include a login form as well as a link for OTP

Username:



Password:

Login

The authentication page. For now, it always redirects to the login_otp page but once the server is fixed (postgres is causing problems), we'll be able to route accordingly.

Login processing page

Insert your yubikey and press the golden pad for one second

The login processing page. This page reads the OTP generated by the Yubikey through some simple javascript, requests the verification from our local machine, and then sends that request to the server. This allows for obfuscation of the authentication process and of the data being exchanged.

Upload page

In theory, the only way to access this page is to go through the authentication system

The upload page, once the authentication system will be complete (by the end of next week), this page will be for uploading documents to folders

and updating already uploaded files.

7. Search Engine Feature Sample Results:

This feature was just finished and requested for a code review from the other team members. Meanwhile, I attached the results post the inclusion of this feature.

Case 1: Input query is “Foreign Policy”

MillieFetch Search

Enter:

Search Results

folder_id	name
1	Donald Trump's Foreign Policy
2	President Donald J. Trump's Foreign Policy Puts America First
3	The Hill: The Memo: Tillerson Flap Puts Spotlight ON TRUMP's FOREIGN POLICY moves
13	NGO assessment of the human rights requirements in foreign military financing assistance to mexico

Case 2: Input query is climate change

MillieFetch Search

Enter:

.

Search Results

folder_id	name
9	United nations framework convention on climate change
10	U.S. views on global climate change .
11	Argentina: A useful partner on climate change

Pivotal Tracker: [here](#)

Git Repository: [here](#). Note: Please refer to SearchEngine feature branch for the search engine [user story](#).

Note: We are having some migration rollback issues while deploying to heroku. Hence have no deployed application. But Please refer to the main and [SearchEngine](#) feature branch for more information