PHASE 2 REPORT

Team 01: ***NewsFlash***

Members:

Miranda Xu
Nicholas Kazlauskas
Dickson Wong
Chandni Sehgal
Amna Iqbal

State of the Project

The release plan from the previous phase was completely rewritten to be in sync with the standards of business protocols. As such, we are not following the release plan from the first phase. Previously, the release plan was catered more for the programming requirements, over the business requirements. However, we did use and build on the content of our user stories. This time around, we effectively prioritized each user story and placed accordingly into one the three phases (one of which is this current one).

Moreover, although the previous costs (for user stories) were reasonable, the cost model itself was very strict and did not provide much flexibility. To elaborate, the current cost model assigns a cost based on the number of "programming days". For this group, it was reasonable to assume that not all the group members would be working at the same time and a programmer would only spend a few hours over a number of a days working on the project.

In addition, the original plan only consisted of three personas. As such, another persona (the library programmer) was included during this phase's personas; consequently, some user stories were reorganized for different personas.

NewsFlash had to overcome the following obstacles:

1.
The group had to decide the interaction of different classes. For example, because the interface-responsible class had to get information from the database-class, it was necessary that the proper information had to be relayed between the two. How would the database class get its information and what kind of information would it return (and in what form)?

2.
The second issue pertained to the appropriateness of a web framework over a standard GUI interface. There were significant pros and cons for each choice, however, ultimately, the difficulty in implementing a GUI and the strengths of the group (database, web familiarity) steered the project towards the standard GUI.

As a group, we were able to meet most of this phase's requirements. The only change that had to be made to the iteration plan was due to the miscalculations of the costs of implementing some user stories, and thus broke it down to several more stories with cost re-estimation. (See **re-estimation**)

Iteration Plan for Phase 2

We used the release plan from the previous phase to build on the current iteration. In keeping with that, the user stories in this iteration comprise of the following categories of tasks:

1. Adding/removing items to the database
2. Keeping track of the related items in the database - (i.e.table management)
3. Displaying lists of items
4. Parsing HTML files (articles) for references to sources

The cost estimations for each task are separated into the following parts:

i) Implementation of the function (.7 of a task)
-Database backend implementation (.3 of a task)
-UI frontend implementation and interaction with database (.4 of a task)

ii) Testing of the function (.3 of a task)

Exploration Phase

1. Functions that involve addition and removal of entries to the database require:

a) database.add_item
b) database.remove_item
c) The implementation of the different classes are a prerequisite for these tasks

This corresponds with each type of item (Source, Article, Reference classes) that is to be added into the database. Furthermore, these programming tasks require the implementation of the elements from the UI and the server.

d) Views for listing sources and articles
e) Modification pages for sources and articles
f) Interfaces for adding keywords and references

In total, the implementation of these functions will require: 6 * (.7) = 4.2

2. To Keep track of lists of items, the following is required:
a) database.intialize_tables

This will be done via the implementation inside the database: 6 * .7 = 4.2

3. The displaying of lists will be done through:
a) database.get_items

From the UI and server backend:
b) server.db_get_items
c) server.render_views

The cost for the implementation of the display article and display reference requirements is:
3 * .7 = 2.1

4. To locate references inside of an article, the algorithm is to be implemented. There are no corresponding UI related components:
database.find_references(articles)

The cost for this task is 5 * 0.7 = 3.5

Commitment Phase:
The implementation stages, will contain the tasks division as follows:

Dickson: 1abc, 3a - Database management, and accessing lists of entries
Nick: 1cde, 3bc - Server interaction and the implementation of UI elements
Amna: 2 - Initialization of the proper database tables
Miranda: 4 - Search algorithm for references within articles
Chandni: 1c - Classes implementation

Steering Phase:
Before the completion of the implementation, tests need to be written for each of the aforementioned tasks.

Dickson: Testing the correct working of lists of entries

Nick: Testing the UI works
Miranda: HTML parsing
Amna: Adding and removing functions
Chandni: Adding and removing functions
The unit tests will be divided between various functionalities, dependent on relevant table interaction. Examples include, the testing of interaction with the sources table, and the interaction with the articles table. Testing UI and server elements will incorporate the verification of the acceptance criteria. This will be done via sequences of operations to test the correct result.

Testing will cost 20 * 0.3 = 6 units of work.

In total, the cost for this phase is 4.2 + 4.2 + 2.1 + 3.5 + 6 = 20 units.

Re-Estimation of the Iteration Plan

There were two user stories that could not be completed for the release data:

1. Reference search inside an article and recursive search on article links inside

The team did not foresee the potential problems of the recursive search:

a) Adding a reference from article 1 to article 2 requires that both articles be inside the database already and adding an article to a system entails that the source of the article is inside the database already.

Therefore, more time must be spent ensuring that both conditions are met before performing a recursive search.

User Story 10a: As John (a researcher), I want to search for references inside an article, so that I can research the influence of sources. (Cost: 5, Priority: 10)
User Story 10b: As John (a research), I want to search links to articles inside of an article, so that I can research the influences of sources. (Cost: 5, Priority: 10)

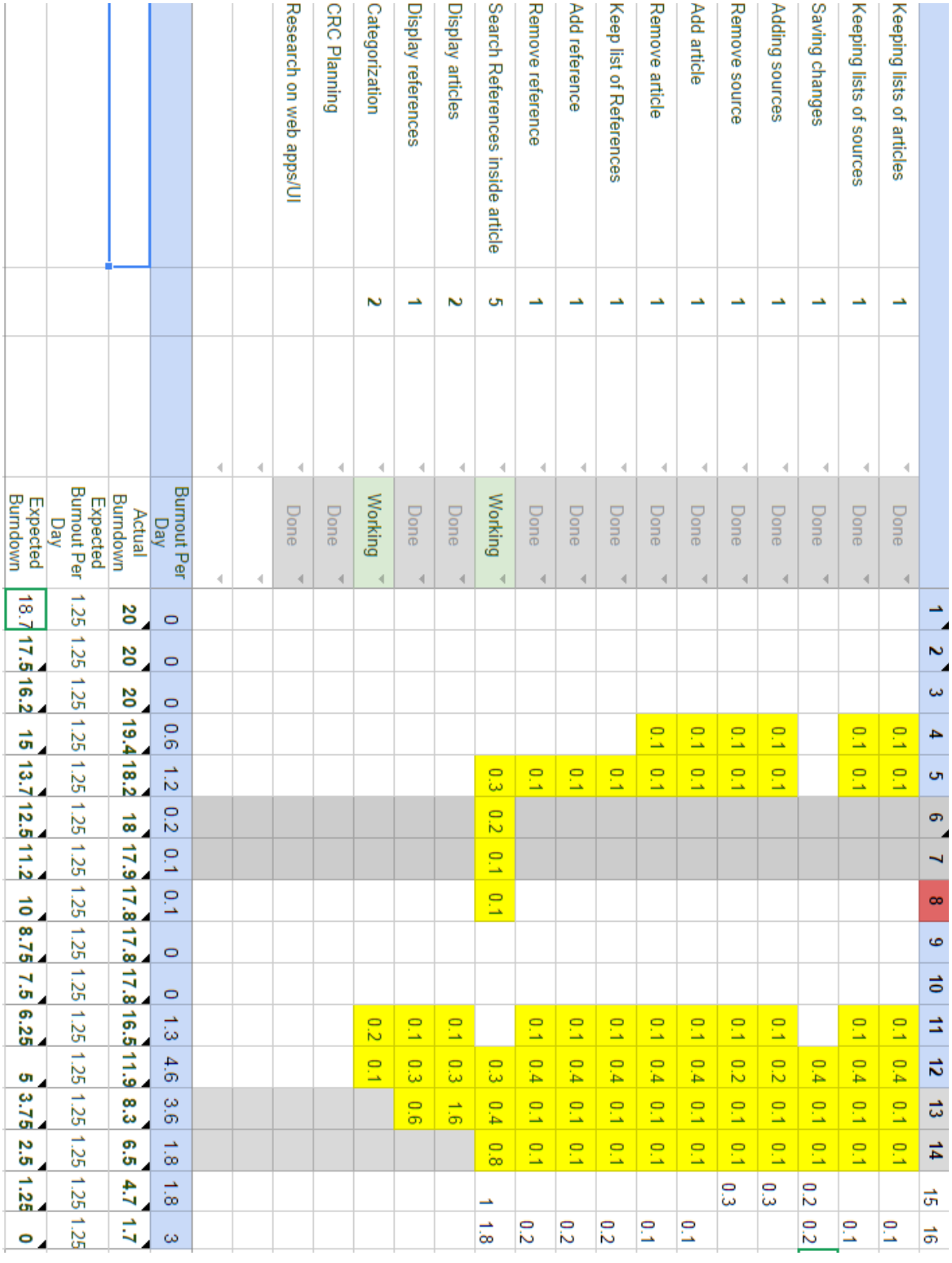
For this phase, a working version of 10a has been completed.

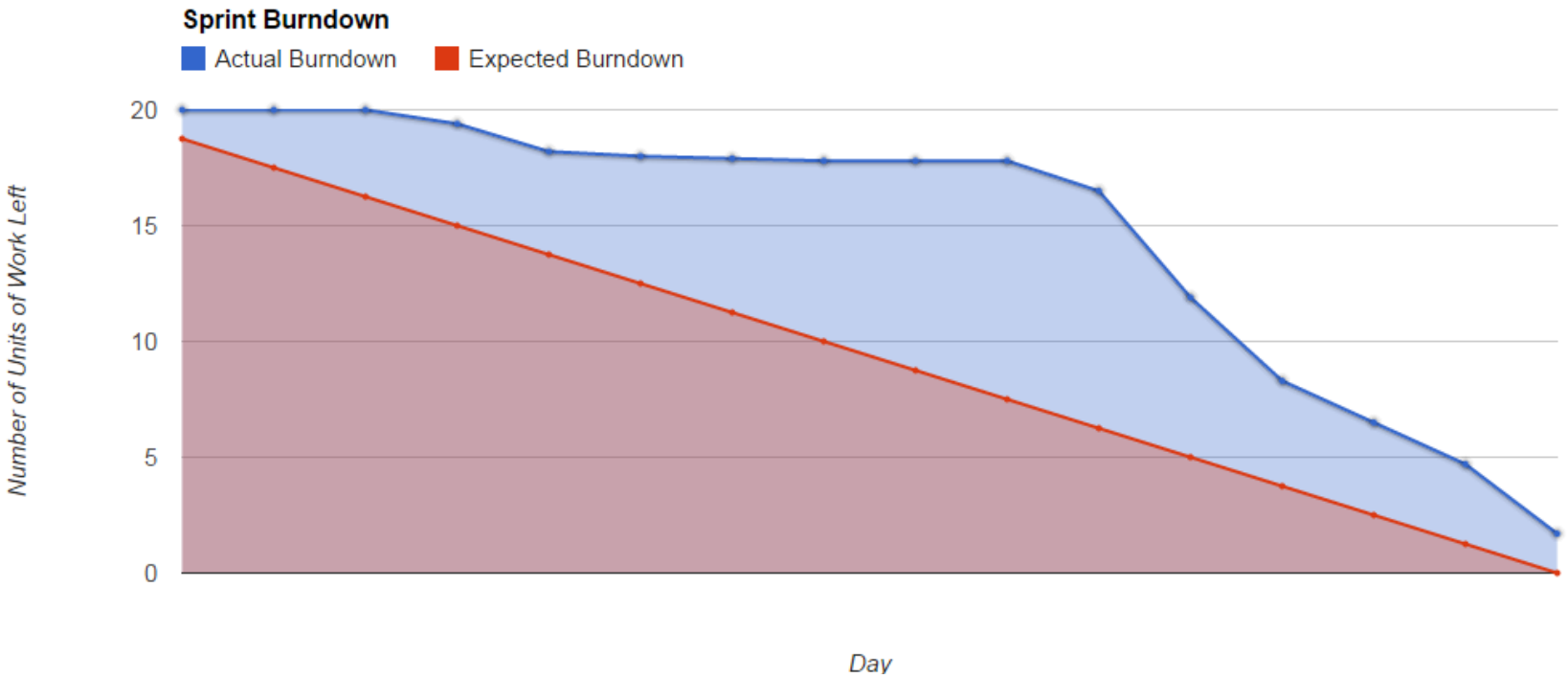
2. Categorization was not completed due to time constraints - moved to next phase.

Task Board and Burndown Chart

For the Task Board, please see the Trello Link: https://trello.com/b/EWA48O1o/cscc01-project-team01 (we have sent you the access of our Task Board, you can access this board after you confirm the system email.)

Burndown Chart




Sprint Burndown
Actual Burndown
Expected Burndown
20
15
10
5
0
Number of Units of Work Left
Day

CRC Model

Database	
add_item remove_item modify_item get_items	Article Source Reference Keyword User

Article	
Source Title Author URL Date	

Source	
Name Source (URL)	

Reference	
Article Name Source to Article reference to	

Keyword	
Source Name	

Visualizations	
Render Export	Database

Network Graph Visualization	Subclass of Visualizations
References Source Nodes Article Nodes Edges	

Server	
add_item remove_item modify_item get_items users login render_page get_visualization get_article_references	Database Visualization Web

Web	
find_article_references	Database

User	
Username Password add_item remove_item modify_item Save Categorize Search Visualize List(s) of data	Server

Librarian	Subclass of User
Retrieves data Export data	

Improvements from Phase 1

The improvements from Phase 1 are listed below:

Personas

John Bates

John Bates is a 35 year-old researcher. He has a high interest in research- specifically research that is oriented towards languages and media. He views his research as his hobby, and this enables him to devote more hours to it. He likes to live a stable and a predictable lifestyle. This is reflected in his choices of research tools. Most of the tools that he works with, do not crash often, and require minimal updates.

He works with his research tools for a substantial part each day. During the day, he works as an Assistant Professor. In the late evening, he contributes productively to his research. His research requires him to look at very large networks of information flow. As such, he prefers to work with large graphs at a single time. He enjoys looking at, and analyzing the "big picture". This helps him in simplifying the visual representations of the gathered data. These visualizations are a core part of his research. They enable him to gauge the flow of the information, and the influence of the plethora of news articles in various scenarios.

John's work requires him to spend ample time with computers. However, he would not consider himself to be technically savvy. As well, although he is adaptable to change- he prefers that everything works like he would expect it to. He likes his tools to be accessible through a web browser on a desktop computer. He is not very comfortable using his smartphone.

Jessica Walker

Jessica Walker is the head librarian. She has a strong knowledge of Linguistics, History, and English. She possesses advanced degrees in Language and Literature, and Library and Information Science. This is beneficial in her role with data storage and categorization. She enjoys her work, and is always on time. She works in the mornings, and all of her tasks are completed during her work hours.

Jessica's work with computers includes electronic resources, metadata, archives, and knowledge management. She has prior knowledge of digital repositories and open source digital tools. She has familiarity with markups, and graphical interfaces. Her days are diverse and include various tasks. She retrieves, and categorizes information according to the given requirements. She also manages

data storage, and ensures that it is in a compatible format. In her breaks, she likes to read news and articles from a range of sources.

Jessica is technically sound, and if she runs into a technical issue- she would rather communicate with support staff (over a manual). This is useful, as she works during the day- when access to support staff is convenient. She is also familiar with the latest Smartphones, and open to learning about new tools and technologies. As such, she is eager to find out the best and most efficient resources to carry out her tasks. She is familiar with console interfaces, as well as graphical interfaces. However, in terms of preference- she would much rather work with visualizations and GUIs.

Candice Walters

Candice Walters is a lawyer and a PhD student in the Graduate School of Journalism. Her research is focused on journalism ethics and the impact of media and technology on politics and public. She wants to impact the younger generation with her work. Walters began her career serving as a lawyer in the areas of digital media, communications and sports matters at a private law firm.

Candice believes that the education system needs to be dynamic and constantly improving with current trends. Mrs. Walters will be teaching a new undergraduate course called "Legal Constraints on Digital Activity" for the first time. She likes her classes to be interactive; therefore, she emphasizes the use of graphics and multimedia in her teachings. She also understands the value of digital resources and the importance of their reliability.. She sees herself playing a larger role in modern education. She believes that education should evolve with available technology. She likes to familiarize herself with every technical instrument, including popular and trendy products. On the other hand, the tools that she uses are not too sophisticated, they work well on the mobile platform, and are also presented well.

Eda Smith

Eda Smith is a library programmer. Most of her work hours are spent within the library. Her tasks involve the maintenance of the library systems. She has to ensure that all the systems within the library environment are up and running. As such, she possesses deep technical expertise, and is eager to use it in her tasks. She is the first point of contact for any open library technical ticket. As such, she is familiar with the existing tools and software present for library use.

Eda is also very familiar with the various technical platforms, including web and mobile interfaces. She always has new ideas on how to better implement existing technology. Within the library, she works with the current repositories to streamline relevant data. She is comfortable with both GUI and command line. However, in terms of preference she would much rather work with command line.

Her other tasks revolve around working with Jessica Walker. She helps out with the technical components of ongoing library research. Therefore she has exposure to new tools that are implemented within the library. Furthermore, Eda ensures that the data that the library works with is standardized and finds potential ways for different pieces of software to use them. Eda wants data searches to be relevant and fast. In keeping with this, any code that she works with is clean, well documented and optimized.

User Stories

Preface

Some of the key concepts mentioned in this document are listed below. These pertain to client needs, and the corresponding software.

- *Source:*
A publisher of news media. These include, text, audio, video, websites, blogs, Twitter accounts, YouTube channels, and many other existing and emerging forms.

- *Keyword:*
A word or phrase that is associated with a source. For example, the "Associated Press" is often referred to as "AP". Both “Associated Press” and “AP” are both keywords for the same source. The URL of the main page of a news article website could also be used as a Keyword to match the sources from references.

- *Article:*
A news piece published by a source. Since sources publish many different forms of media, this could be an article on a website, a blog post, a Twitter discussion, a video, a podcast, or anything else on the internet.

- *Tags:*
Information associated with an article - such as author, date, number of hits, etc

- *Link:*
It is the connection between two articles. This is formed when an article directly cites an article from another source.

- *Reference:*
A connection formed when an article cites a source, either by citing an article from another source, or the name of the source. The connection is from an article to another source. References may not imply a link, as not all references (such as quotations) may not contain a link.

- *Visualization:*
A representation of stored data in the software. Examples include, lists, tables, graphs, plots and pie charts.

Note: The cost model will be explained in the release plan. Priorities are defined as following: 10 being the highest, and 1 being the lowest.

User Stories

1.
As John (a researcher), I want to keep a list of articles (with appropriate tags), so that I can research the influence of certain articles and sources.

Priority: 9
Cost: 1

Acceptance Tasks:
- Sources should contain the following information:
    - Mandatory Keyword: Name
    - Optional Keyword: URL, aliases, etc
- If a mandatory field is missing, then that source should not be added.
- Keywords should be unique for all sources

Success
- A list of articles will be displayed and each article in list will contain a unique URL, date and source

Fail
The following message will be displayed:
- Empty List!

2.

As John (a researcher), I want to keep a list of sources (with appropriate keywords), so that I can research the influence of sources.

Priority: 9
Cost: 1

Acceptance Tasks:

- Sources should contain the following information
    - Mandatory Keyword: Name
    - Optional Keyword: URL, aliases, etc
- If a mandatory field is missing, the source should not be added.
- Keywords should be unique for all sources.

Success

Clicking 'List Sources' would display a list of sources with relevant names and corresponding information

Fai

The following message will be displayed:

- Empty List!

3.

As John (a researcher), I want to save changes to the list of articles and sources, so that I can retain the old data for future use.

Priority: 9
Cost: 1

Acceptance Tasks:

- The user should be able to store data (save button).

Success

The session is resumed at the user's next login.

Fail

The following message will be displayed:

- Failed to save. Try again!"
- Previous session is corrupt."

4.

As John (a researcher), I want to add a source to the application, so that I can research the influence of sources.

Priority: 9
Cost: 1

Acceptance Tasks:
- User must add a name

Success
The chosen action would be performed, followed by the displaying of the message:
"Source Added!"

Fail
The following message would be displayed:
"This name exists already, please try a different source."

5.
As John (a researcher), I want to remove a source from the application, so that I can manage my sources.

Priority: 9
Cost: 1

Acceptance Tasks:
- The user must include a keyword for a source in order to remove it.

Success
Performs the chosen actions, and displays the message: "Source Removed!"

Fail
The following message would be displayed:
"Keyword was not found; try again."

6.
As John (a researcher), I want to change keywords associated with a source, so that I can research the influence of sources.

Priority: 7

Cost: 1

Acceptance Tasks:

- Keywords can be used to match sources.
- Keywords need to be unique across sources.
- Upon clicking "Change Keyword" - User will be prompted to enter relevant information

Success

The following message would be displayed:
“Changes made successfully”

Fail

The following message will be displayed:

- “Keyword changes not saved. Please try again!”

7.

As John (a researcher), I want to add an article to the application, so that I can research the influence of sources in this article.

Priority: 9
Cost: 1

Acceptance Tasks:

- Clicking the Add button would allow:
    - the user to enter the article via the browse button (upload article) or through the text box. The latter would allow for the copy and paste functionality.
    - if the user wants to add an article, then the associated tags (such as author, name, etc) would appear. The user is also able to change these tags.
    - file/article size limit and support formats

Success

The following message would be displayed:
"Changes saved"

Fail

The following messages would be displayed:

- "Failed to add article due to size limit exceeded. Please try again!"
- "Failed to add article due to format. Please try again!"
- "Empty field. Please add article to modify"
- "Article already exists."

8.

As John (a researcher), I want to remove an article from the application, so that I can manage the articles I am interested in.

Priority: 9
Cost: 1

Acceptance Tasks:

- Clicking the remove button would allow the user
    - to enter the article that they want removed

Success

The following message would be displayed:
"Article removed!"

Fail

The following messages would be displayed:

- "Failed to remove article. Please try again!"
- "Empty field. Please enter an article to remove"
- "Article does not exist. Please try again!"

9.

As John (a researcher), I want to change the tags of an article in the application, so that I can research the influence of sources in this article (and to other articles).

Priority: 7
Cost: 1

Acceptance Tasks:

- Clicking the Change button would allow the user to:
    - enter the article that they want to remove
    - specify the tags that they want changed

Success
The following message would be displayed:
“Article Changed!"

Fail
The following messages would be displayed:
- "Failed to change the article. Please try again!"
- "Empty field. Please enter an article to change"
- "Article does not exist. Please try again!"

10.
As John (a researcher), I want to keep track of a list of references made inside an article, so that I am able to analyze the different sources influencing the article.

Priority: 10
Cost: 1

Acceptance Tasks:
- Each article should contain a list of references to other sources.
- The source should be a source that has already been included by the user.
- There should not be duplicates every time a reference is made inside the article to the source.

Success
Clicking "List References" would display the list of references made inside article

Fail
The following message would be displayed:
- No references found. Please try again.

11.
As John (a researcher), I want to be able to manually add references to an article in the application, so that the software can help me research the influence of sources.

Priority: 8
Cost: 1

Acceptance Tasks:
- The user should be able to choose an article from a list.

Success
The following message would be displayed for the added reference:
"Reference added"

Fail
The following messages would be displayed:
- Source field empty. Reference cannot be added/unmodified."
- Source not found. Try again."

12.
As John (a researcher), I want to manually remove references from an article in the application, so that the software can help me research the influence of sources.

Priority: 8
Cost: 1

Acceptance Tasks:
- The user should be able to choose an article from a list.

Success
The following message would be displayed for the reference removal:
“ Reference removed.”

Fail
The following messages would be displayed:
- Source not found. Try again."
- Source field empty. Reference cannot be removed/unmodified."

13.
As John (a researcher), I want to see a list of articles that a specific article links to, so that I can research the article.

Priority: 8
Cost: 1

Acceptance Tasks:

Success

- The "list of reference articles" can be shown after clicking the "reference list" button.
  - If the list is not empty:
    - Display list of links to other articles
    - If there are multiple articles with the same name, the user prompted for additional information about each article.

Fail

The following message would be displayed:

- "Sorry, no article is given. Please input article name/URL for the required article"

14.

As John (a researcher), I want to search (for references and sources) within the articles links(hyperlinks/web links) recursively, so that I can research more thoroughly.

Priority: 10
Cost: 5

Acceptance Tasks:

- On addition of article, the user needs to specify whether automated searching occurs
- Each link to another article can be searched recursively
- Every reference to a source and every link to another article is saved
- Each reference and link should be unique and never duplicated

Success

The references and sources within the article will be displayed

Fail

The following messages would be displayed:

- "No such article found. Please try again"
- "No references or sources found. Please try again"

15.

As John (a researcher), I want a display of the list of the references a source makes to other sources, so that I can research the influence of different sources and their relation to one another.

Priority: 9
Cost: 2

Acceptance Tasks:

- The user is prompted for the source, and only sources specified can be displayed.

Success

A list of references is displayed.

- For a non empty list, the name of the source is displayed for each reference

Fail

The following message would be displayed:

- "Sorry, no source is given, please input the required sources”
- "Sorry, this source cannot be parsed. Please check the source name/URL and try again."

16.

As John (a researcher), I want a graph visualization of references from a source to other sources, so that I can analyze the references that are made in articles from a certain source.

Value: 8
Cost: 4

Acceptance Tasks:

Success

- A source to sources graph contains:
    - each node represents a source
    - every directed edge represents a reference inside an article made from one source to another
- If there are no correlations between articles/sources then there would be no edges
- Edges could show the number of references made

Fail

The following message would be displayed:

- "Sorry, no source is given, please input the required sources/article”
- "Sorry, some of the sources/articles cannot be parsed. Please check the names/URLs and try again."

17.

As John (a researcher), I want to see graphical representation (network connections) of a specific article with the reference sources of this article, so that I can see the sources referred by the article.

Priority: 6

Cost: 3

Acceptance Tasks:

Success:

- An article to sources graph contains:
    - each node represent either the article in question or the sources
    - every directed edge would represent the reference the article makes to source
- If there are no correlations between articles/sources then there would be no edges

Fail

The following message would be displayed:

- "Sorry, no source/article is given, please input the required sources/articles"
- "Sorry, some of the sources/articles cannot be parsed. Please check the names/URLs and try again."

18.

As John (a researcher), I want a graph visualization of references made between different sources, so that I can research the relationships between sources.

Priority: 8
Cost: 3

Acceptance Tasks:

Success

- Nodes represent sources
- Edges represent references made from one source to another
- The number of references could be shown beside the edge

Fail

- "No sources were specified."

19.

As John (a researcher), I want to see graphical representations (network connections) of related articles and sources, so that I can better understand the big picture/public sphere.

Priority: 6

Cost: 4

Acceptance Tasks:

Success

- Combined sources and articles graph contain:
  - Sources being represented by larger nodes
  - Articles being represented by smaller nodes
  - Edges being represented by connections from the previous types
  - Every y node at the reference level (reference are not shown completely on the graphical representation), can be set as a core source. This would allow the graphical representation to completely show the references.
  - If there are no correlations between articles/sources then there would be no edges

Fail

The following message would be displayed:

- "Sorry, no source/article is given, please input the required sources/articles"
- "Sorry, some of the sources/articles cannot be parsed. Please check the names/URLs and try again."

20.

As John (a researcher), I want to see graphical representations (2D line plots) of the number of references made to sources, so I can analyze the changes of popularity of sources over time.

Priority: 8
Cost: 4

Acceptance Tasks:

Success

- On the graph, each curve represents one reference source site, labeled with sitename.
- Different curves should have different colors.
- The ordinate should be the frequency of the sources are referenced by.
- The abscissa should be the time line of the reference time.
- The abscissa time unit can be set to month(default), year, week and season.

Fail

The following message would be displayed:

- "Sorry, no source is given, please input the sources as you need."
- "Sorry, some of the sources cannot be parsed. Please check their names/URLs and try again."

21.
As Jessica (a librarian), I want to store collections of news, discussions, and articles, so that it can be used in future references.

Priority: 8
Cost: 4
Acceptance Tasks:
- Each entry should contain different information and can only be saved with the presence of mandatory tags.
- The data format should be consistent throughout all entries.
- There should be enough data for each entry, dependent on the articles referenced.
- The format should be compact, and organized in a reasonably.

Success
The following message would be displayed:
“Data stored successfully"

Fail
The following message would be displayed:
- "Mandatory information is missing. Please try again"
- "Data failed to store. Please try again"

22.
As Jessica (a librarian), I want to categorize the data by date, title, or keywords/tags, so that I can retrieve certain pieces faster.

Priority: 8
Cost: 2

Acceptance Tasks:
- There will be three options for data categorization.
    - Date
    - Title
    - Tags

Success
The following message would be displayed:
"Data categorized successfully"

- data can be sorted by the category chosen;

Fail
The following message would be displayed:
"The following was missing X" (X is the chosen sorting criteria)

- Data that cannot be sorted by the category will be listed separately.

23.
As Eda (a librarian programmer), I want to export data in a compatible format (e.g. XML) with certain other software, so that the data will be accessible in the future on any systems and software.

Priority: 7
Cost: 3

Acceptance Tasks:
- The data format will follow the requirements set out by the client as the project evolves.
- Other software (perhaps visualization) will be able to use the database files.

For now:
- Upon clicking the "Export" button, the user will be shown 'save-as' options

Success
- the data will be exported and saved wherever user specifies
- The following message would be displayed:
"Data exported successfully"

Fail
The following message would be displayed:
- "An error occurred while exporting. Please try again"
- "Incompatible format. Please try again"

24.
As Jessica (a librarian), I want a login screen to authenticate and authorize users, so that the data access is restricted and secure.

Priority: 5
Cost: 2

Acceptance Tasks:
- There should be a login screen, to allow the users to input their credentials.

Success
The following message would be displayed:
"Login successful" and "User authenticated"

Fail
The following messages would be displayed:
- "Login failed. Please try again!"
- "Authentication failed. Please try again!"

25.
As Candice (an educator), I want to see table visualizations on various categories, such as authors, source location, or references within articles, so that I can have a comprehensive view of these articles.

Priority: 4
Cost: 2

Acceptance Tasks:

Success
- the top row indicates every reference sources name, and the right side column labels the time unit(month, year, week or season).
- The time range can be set manually by the users
- The bottom row will show the number of times/citations for each reference within a defined time range.

Fail
The following message would be displayed:
- "Category not specified."
- "Articles with this category do not exist"

26.
As Candice (an educator), I want to allow student level accounts to only view and access visualizations and queries, so that it will prevent unauthorized changes to the database.

Priority: 3
Cost: 2

Acceptance Tasks:

- An Educator login screen will have checkbox options to allow Import/Export data

Success

- A checkmark beside "Allow Import/Export data" will allow data sharing

Fail

- if a student authenticator will try to import/export data - the following message would be displayed: "Error! Data Sharing not allowed. Please contact admin"

27.

As Candice (an educator), I want to export the data (sources, articles, visualizations) through a web interface, so that the data is more readily available to larger groups of people.

Priority: 6
Cost: 4

Acceptance Tasks:

Success

For users with access to data sharing.

- Select export data: prompts user to save

Fail

The following message would be displayed:

- "Sorry, you don't have the access to this data. Please check with the administrator

28.

As Candice (an educator), I want to import the data (sources, articles, visualizations) through a web interface, so that the data is more readily available to larger groups of people.

Priority: 6
Cost: 4

Acceptance Tasks:

Success

For the users with access of data sharing:

- Select import data: set current database to data

- Indicate whether marked the data with data author.

Fail

The following message would be displayed:

- "Sorry, you don't have the access of this data. Please check with the administrator(with email address)."
- "Please select the data you need”
- "Sorry, your data type is incompatible, please check again."

System Design

The software manages a collection of articles, published by sources, and the links and references between them. The user can input sources, articles, and can add links and references. The managed articles can be categorized to filter and find specific content. With the managed data, it creates visual representations of the links and references.

The main components of the software are:

1.
The database that stores the source entries and article entries, each of which may contain different types of data.

2.
The interface that allows the user to load from a database, modify the database by adding new sources and articles.

The interface also allows for a variety of visualizations to be created from the data stored. In the system, the user must interact with the interface and specify sources and articles. The interface will then find connections and store both the sources/articles and connections that it has found into the database.

System Metaphor

Every store has a list of different groups of consumers that often buy from the store. These include various groups, from the elderly to the hip young kids (which is a very basic example). However, every group of consumers also visits different types of stores.

A company hires an analyst to understand these relationships - the company will tell the analyst exactly which groups of consumers the company is interested in and will also specify which stores (whose consumer traffic) are particularly of interest to the company. The analyst will then research a store, taking note of the consumer groups that are of interest to the company.

Sources

We liken the sources of information (that we want to analyze) to be the different stores. We expect that a certain consumer group will likely visit many stores - just like how an article may contain many references to sources. We can also draw connections between two stores if they happen to have the same consumer groups visit them - just like how we draw connections between sources of

information if there is an article from one source that references from another. Information about stores is stored inside the filing cabinet (database).

Articles

We liken the articles to be different consumer groups. Different consumer groups visit different stores -similar to how different articles make references to different sources. Information about consumer groups is stored inside the filing cabinet (database).

User

The company is the user, providing a list of sources (consumer groups), and articles to analyze (stores). The company interacts with the analyst.

Analyst

- The analyst is similar to the interface - it is who the company (user) interacts with.
- The analyst is responsible for interpreting the information from the company (similar to how the interface interprets user input).
- The analyst will seemingly "analyze" the relationships between the consumer groups and the stores (similar to how the connections between articles and references just appear on screen for the user to see when they give input).
- The analyst is responsible for keeping a filing cabinet (or of the sort) to keep track of which consumer groups and stores the company wants to look at (like how the interface will interact with the database).
- The analyst is also responsible for presenting visualizations of the data that he has collected to the company, depending on what visualizations the company would like to see. (just like how the user can choose from a few visualizations)
- The analyst will interact with the company (the user) and keep track of information with his filing cabinet (database).

Database

It is what the analyst uses to keep track of the researched information, as well as all the consumer groups and the stores that have been provided by the company (it could be a filing cabinet).

Release Plan

== Cost ==
Unit cost - one member of the team working on the project for a day.

We have 5 members, so in a day, we have 5 units of work that can be completed. But we are not working every day in the week, so we expect to only have 2 days of the week to work on the project.

For example, in one week,
5 units per day * 2 days per week = 10 units of work per week
==========
Release 1 - 2 weeks (01/10/2014 - 15/10/2014)

The main functionalities needed by the business are implemented during this phase - searching for references to sources inside an article and keeping a list of references that are made inside articles.

Furthermore, other important functionalities (like, keeping track of a list of sources, articles, and keywords) implemented during this phase:

Note that although user stories 10 and 14 have priorities 10, the other user stories (except 13 and 22) are prerequisites for 10 and 14. As such, iteration planning will be carried out so that the prerequisite stories are implemented first.

10.
As John (a researcher), I want to keep track of a list of references made inside an article, so that I am able to analyze the different sources that influence the article.

Priority: 10
Cost: 1

14.
As John (a researcher), I want to search (for references and sources) within the articles links (hyperlinks/web links) recursively, so that I am able to research it more thoroughly.

Priority: 10
Cost: 5

1.

As John (a researcher), I want to keep a list of articles (with appropriate tags), so that I can research the influence of certain articles and sources.

Priority: 9
Cost: 1

2.

As John (a researcher), I want to keep a list of sources (with appropriate keywords), so that I can research the influence of sources.

Priority: 9
Cost: 1

3.

As John (a researcher), I want to save changes to the list of articles and sources, so that I can retain the old data and continue from there.

Priority: 9
Cost: 1

4.

As John (a researcher), I want to add a source to the application, so that I can research the influence of sources.

Priority: 9
Cost: 1

5.

As John (a researcher), I want to remove a source from the application, so that the I can manage my sources.

Priority: 9
Cost: 1

7.

As John (a researcher), I want to add an article to the application, so that I can research the influence of sources in that article.

Priority: 9
Cost: 1

8.
As John (a researcher), I want to remove an article from the application, so that I can manage the articles I am interested in.

Priority: 9
Cost: 1

15.
As John (a researcher), I want a display of the list of the references a source makes to other sources, so that I can research the influence of different sources and their relation to one another.

Priority: 9
Cost: 2

11.
As John (a researcher), I want to manually add references to an article in the application, so that the software can help me research the influence of sources.

Priority: 8
Cost: 1

12.
As John (a researcher), I want to manually remove references from an article in the application, so that the software can help me research the influence of sources.

Priority: 8
Cost: 1

13.
As John (a researcher), I want to see a list of articles that a specific article links to, so that I can research the article.

Priority: 8
Cost: 1

22.
As Jessica (a librarian), I want data categorization by title, or keywords/tags, so that I can retrieve certain pieces faster.

Priority: 8
Cost: 2

The total cost of this is 20 (or 10 units per week * 2 weeks)
==========

Release 2 - 4 weeks (16/10/2014 - 13/11/2014)

The second release will implement the next most important requirements for the business: Graph visualizations and plots representations for different sources, articles, and the references within them.

Furthermore, the implementation of different types of data exports occurs during this phase. This includes, creating an XML format database and the storage of different news articles and discussions.

16.
As John (a researcher), I want a graph visualization of references from a source to other sources, so that I can analyze the references that are made in articles from a certain source.

Value: 8
Cost: 4

18.
As John (a researcher), I want a graph visualization of references made between different sources, so that I can research the relationships between sources.

Priority: 8
Cost: 4

20.
As John (a researcher), I want to see graphical representations (2D line plots) of the number of references made to sources, so that I can analyze the changes to popularity of sources over time.

Priority: 8
Cost: 4

21.
As Jessica (a librarian), I want to store collections of news, discussions, and articles, so that it can be used in future references.

Priority: 8
Cost: 4

20A.
As John (a researcher), I want to see graphical representations (bar graph) of the number of references made to a target list of sources, so that I can analyze the changes to popularity of sources over time.

Priority: 7
Cost: 4

23.
As Eda (library programmer), I want to export data in a compatible format (e.g. XML) to certain other software, so that the data will be accessible in the future on any systems and software.

Priority: 7
Cost: 4

6.
As John (a researcher), I want to change keywords associated with a source, so that I can research the influence of sources.

Priority: 7
Cost: 1

9.
As John (a researcher), I want to change the tags of an article in the application, so that I can research the influence of sources in this article (and to other articles).

Priority: 7
Cost: 1

19.
As John (a researcher), I want to see graphical representations (network connections) of related articles and sources, so that I can better my understanding of the big picture of the public sphere.

Priority: 6

Cost: 4

The total cost for this phase is 40. (or 4 weeks)

==========

Release 3 - 2 weeks (13/11/2014 - 26/11/2014)

The last release involves implementation of multiple user access tools. These include, a login system and user tools for data exportation and import.

27.
As Candice (an educator), I want to export the data (sources, articles, visualizations) through a web interface, so that the data is more readily available to larger groups of people.

Priority: 6
Cost: 4

28.
As Candice (an educator), I want to import the data (sources, articles, visualizations) through a web interface, so that the data is more readily available to larger groups of people.

Priority: 6
Cost: 4

24.
As Jessica (a librarian), I want a login screen to authenticate and authorize users, so that the data access is restricted and secure.
Priority: 5
Cost: 2

25.
As Candice (an educator), I want to see table visualizations on various categories, such as authors, source location, or references within articles, so that I can have a comprehensive view of these articles.
Priority: 4
Cost: 2

26.

As Candice (an educator), I want to allow student level accounts to only view and access visualizations and queries, so that it will prevent unauthorized changes to the database.

Priority: 3
Cost: 2

The total cost for this phase is 14.