PHASE 3 REPORT

Group 01: NewsFlash

Team Members:
Nicholas Kazlauskas
Dickson Wong
Miranda Xu
Chandni Sehgal
Amna Iqbal

State of the Project

The Release plan underwent some major changes. These include:

We moved some user stories from iteration 2 to iteration 3. As well, some stories previously in iteration 3 were shifted to iteration 4. These changes were made based on the project velocity of the team at the time.

1. User stories/requirements have been added at the time between iteration 2 and 3, which were mostly of lower priority (see iteration 3 and adjusted for 4)
2. The release plan is structured around fixed length iterations with fixed release dates set by the client.

To expand on 1, the project velocity was re-estimated during iteration 2 and it was deemed that the team would not have enough time to complete stories 14B, 16, 18, and 20 – requirements involving graphical visualizations and reference searching. These were shifted to iteration 3 and stories relating to data exports and updating data for articles inside the watch list (all stories originally planned for iteration 3) were moved to iteration 4.

On other hand, iteration 3 went more smoothly based on the amount of time the team had to complete the project.

To expand on 2, new requirements were realized right before iteration 3: watching articles for updates, WARC exports, and RSS feeds to automatically generate data.

From the previous phase, the major changes pertained to:

1. Release plan – changes above

The challenges faced during this phase as are follows:

1. The project velocity of the team was re-estimated to be much lower at one point during iteration 2. This was mostly because of other commitments team members had (exams and assignments during those weeks). Consequently, the project velocity reduced to less than half of what it was expected (see iteration plan 2).
2. Furthermore, changes to the database were constantly made were several reasons:
    a. The error messages displayed to the users did not always reflect the actual error. Different types of error handling had to be added.
    b. It was neither entirely obvious nor easy to work with the database to generate data for visualizations. Major changes to the database were made throughout the entire project.

Updated Release Plan

== Cost ==

Unit cost –refers to 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 can expect to only have 2 days of the week to work on the project.

For example, in a given week:
5 units per day * 2 days per week = 10 units of work per week

==========

Iteration 1 (2 weeks) - Release Date: 15/10/2014

This iteration takes into account the implementation of the essential functionalities needed by the business. These include:

- The search for references to sources within an article and then retaining a list of referencesthat are made inside articles.

This iteration also focuses on the implementation of:

- Keeping track of a list of sources, articles, and keywords.
- Adding the list of sources, articles and keywords to the system,

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

14A.
As John (a researcher), I want to search (for references and sources) within the article, 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 canresearch the influence of sources.

Priority: 9
Cost: 1

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

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 thatI can manage the articles that interest me.

Priority: 9
Cost: 1

15.
As John (a researcher), I want a display of the list of 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 specificarticle links to, so that I can research the article.

Priority: 8
Cost: 1

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

Priority: 7
Cost: 1

The total cost for this iteration is 20.

==========

Iteration 2 (2 weeks) –End Date: 30-10-2014

The second release will focus on the implementation of:

- Association of tags to an article.
- Graph visualizations of sources and associated references to one another (including both network graphs and 2d plots)

In addition, this iteration will cater to:

- Recursive reference searching inside an article
- Sorting by category

14B.
As John (a researcher), I want to search (for references and sources) within links of an article, so that I am able to research it more thoroughly.

Priority: 10
Cost: 3

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

Priority: 8
Cost: 2

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

Priority: 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 the popularity of sources over time.

Priority: 8
Cost: 4

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

The planned cost of this iteration is 20.

==========

Adjustments made:

User stories 14B, 16, 18, 20 are moved to iteration 3 (see iteration 2 plan and its re-estimations).

Iteration 3 (2 weeks) - Release Date: 14-11-2014

Iteration 3 is geared towards completing the visualizations and updating entries inside the database with RSS feeds and article watch list capabilities. Furthermore, implementation of various data exports are planned to be completed.

This includes:

- Graph visualizations and plots representations. These pertain to different sources, articles, and the references.
- Watch list of articles
- RSS Feed from websites
- XML and WARC exports

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 the popularity of sources over time.

Priority: 7
Cost: 4

30.

As John (a researcher), I want to specify the articles that I can getupdates on (watch list), so that my research is relevant.

Priority: 7
Cost: 1

31 .
As John (a researcher), I want to receive updates on thearticles that are on my watch list, so that my research is relevant.

Priority: 7
Cost: 2

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

Priority: 7
Cost: 3

35.

As John (a researcher), I want to add a website to automatically get articles from the RSS feed, so that I my research and information is more complete.

Priority: 7

Cost:4

32.

As Jessica (a librarian), I want to be able to store archives of articles and related information in a WARC format, so that the data will be accessible in the future on any system.

Priority: 6
Cost: 3

23.

As Jessica (a librarian), I want to be able 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 system.

Priority: 7
Cost: 3
==========

Adjustments made:

From iteration 2, user stories 14B, 16, 18, 20 are completed during this iteration instead.

User stories 20A, 31, 32, 23 are moved to iteration 4. (see iteration 3 plan and its adjustments)

Iteration 4 (2 weeks) - Release Date: 28/11/2014

The last release will focus on multiple user access tools. These include the implementation of a login system and the provision of data export and import tools for the users.

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 for the authentication and authorization of users so that the data access is restricted and secure.

Priority: 5
Cost: 2

25.
As Candice (an educator), I want to see the table of 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 the allowance for student level accounts to onlyviewing and accessing of visualizations and queries, so that there will be a prevention of unauthorized changes to the database.

Priority: 3
Cost: 2

The total cost planned for this phase is: 14.

==========

Adjustments made:

User stories 31, 32, 23 were moved to this iteration.

Iteration Planning and Re-estimations

Iteration Plan 2:

The following user stories are planned to be completed and the breakdown is as follows:

(The breakdown and costs are all made in programming days, as described in the release plan)

14B – recursive search of references inside links of articles
Cost: 3

a. Reference search –finding new articles and generating references: 1 - Miranda
b. Database functionality: adding new articles and sources if they do not exist already: 1 -Dickson
c. UI/testing: 1– Nick, Miranda

22 – Categorization of data in the UI
Cost: 2

a. UI: 2 – Nick

16 – Graph visualization of a source to other sources
Cost: 4

a. Database functionality: matching references made between sources: 1 – Dickson
b. Generating data and visualization: 2 – Amna
c. UI/Testing: 2 - Nick/Amna

18 – Graph visualization of references between different sources
Cost: 4

a. Database Functionality: getting articles from sources and matching articles and sources to get references: 1 – Dickson
b. Generating data and visualization: 2 – Miranda
c. UI/Testing: 1 - Nick/Miranda

20 – 2D line plot of references made to sources
Cost: 4

a. Database functionality: adding functionality to get references from sources: 1 – Nick
b. Generating data and visualization: 2 – Chandni
c. UI/Testing: 1 – Nick/Chandni

9 – Article tagging
Cost: 1

a. Database changes: 0.3 – Dickson
b. UI/Testing: 0.7 – Nick, Amna

Furthermore, additional requirements for this iteration were:

a. Improving database to return more meaningful data for the server: 2 – Dickson

Project Velocity and Costs

The expected project velocity for this iteration was10 units per week.

However, over the iteration, the team had to deal with other commitments (exams, assignments, space shuttles), and as a result, the project velocity reduced to 4units per week.

As a result, user stories14B, 16. 18. 20 were not completed and these were moved to iteration 3.

The cost of maintaining the database to produce better return values took more time than expected: instead of 2 programming days, it took 5 programming days to fully complete and fully implement and test.

The costs of the other requirements were as expected.

Iteration Plan 3:

The following user stories are planned to be completed and the breakdown is as follows:

(The breakdown and costs are all made in programming days, as described in the release plan)

==== From Iteration 2 Shifted Over ====

14B – recursive search of references inside links of articles
Cost: 3

a. Reference search – finding new articles and generating references: 2 - Miranda
b. Database functionality: adding new articles and sources if they do not exist already: 1 -Dickson
c. UI/testing: 2 – Nick, Miranda

16 – Graph visualization of a source to other sources
Cost: 4

d. Database functionality: matching references made between sources: 1 – Dickson
e. Generating data and visualization: 2 – Amna
f. UI/Testing: 2 - Nick/Amna

18 – Graph visualization of references between different sources
Cost: 4

d. Database Functionality: getting articles from sources and matching articles and sources to get references: 1 – Dickson
e. Generating data and visualization: 2 – Miranda
f. UI/Testing: 1 - Nick/Miranda

20 – 2D line plot of references made to sources
Cost: 4

d. Database functionality: adding functionality to get references from sources: 1 – Nick
e. Generating data and visualization: 2 – Chandni
f. UI/Testing: 1 – Nick/Chandni

========

20A – 2D bar graph of references made to target article
Cost: 4

a. Database functionality: getting references from specific combination of articles, sources: 1 – Dickson

b. Generating data and visualization: 2 – Chandni
c. UI/Testing: 1 – Nick/Chandni

30 – Watch list specification of articles
Cost: 1

a. Database: 0.5 – Dickson
b. UI/Testing: 0.5 – Nick/Dickson

19 – Graph visualization of sources and articles
Cost: 3

a. Database functionality: 0.5 – Dickson
b. Generating Data and visualization: 2.5– Miranda
c. UI/Testing: 1 – Nick/Miranda

35 –RSS Feeds

Cost: 3

a. RSS feed information and parsing: 1.5 - Amna
b. UI/Testing: 1.5 – Nick/Amna

==== Shifted Over to Iteration 4 because of Iteration 2 changes====

31 – Watch list update on articles
Cost: 2

a. Database: 0.5 – Dickson
b. UI/Testing: 0.5 – Nick/Dickson

32 – WARC archiving

Cost: 3

a. Generation of archive format: 2 - Amna
b. UI/Testing: 1 - Nick/Chandni

23 – XML Formating

Cost: 3

a. Generation of formatted data: 2 – Amna
b. UI/Testing – Nick/Chandni

==========

Other requirements for this iteration are:

a. Improving the UI so that it is more intuitive – users can add URLs to sources: 1- Nick
b. Making changes to the database to suit these needs and testing them: 1 – Nick/Dickson

Project Velocity and Costs

The expected project velocity for this iteration was 14 units per week. Members were expected to and able to complete more work than before because members had less commitments (exams, assignments) to tend to than before.

The team was not able to meet the requirement 20A; the amount of time maintaining the UI and the database was also more than expected (see below). So although the team did not finish user story 20a, the project velocity of the team was higher than was expected.

Consequently, user story 20A is moved to iteration 4. Furthermore, although the functionality has been completed to generate data and visualizations; there was not enough time to connect this to the database. Consequently, the current release does not contain the visualizations. Similarly, functionality to pull articles from RSS feeds has been implemented, but has not been integrated into the UI. These will be completed at the beginning of iteration 4.

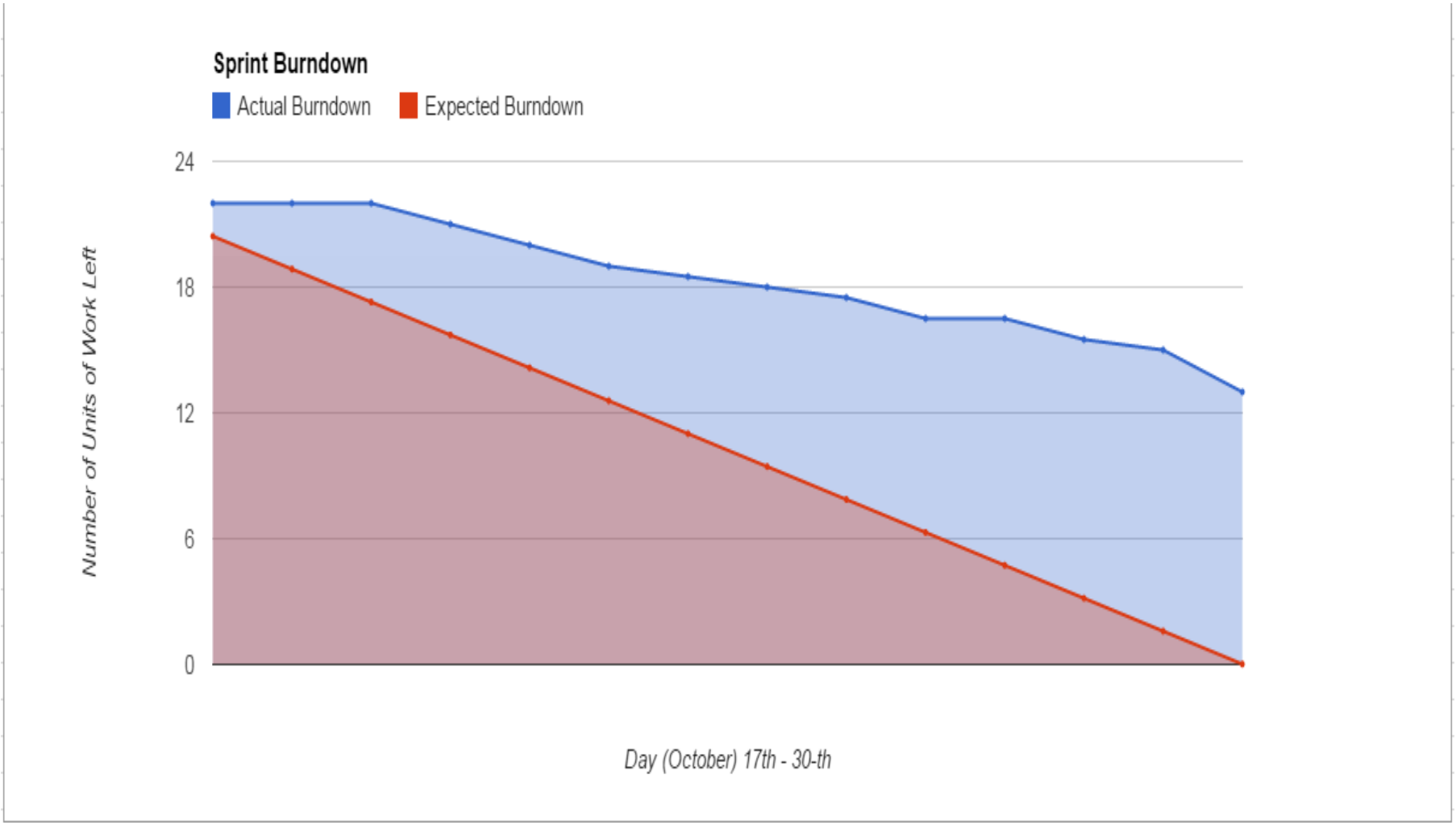
Because of the shift from iteration 2, user stories 31, 32, 23 are shifted to iteration 4.

The costs of the user stories were as expected; however, as mentioned before, the cost of maintaining the UI and database were each respectively 3 (2 more than expected).

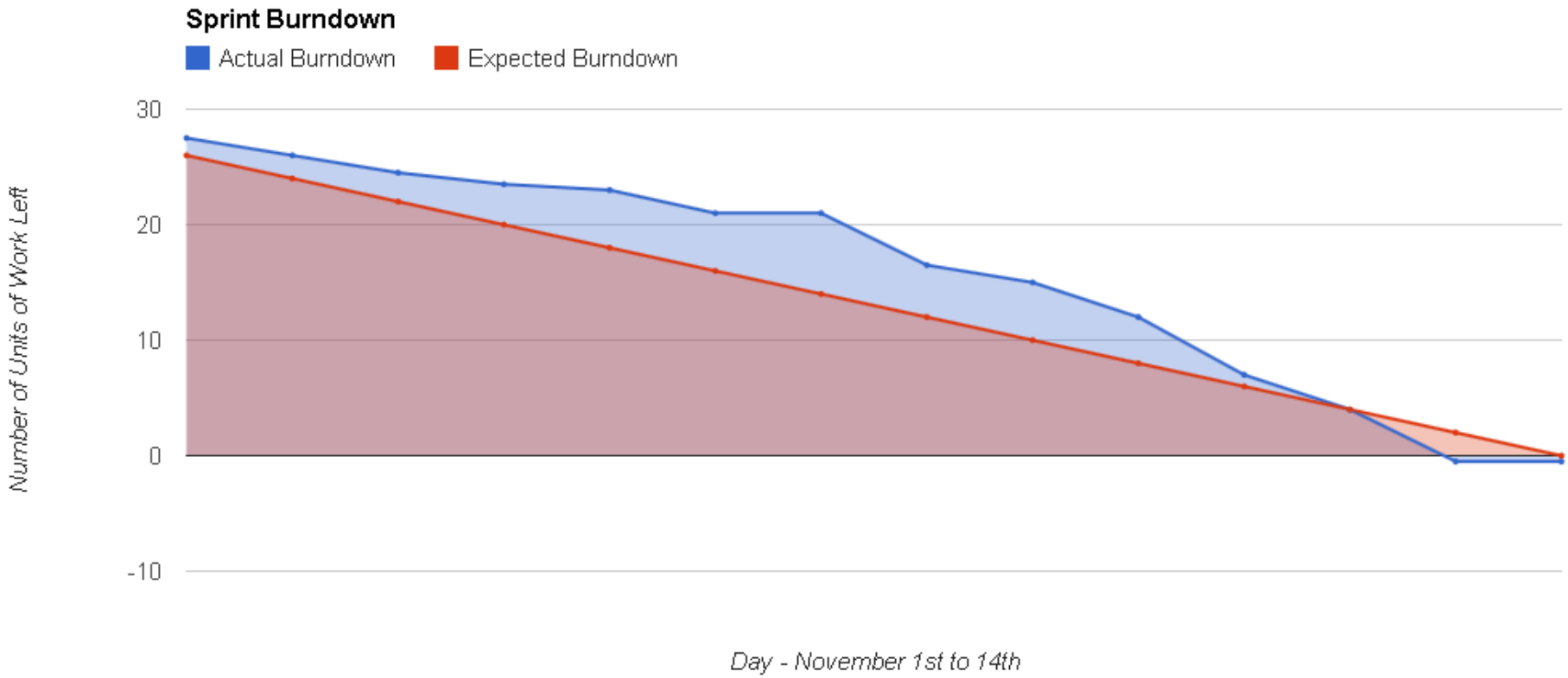
Task Board and the Burndown Chart

Our Task Board is available via the Trello Link:
https://trello.com/b/EWA48O1o/cscc01projectteam01
(The access to our Task Board is through the previously sent system email). See below for burndown charts. The first two diagrams correspond to iteration 2 and the latter iteration 3.

User Story	Time (est)	Assigned	Status	Work Done per day per task [units]													
				October													
				17	18	19	20	21	22	23	24	25	26	27	28	29	30
14B - recursive search	5		Incomplete														
22 – Categorization of data in the UI	2		Done				1	1									
16 – Graph visualization of a source to other sources	4		Incomplete														0.5
18 – Graph visualization of references between different sources	4		Incomplete														
20 – 2D line plot of references made to sources	4		Incomplete														0.5
9 – Article tagging	1		Done						1								
Improving database to return more meaningful data for the server:	2		Done							0.5	0.5	0.5	1		1	0.5	1
			Burnout Per Day	0	0	0	1	1	1	0.5	0.5	0.5	1	0	1	0.5	2
			Actual Burndown	22	22	22	21	20	19	18.5	18	17.5	16.5	16.5	15.5	15	13
			Expected Burnout Per Day	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57	1.57
			Expected Burndown	20.4	18.8	17.2	15.7	14.1	12.5	11.0	9.43	7.86	6.29	4.72	3.15	1.58	0.01


Sprint Burndown
Actual Burndown
Expected Burndown
24
18
12
6
0
Number of Units of Work Left
Day (October) 17th - 30-th

				November													
				1	2	3	4	5	6	7	8	9	10	11	12	13	14
14B - recursive search	3		Done		0.5	1.5										0.5	
16 – Graph visualization of a source to other sources	4		Done										1	1	1	1	
18 – Graph visualization of references between different sources	4		Done										1	1	1	1	
20 – 2D line plot of references made to sources	4		Done								1	1			1	1	
20A - bar graph	4		Incomplete														
30 – Watch list specification of articles	1		Done								1						
19 – Graph visualization of sources and articles	3		Done										1	1		1	
35 –RSS Feeds	3		Done						1		1			1			
mproving UI elements	1		Done				0.5	0.5	1		1						
Database Maintenance and Testing	1		Done	0.5	1		0.5				0.5	0.5		1			
			Done														
			Done														
			Burnout Per Day	0.5	1.5	1.5	1	0.5	2	0	4.5	1.5	3	5	3	4.5	0
			Actual Burndown	**27.5**	**26**	**24.5**	**23.5**	**23**	**21**	**21**	**16.5**	**15**	**12**	**7**	**4**	**-0.5**	**-0.5**
			Expected Burnout Per Day	2	2	2	2	2	2	2	2	2	2	2	2	2	2
			Expected Burndown	26	**24**	**22**	**20**	**18**	**16**	**14**	**12**	**10**	**8**	**6**	**4**	**2**	**0**


Sprint Burndown
Actual Burndown
Expected Burndown
30
20
10
0
-10
Number of Units of Work Left
Day - November 1st to 14th

Verification

This was done by unit testing, and code reviews. The former are included with our Phase 3 submission. The focus was primarily on unit testing for each of database functions, web functions – reference searching, RSS and visualizations (at least the data generators). See the unit tests in /src.

The code reviews are listed below. We used a checklist based inspection, taking care to address the elements listed below. See /course/code_reviews