

6-15-2016

## Founders Keepers Newsletter 2016-06-15

NIU Libraries

Follow this and additional works at: <https://huskiecommons.lib.niu.edu/ua-library>

---

### Recommended Citation

NIU Libraries, "Founders Keepers Newsletter 2016-06-15" (2016). *University Libraries Archive*. 52.  
<https://huskiecommons.lib.niu.edu/ua-library/52>

This Newsletter is brought to you for free and open access by the Teaching & Learning Archives at Huskie Commons. It has been accepted for inclusion in University Libraries Archive by an authorized administrator of Huskie Commons. For more information, please contact [jschumacher@niu.edu](mailto:jschumacher@niu.edu).



# Founders Keepers

## Upcoming Events

- June 23-28: ALA Annual Conference
- July 4: Independence Day (Libraries closed)
- August 7: Summer Session ends
- August 22: Fall classes begin

## and the winner is...Reflection Room!

You were asked to suggest and vote on a name for the quiet room we designated on the 1st floor. "Reflection Room" won, but there was considerable participation, and the votes were spread over eight choices. Thanks to all who took the time to respond to our initial call for names and, then, vote on your fa-

vorite. We'll put a sign on the room, but it is already open for use. It is furnished with four soft chairs and a small table and lamp. We have chosen some photos from Regional History that will be printed and mounted for display. The door will remain open, and you are all welcome to stop in and use the space. We will be pub-

licizing the availability of the Reflection Room. If you are asked, please remember that it is for individual use (not group study) and cannot be reserved. Students are welcome to bring in a prayer rug or reading material.

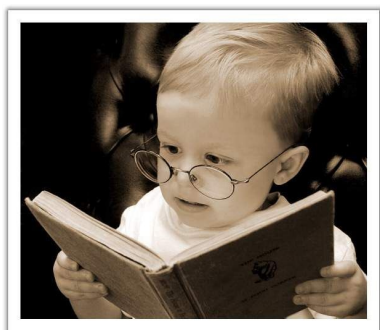
Thanks again for your participation!

## Where have the READ posters gone???

Well, the READ posters are still on the 3rd floor in the Leisure Reading area, but we haven't had a new one for quite some time. Why not? No nominations! Anyone can nominate another employee (they don't have to be students) for this recogni-

tion. After the poster has hung in the Leisure Reading area for a month or so, it is given to the employee. Some want to keep it, and some want to give it to a relative or friend. Who would you like to nominate? Ask them if they would be willing to

have their photo made into a poster, and then send me their name. The book with which they pose is entirely up to them, although we can provide some suggestions if they wish. Send names to [rcordell@niu.edu](mailto:rcordell@niu.edu) any time!



## Getting to Know...Tamara Boston



Tamara taking a selfie.

I have been working in the library since July 2012. This is my fourth and final year in the library. I started in Gov. Pubs. and will finish in library security. The thing that I enjoy most about my job as security is that I have the opportunity to interact with the diverse student population. I delight in being able to greet visitors to the university, to assist students when they are perplexed about where to locate items in the library, and talking to my former and current professors.

When I'm not in the library as an employee, I'm here as a student a lot during my off time (LOL). Seriously, I enjoy reading books and articles on feminism and feminist pedagogical

techniques and watching HGTV shows and Frazier on Netflix. Currently, I am reading *Discourse: The New Critical Idiom* to help prepare me for doctoral studies. I'm listening to Vivaldi's Four Seasons: Spring 3 which was recomposed by Max Richter. I love the string articulation in this piece.

The last good movie that I watched was *The Women's List*. It was empowering and encouraging movie.

I have pet guinea pig named Honey Bunny. If I won the lottery today, I would go into hiding then resurface in time for graduation and travel for the rest of the summer. I would write checks to various NIU departments (Political Science, Library, Center

for Black Studies, and College of Education) for necessary updates and scholarships.

There are three places that I would like to visit in the near future:

France, UAE Dubai, and Washington DC.

My favorite restaurant in the area is Pizza Pro's.

The have the best wings ever.

People don't know that I'm actually kind of shy. Since these are my last months in the library, I would like to thank you all for the kindness given to me throughout my tenure in the library. I would send a special thank you to Mark, Debbie, Rachel in Gov. Pubs. and Margaret in Regional History (she hired

me 😊 ) and to Joel and Bill in Security. I'll miss you all.

## Upcoming Issues of *Founders Keepers*

The July issue of *Founders Keepers* will be late in arriving in your mailbox. The deadline for submissions will be July 19, and you can expect that issue to go out around July 22.

The August issue of *Founders Keepers* should be on time, with

a deadline for submissions of August 12 for the August 15 publication.

Thank you for your patience!



## Text-mining, by Drew VandeCreek

During the spring semester of 2016 I supervised a team of students (Marcos Quezada, a graduate student in Operations Management and Information Systems; Fredrik Stark, a PhD candidate in NIU's English Department, and Mitchell Zaretsky, a junior Computer Science major) as they explored text-mining in the context of Northern Illinois University Libraries' large online collection of late nineteenth and early twentieth century dime novels (<http://dimenovels.lib.niu.edu>). We worked in the format of an experiential learning activity, meaning that we addressed a problem brought to us by a client. In this case Matthew Short, NIU Libraries Metadata Librarian and Cataloger, served as the client.

In the experiential learning format, the client presents the student team with a set of goals. Mr. Short asked the team to develop a text classification application or tool to help library catalogers to determine the genre of the approximately 1,900 digitized texts in the collection. In traditional cataloging activities, the cataloger inspects a work manually in order to derive basic information necessary to catalog it accu-

rately. This can be a lengthy process. Perhaps text-mining technology could help catalogers to improve the speed and efficiency with which they catalog a very large collection.

Mr. Short's goals also included the compilation of a list of genres and related subject terms for possible use in reclassifying online digitized collections; investigating text-mining tools for the future development of the prototype classifier application and future studies of the collections.

The team began work by using Weka, an open-source data and text-mining application. Mr. Short selected it because it enables users to acquaint themselves with the separate activities that make up text-mining and construct original applications using blocks of existing Java code.

Mr. Short introduced the students to a typical text-mining work flow. He had been working to achieve his goals prior to engaging with this group, and for all intents and purposes led the team's activities. As the team's official coach, I attempted to facilitate discussion, scheduled activities, and complet-

ed paperwork.

The students began by gathering text files of digitized dime novels cataloged as belonging to the collection's better-represented genres. These genres included detective and mystery stories; western stories; sea stories; historical fiction; adventure stories; and bildungsromans (coming of age) stories.

The team next engaged in pre-processing activities in order to produce the most accurate text possible. NIU Libraries staff members originally produced the digital texts in the digital dime novel collection by the use of Optical Character Recognition software and did not attempt to correct any mistakes within them. Pre-processing began with the removal of stop words (such as the, an, and, etc.) and also included tokenization (identifying groups of characters as words) and stemming (reducing different inflections of a word to their root form) of words. We also used Weka to render the text materials as a bag of words (i.e., set aside grammar and word order) and transform words into vectors, or numerical representations.

The team then moved on to text classification. They began by using a set of already-cataloged works to train Weka to identify specific words or sets of words with the individual genres mentioned above. Of the algorithms available in Weka, Naive Bayes proved most effective. They found that in 65% of works examined, Weka's classification agreed with that of a human cataloger. Investigating this discrepancy, the team found that the use of additional filtering techniques, including the use of TF-IDF (a process to determine how important a word is to a document in a collection or corpus); a better stemmer (the open-source product Snowball); a list of nineteenth-century stop words composed by Matthew Jockers, a scholar of the period's literature; rendering all letters in lower-case; and setting the number of words in each text to be analyzed to 500 improved accuracy, i.e., Weka agreeing with a human cataloger's genre classification, to 75%. They also discovered that a number of texts in the training set had been cataloged as belonging in two different genres. Removal of these works improved accuracy to 83%.

*(Continued on p. 4)*

## Text Mining, continued

With the information above, Mitchell Zaretsky used Weka's Java API to construct an original classifier application. It reported the probability of a work fitting in one of the several genres. Working with a new test corpus of 214 digitized dime novels, the team found that their classifi-

er agreed with human catalogers 71% of the time.

On the basis of this test, the team determined that their application can help catalogers to determine a dime novel's genre. It can also serve as an effective tool for evaluating the

genre determinations of catalogers not using the application in their work. They also suggested that text-mining activities uncovered details about the form and content of works in NIU's digitized dime novel collection that invite further research.

## The Last Word

I'll be leaving for ALA next week, with several days of meetings and programs to attend. Annual is always a busy conference with much to do and see: many more programs I would like to attend than I can, meetings that conflict, and more vendors to visit than at Midwinter. One of my committees is the ALA Committee on Professional Ethics (COPE), which has hit on a popular format for its programs: skits, showing ethical dilemmas that library employees face, with discussions involving the audience and eliciting their suggestions for handling the situations. Audiences probably find this type of program a welcome relief from the usual "talking heads" program, but it is also a good way to delve into the complexities of ethical problems in librarianship since the audience members would usually

present several different viewpoints in the discussion. This year, the committee has prepared skits on serving transgender youth, a patron being disrespectful to a Muslim librarian, and a request to use a library meeting room for a meeting that is for African-Americans only. Some ALA attendees may think the committee is attempting to capitalize on the recent debate on North Carolina's "Bathroom Bill" legislation and the shootings in Orlando, but these skit ideas were developed months ago in response to more mundane issues that libraries have faced. In truth, what we deal with in libraries is always a reflection of what is happening in society; we are not isolated from the culture around us and are affected by societal shifts and movements. The difference in libraries is the emphasis we give to

educating oneself about issues. As James Madison said:

A popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or, perhaps both. Knowledge will forever govern ignorance: And a people who mean to be their own Governors, must arm themselves with the power which knowledge gives.

This quote is often used in defense of intellectual freedom, and it is one of my personal favorites. However, Madison was probably talking about the importance of education, not specifically the First Amendment. That understanding makes bringing facts into public debate a long-held good idea... certainly one ALA can support!

