Undergraduate Research Support with Optical Character Recognition Apps

Jim Hahn, University of Illinois at Urbana-Champaign
jimhahn@illinois.edu

Acknowledgements

The author acknowledges the Campus Research Board of the University of Illinois at Urbana-Champaign and the University of Illinois Library Research and Publications Committee, which provided support for the completion of this research. Many thanks to Chris Diaz, Residency Librarian, Scholarly Communications and Collections, University of Iowa, for help with participant recruitment, observation and interviewing support in the user studies; Mayur Sadavarte, Graduate Student in Computer Science at the University of Illinois and Nate Ryckman, Graduate Student in Information Systems Management at Carnegie Mellon University for Optical Character recognition programming support; Yinan Zhang, PhD Candidate in Computer Science at the University of Illinois; Sherry (Mengxue) Zheng, Graduate Student in Computer Science, for help developing the search and suggestion functionality of the Deneb near-semantic index; Maria Lux, Graphic Designer, for laying out the polished recommendations and prototyping Text-shot integration as a Minerva module.

Introduction

Mobile applications are nearing a maturation stage as a technology that undergraduates use and rely on for everyday tasks. This near-maturation is underscored compellingly in the sheer number of mobile applications available; consider the datum that total installs of mobile applications (or apps) now number in the billions\(^1\). Total apps that are available for quick and often free access to information are nearly a million strong\(^2\). These numbers are indicative of a profound shift since the development of application phones.

This paper reports findings from formative user testing results of a library mobile application and its related functionalities for supporting first-year students as they make the transition to research at the university level. The Text-shot module


\(^{2}\) There are over a million apps on Google Play: [https://play.google.com/about/features/](https://play.google.com/about/features/) (accessed 13 December 2013)
reported here is designed for integration as a component within the Minrva mobile app (Hahn & Ryckman, 2012a). Text-shot is one of the first Minrva modules to be developed using optical character recognition (OCR). One of the advantages of designing modularly is that component functionalities can be developed independently and dependencies in the program are isolated. This contributes to a more stable and functional mobile application.

Minrva (https://play.google.com/store/apps/details?id=edu.illinois.ugl.minrva, accessed 13 December 2013) is designed modularly in order to protect the app from future obsolescence. By choosing to design modularly the following advantages are gained:

- The Minrva app grows in usefulness over time as additional modules are developed and integrated.
- Modular design is the best design approach when employing student programmer talent; programmers can work on components of the app independently and complete tasks in short time frames.
- Modules are portable to other library systems; the Minrva 2.0 release available on the Android Play store offers consortia location select on first load, which then displays the available modules for that campus location.


Text-shot is a prototype that will become a Minrva module. The prototype is uses OCR software and a backend search system for subject and title recommendations. The subject recommendations are from a near-semantic subject suggestion index (Hahn & Diaz, 2013), which includes data derived from Library of Congress Subject Headings. The title suggestions are from the same index. The choice to recommend library content to users from the app stems from the objective to connect students with library resources, and to help students integrate library resources into their work.

As a software development best practice, development for the Minrva modules involves gathering student use preferences for mobile components early in the design phase through both user testing and interface modification based on these tests. This is referred to as formative user testing. This paper details the evaluation of the prototype with a set of users who are in their first year of study at the University of Illinois (Illinois). With this original data, Minrva developers are able to shape a final interface (see images 5 and 6) with desired functionality that researchers know will be useful to undergraduate students in their first year of study.

There are a variety of uses for OCR in library settings. Most often in libraries, text recognition is employed for scanning the content of books using high quality
scanning tools that are made for this purpose. However, mobile technology now ubiquitously bundled with a digital camera has created the opportunity for image capture as well. With image capture from a phone, and the accompanying software development resources available, it is possible to programatically pull out text strings from images and use these text strings for searching in library resources such as the online catalog (Hahn & Ryckman 2012b). The software tool from which the OCR identification in this study is based is the OpenCV programmer tool kit, which is available as an open source computer vision software package that can be integrated into Android applications. More information on the OpenCV toolkit is available here: http://opencv.org/platforms/android.html (accessed 13 December 2013).

This article details how students used the prototype to scan the content of their assignment sheets, their syllabus, or other text from courses and then get suggested resources from the app based on the scanned content. As an example see the scanned text image below as well as the result book and subject suggestions, which are part of the prototype Text-shot module.

Image 1 is an example of the optical character recognition screen as it appears on the phone when the OCR app recognizes a string of letters.
**Image 2** is a prototype example of subject suggestions based on the captured string of letters. Subject suggestions are from the Illinois near-semantic search index available at: [http://dunatis.grainger.uiuc.edu/deneb-2](http://dunatis.grainger.uiuc.edu/deneb-2) (accessed 13 December 2013).

**Image 3** is an additional prototype of book suggestions from the Illinois near-semantic search index.
The paper progresses next with a literature review on first-year student support as it relates to digital resources and a review of mobile computing literature in educational settings. Methodology of the study follows, and the paper then presents an analysis of results and a discussion of next steps in development. The report ends with a conclusion analyzing the broader implications of OCR software in library settings for both students completing assignments and librarians involved in reference and instruction.

**Literature Review**

The literature review that follows provides digital learning and mobile computing background that contextualizes motivations for the current study. Examples of mobile app uses by first-year students, and a review of other optical character recognition software available through mobile app stores are discussed. The review concludes with a survey of available application programming interface (APIs) tools that developers may find useful for designing optical character recognition services in library settings.

The 2013 Horizon Report (http://www.nmc.org/publications/2013-horizon-report-higher-ed, accessed 13 December 2013) a resource for higher education technologists, includes a list of significant challenges facing higher education in the near term and long term horizon. One of the challenges noted for 2013 includes – “the demand for personalized learning is not adequately supported by current technology or practices,” going on to note that “the increasing demand for education that is customized to each student’s unique needs is driving the development of new technologies that provide more learner choice and control and allow for differentiated instruction,” (Johnson et al., 2013, p.10). The development of the OCR mobile app in this study seeks to meet these challenges with assignment specific library resources bundled with functionality that will help to meet student needs in the research process.

The Pew Research Center Internet & American Life Project reported “56% of American Adults now own a smartphone of some kind; Android and iPhone owners account for half of the cell phone user population” (Smith, 2013, p1). The report goes on the mention that – “fully half – 49% -- of cell owners with a household income of $150,00 or more say that their phone is an iPhone. And African-American cell owners are more likely than whites or Latinos to say that their phone is an Android device as opposed to an iPhone” (Smith, 2013 p6). These statistics underscore the motivation for development of mobile apps in library settings. They also point out that outreach to traditionally underserved populations may be possible by choosing one type of mobile operating system (Android) over others. The prototype developed in this study is coded for Android, initially – but researchers anticipate developing an iOS version in the future.

The EDUCAUSE Center for Applied Research report *Student Preferences For Mobile App Usage* -- “The survey asked students to report the relative amounts of time they
spent using mobile apps versus using a smartphone browser. Overall, students reported spending more time using mobile apps, and as students become more advanced in their use of smartphones, the gap widens—the amount of time spent using mobile apps increases, while the amount of time spent using a smartphone browser remains relatively consistent.” (Bowen & Pistilli, 2012 p .5). Bowen and Pistilli also report “The survey asked students to evaluate the ease of use and speed of accessing information on their smartphones. The largest percentage of students indicated that mobile apps are both faster (68%) and easier to use (70%) when compared to accessing information via the browser” (2012, p.8). These data provide additional evidence for the development of mobile applications in order to reach students. Specifically, mobile applications will help academic units like libraries integrate into the preferred information environments of undergraduate students.

First-year learning support

In a research study on iPad 2 uses by a first-year undergraduate learning community (Hahn & Bussell, 2012) found that mobile computing has multiple curricular uses at the university level. Specifically, researchers found that students’ value the capability of Internet searching and accessing course specific content through campus wireless networks during lecture or small group course sessions. In meeting other first-year student needs, there are also examples of mobile computing replacing other class-based technologies with personal mobile tools that students bring to class, like clickers.

In the case study on using mobile phones as a replacement for clicker software, research indicates that using students’ phones may engage student inquiry (Burkhardt & Cohen, p194) at a higher degree than standard clickers in classrooms. The authors found that “ease of use paired with its dynamic interactivity makes integrating this technology into the classroom fun for both students and librarians” (Burkhardt & Cohen, p200.) As mobile technologies relates to first-year undergraduate students, the authors also note “today’s college freshman are a generation that communicates primarily through their mobile phones, more specifically through use of text messaging. According to the Pew Center, 77% of 17 year olds talk with their friends by text daily” (Burkhardt & Cohen, 192).

Optical character recognition apps

In addition to the prototype developed for this study, there exist OCR applications available for download from popular app stores like iTunes app store and the Google Play store. A few of the most visible and popular of these software programs are reviewed here. These include language learning and translation apps like the Wordlens app (http://questvisual.com/us/, accessed 13 December 2013), which can translate words from different languages using a digital camera feed. The free version shows functionality for optical character recognition, the Wordlens app may be useful for academic settings when students travel abroad, or in completing first and second year required language courses at the undergraduate level.
Another mobile search by camera image app includes the *Google Goggles app* available from Google: (https://play.google.com/store/apps/details?id=com.google.android.apps.unveil, accessed 13 December 2013) – while the app supports OCR scanning – the user of this app could also take a picture of a book cover, or other artifact – like a painting, which the app then runs a Google search on the identified object. This is useful especially in shopping or commerce, but its uses in academic settings may be limited due to the limitations of solely searching the Google search index that may not directly guide students into the curricular content of peer-reviewed articles to which the library provides access.

*Camscanner* is a free mobile app that allows library patrons to digitize their print documents with the camera on their mobile device. After taking a picture of the document the *Camscanner* platform supports a number of storage options and features. These include the ability to add notes and tags to scanned documents, as well as functionality to share scanned documents with collaborators. With cloud-based infrastructure the *Camscanner* platforms supports the ability to view any content on any device – from tablet to smartphone – any edits are uniform across platforms (http://www.camscanner.net/user/guide#guide1, accessed 13 December 2013).

**Optical Character Recognition APIs**

APIs are an additional approach that could be utilized in development of optical character recognition services in library settings. The *Evernote app* (http://evernote.com/, accessed 13 December 2013) includes OCR functionality. *Evernote* also makes available an API for development of OCR tools and services: http://dev.evernote.com/doc/ (accessed 13 December 2013). *Evernote* adds a set of functionalities that would connect to a number of resources that the app provides. These resources include integrating with *Evernote’s* organizing and free text search tools; this would be of great use for library patrons who are already using the tool for their day-to-day note keeping and organizational needs.

*Google Drive* is the cloud based document platform from Google. It supports conversion of images with text into documents of text. More information about this capability is available from the drive support documents: https://support.google.com/drive/answer/176692?hl=en (accessed 13 December 2013). *Google Drive* offers an API that developers may be interested in utilizing, since the processing of Google’s cloud may be of use in any service development setting.

For libraries with systems departments that have capacity and staff to develop OCR apps on their own, the Vuforia 2.6 SDK release includes OCR functionality. System
librarians can download the developer kit and related set up files from the Vuforia developers’ page here: [http://developer.vuforia.com/resources/sdk/android](http://developer.vuforia.com/resources/sdk/android) (accessed 13 December 2013). The technical development of the Text-shot OCR project is still underway, but at the time of the study (Spring 2013) Vuforia had not yet released their OCR tool leading the Illinois development efforts to create their own character recognition solution leveraging the OpenCV software library. The research and development of the homegrown scanning and recommendation solution will result in a code base that researchers at Illinois intend on releasing into open source. Interested developers can monitor our source code from the Minerva Developers Source Code page: [http://minrvaproject.org/source.php](http://minrvaproject.org/source.php) (accessed 13 December 2013).

Based on the review of technologies outside of libraries it should be made clear that there are myriad use cases and services that could be developed outside of library environments. Since this research takes place in a library setting on a University campus, with a grant stream of funding specifically targeted to support first year student research needs, the initial scope of development is narrowed to a few test cases of scanning an assignment sheet, a course syllabus, or essay prompt and connecting this with library data. The next section details the methodology for formative evaluation of these use cases.

**Methodology**

Since the specific use case of scanning a course related sheet of text with a smartphone is novel and there are many areas of possible development inside libraries the methodology used here is one of formative evaluation. Formative evaluation\(^3\) starts with a small set of test participants to gather feedback early in the design phase so that the software development process can progress in a direction that will support user requirements for the software. Before the OCR module is integrated into the set of Minerva app modules, researchers tested the stand-alone functionality of scanning an assignment sheet and obtaining suggestions based on image scan. This section is an overview of the test participants, the study process, and the OCR functionality tested.

**Test participants**
The student test participants were recruited from the General Studies 101 (GS101) course. The students of GS101 are in their first year of study at the university and have not yet chosen a major. This presents the study with a cohort of students who are novice searchers of library databases at the university level, and do not yet possess discipline specific research needs.

---

\(^3\) Formative evaluation is a method of rapid prototyping whereby data are gathered from test users early in the design process such that the software is shaped by user preferences and needs throughout the development cycle (Jones & Richey, 2000).
During the Spring 2013 school term student interviewing and observation took place. Interview and observation took place in the Undergraduate Library at Illinois. These were brief assessments lasting no longer than thirty minutes each.

There were a total of five test participants for this first round of study. According to usability expert Jakob Nielsen, a majority of usability issues within a given tool can be uncovered after small numbers of participants complete the study (Nielsen, 1993). While the data are not statistically significant in a quantitative frame, the resulting data from these participants is a set of deep qualitative feedback that allowed researchers to improve both the functionality and layout of the application, as our Results section shows.

**Study Process**

Students were given an Android mobile computing device, with the Text-shot app pre-loaded. Investigators observed the students as they used the OCR mobile software to obtain suggested library resources. Investigators collected two sources of data: observations of how students interact with the software and a debriefing interview. Investigators recorded student reactions to the software application (see appendix for investigator logs).

**Functionality Tested**

Researchers tested the two main functions for the app, which include recognizing a string of text by taking a picture of the word in a student assignment sheet and then suggesting subjects and titles based on the scanned text. The Deneb index (http://dunatis.grainger.uiuc.edu/deneb-2/, accessed 13 December 2013) was used for providing the suggestion data of subject and title suggestions. While the Deneb service is experimental the title and subject suggestions are available as an internal Web API. Web APIs act to extend data (in this case subject suggestions from the online catalog) to multiple programs and views and are a key part of any library prototyping pipeline. Examples of currently available Minerva APIs can be found here: http://minrvaproject.org/services.php (accessed 13 December 2013).

**Results**

This section is organized around three areas of user feedback: a) areas to improve functionality of the optical character recognition software, b) suggestions of library content (book titles and subject areas) provided by the mobile app, and c) additional feature requests for functionality of Text-shot application.

**Improvements to Character Recognition**

Observing student users of the Text-shot module, researchers learned focus for the camera can take 3-4 seconds to complete its autofocus once the students aim it at words – the test participants inquired if the app can get a clearer picture of the words quicker than this. The app uses a file of training characters to identify
characters. Some of this training data looks similar to the pattern matching software; “I” and “L” which look similar in training data, and are identified incorrectly in the tests (as with O’s and Zero). Additionally, OCR training data for some letters, such as “C” looks similar in both lower and upper case, such that the lower case “c” is recognized as an upper case “C” in the text string recognition.

Image 4 – An example of problem characters in the Text-shot scan of a word

Using a continuous scan in the OCR toolkit or using a Google API that can correct the word before it sends that string to the subject suggestion may alleviate these OCR recognition problems. As currently designed, recognition software does not adequately accommodate spacing between words so the strings of text become one long string of letters rather than separate words. By using student’s assignment sheets developed by professors the author found that smaller text is difficult to capture, but also learned that small text should be used when training sets are made since professors are more likely to make smaller font sized assignment sheets to save money on paper costs.

The author found also that the optical character recognition software would benefit from error correction. Currently the target box does not exclude cutoffs e.g. if a partial letter gets scanned that isn’t a part of any word, just exclude that extra portion.

Overall, students are looking to combine multiple word queries. Students also expected the application to “understand” the sentences of assignment pages. While understanding of the semantic meaning may be possible in future releases by leveraging some of the APIs discussed in the literature review, the OCR software itself is simply a pattern-matching program. The “understanding” of text is an area
of research that can be addressed with the index that the characters as words are sent to as a search query.

Themes Related to App Suggestions

After scanning the assignment sheet, the app would display suggestions of library titles and subjects in the library catalog that are relevant to the text string that was scanned. This section details themes related to improvements in both subject and title suggestions (Image 2 and Image 3, respectively):

The common threads researchers observed included students desiring to see the subject suggestions in a more streamlined, easier to understand way, e.g. first show broad subjects and then expand to detailed subjects. The desired subject expansion that students request is similar to the hierarchical broad to narrow facets that exist in modern library search databases. While faceting in mobile app interfaces does not look the same as a standard desktop browser, the expansion could be accomplished with the design of rows of content that detail a broad area and when tapped, the row opens and expands to detailed subjects related to the broader subject query. As this interface was initially modeled the author found that it contained too much text for student comprehension and ease of use.

Related to the subject suggestions, the author found that students questioned what the subjects “do” exactly, they were unsure if these were designed to provide more searching or for some other purpose. In the next iteration of the Text-shot module, the author plans to make the subject suggestion uses more intuitive. To make these more useful they should look like other links in the app that can be clicked. This would achieve the consist design users would expect in any user interface. An alternative would be to use a “tool-tip” hover that could notify a first time user that the subject suggestion text is selectable and will take them back into a catalog search if it is tapped.

In the study there was the option to view recommended titles based on the image scan. This button was not prominently displayed. The author plans to include noticeable search result features from a single view – this is modeled in the Discussion section (Image 6).

Feature Requests

Researchers found that students wanted suggestions for websites, article searches, relevant databases, and book results. Test participants also inquired if they could either type in correct words or additional words to the scanned textual content. Students noted that they didn’t have options to select search parameters. The parameters that they requested included the option to select if the scanned text is a title, subject, or author. As it was initially designed, the Text-shot module was solely a keyword search, which may be too basic for undergraduate research needs. Parameter based searching is being built out in the next version of the Minrva app.
As a final testament to the possible integration into student research workflow, a student commented “the app would be convenient if you don’t have a laptop, if you want to be independent about using the library.”

Discussion and Conclusion

This section details changes made to the interface, unpacks curricular connections the software makes possible by connecting with university assignments, noting the implications of optical character recognition searching. The paper concludes by identifying project next steps.

Based on research findings in the Results section a new mobile interface to display search recommendations was developed.
The first interface a user will see when they tap on the 'Text-shot module will be a simple to understand scanning interface (Image 5). In order to meet students’ desires for a quicker image focus, the development team added to this interface a panning red bar that serves to let the students know that the camera attached to their phone is focusing on the words. There are lighter red brackets at each of the
four corners in the target zone that will provide the students with the freedom to
narrow in on the text that they deem most important to their search.

Image 6 is the reworked search results screen that a student will see after the
software identifies a string of text (image 6). At the top of the interface a white box
will display the identified character strings. In this example the student has chosen
to highlight "Practical Research methods for librarians and information
professionals."

The reworked level two interface offers search results for articles, book titles, and
library web content like LibGuides; it is not unlike a search layer powered entirely
by web APIs. The new interface would have three rows to indicate the type of
content available and subject results that are applicable and searchable in the
catalog module of Minrva. The rows offer an intuitive layout such that users could
quickly and conveniently make use of the suggestions and federated results of
Library Guides, Article/Book Searchers, and relevant Subject areas to the
assignment or scanned text.

---

4 Described and tested compellingly in (Rochkind 2013), APIs can provided data to website views
(e.g. traditional webpages) or a mobile app view; mobile apps are essentially an alternative
presentation layer.
**Image 6** – New Text-shot recommendation interface using streamlined view for showing users available Library guides, Library content, and Subject suggestions.

The Text-shot app described here offers increased opportunities for supporting the research needs of first year students. Undergraduate students often desire self-sufficiency in their research process, and can at times lack the adequate support needed in order to complete course based assignments requiring library resources. An app like the Text-shot module may help to meet this desire, while also
supporting the goals of a library in an academic setting, which exist to support the curriculum with the vast resources of digital content available – resources which first year students in the GS101 course may not be aware.

Library content can be made more relevant and timely when connected to a students’ phone in a manner that is instantaneous and seamless to student research needs and processes. The connection provides students with a closer integration into their information environment, namely mobile devices that are ubiquitously used by new students. Reference specialization and instructional support would still be required, even when options like the Text-shot module are available.

Students’ requirements for in-person consultation are not likely to diminish. Indeed library reference services can help students explore their topics in a more responsive way than automated recommendations may be poised to support. At the current state of development, a Text-shot module starts the student on the road to information literacy, laying foundations for important resources to consult; however the Text-shot module cannot fully replace the advice of a reference or instruction librarian who may be guiding the student in the evaluation of their search results, or recommending resources that may be outside the recommendation models.

Connecting Library Resources

A national study by the Project Information Literacy group (http://projectinfolit.org/, accessed 13 December 2013) on research assignments, the Assigning Inquiry: How Handouts for Research Assignments Guide Todays College Students report found:

“The majority of handouts... placed more attention on the mechanics of preparing a research assignment than on conveying substantive information that students also needed, such as how to define and focus a research strategy within the complex information landscape that most college students inhabit today. Moreover, a large number of handouts in the sample provided only limited guidance about how and where to conduct research and find information. The handouts had few specific details about finding and using sources, making the guidance that was provided often vague and inapplicable” (Head & Eisenberg, p3 2010).

The implications for OCR apps are clear. Libraries require tools and services that can help to support course research assignments. These tools are necessary since students do not have adequate direction or instruction from the course assigned handouts alone. While OCR apps are not the only way to meet this challenge they do represent a compelling software solution that has previously not existed in libraries, and is novel for the new features and services which could be provided to support first year students, especially as they go about completing research papers in their critical first years of study.
Project Next Steps

The author envisions the following next steps for this project - first, to open source the optical character recognition software with recommender tools so that other libraries may repurpose this for use in their own mobile apps. Library technologists are welcome to reuse the Text-shot software as a module in their own mobile applications. It may be the case that new devices like the Google Glass hardware could additionally make use of OCR software for developing library apps for alternative educational uses. These may include literacy support or reading comprehension.

Second, researchers are planning ORC investigation with call number scanning on library books for print collection exploration. A new module for taking a text shot of a book call number is in the research and development stages with a planned use study in spring 2014 semester. With this new initiative the author will be able to study services in the library at point of need in the book stacks, and help to integrate digital content and other library resources into the users’ environment. Lastly, future research and development will address the challenge of integrating library resources into student research assignments in the era of digital information access that is inclusive and anchored in print collections.

References


Appendix

Mobile app usability – Interview questions

How easy is the application to use? What would make it easier to use?
What was hard to do with the application?
What was confusing?
What was surprising?
What do you wish you could have done with the application while you were using it?
How useful do you find the application? What would make it more useful?
Would you recommend it to friends?
What do you actually want from a library app? Is there something else that should be here that is not here?
Is this application a worthwhile tool for the library to develop?

Mobile app usability - Investigator Log (observations)

Please describe any previous experience finding items in the Library.
How easy to use is the application?
Does the student need time to learn how to use the software?
What unexpected things occur?
How do students react when the application does not work as they expect?
Do students make use of the recommendation features?

Note any additional observations of student use of the software: