Abstract

Learning how to read kuzushiji is the most important skill for studying the history of pre-modern Japan and classical Japanese literature. However, the difficulty of acquiring this skill has prevented both scholars and amateurs from other fields and countries to work on classical Japanese writings. The kuzushiji project is an attempt to provide accessible resources and training on mobile devices for learning kuzushiji, and available for free. KuLA, the learning app we developed, has already been downloaded more than 36,000 times since its release in February 2016. In this paper, we describe our background, aims, and approach of our project, as well as the implementation of KuLA.

Introduction

The 20th century saw drastic changes in writing systems in many East-Asian countries; Mongolia, for instance, introduced the Cyrillic script in 1941 to replace its traditional script under the political pressure of the Soviet Union. Vietnam switched its public writing system to a Latin-based one from its traditional Chinese-based writing system, Chữ nôm. Malaysia also introduced a Latin-based script (Rumi) at the beginning of the 20th century. Since the skill of reading historical writings is directly linked to familiarity with past writing systems, these countries must equally suffer from difficulties in handing down past knowledge written in historical resources to their younger generations. In the case of Japan, kuzushiji, a general term for classical calligraphic renderings of Japanese characters which used to be common for both publishing and handwriting before the end of 19th century, has become an obstacle to accessing knowledge recorded in the past.

The current Japanese writing system consists of two types of characters: 46 syllabic kana and thousands of logographic kanji, which are adapted Chinese characters. The modern system has been developed since the Meiji Restoration of 1868, and in this process Japanese society obsoleted some types of characters and writing styles that used to be common before the Meiji period. First, pre-modern Japanese had the option to use different characters other than the current kana to denote a single syllable. For instance, a kana “ず” (pronounced "su") had more than five alternative characters (see Figure 1). These variants of kana are called hentaigana, which were officially obsoleted by the amended education law announced in 1900. Second, in most publications and writings before the Meiji era, kanji had been written mostly in cursive forms that appear today only in calligraphy (See Figure 2). As letterpress printing had replaced wood-block printing after the Meiji period, this cursive style of writing gradually declined. Kuzushiji is a general term (including hentaigana and cursive kanji) for characters that are not used today.
Through the process of modernization in the 20th century, modern Japanese (except trained experts) have lost the ability to read *kuzushiji*. A scholar of Japanese classical literature Mitsutoshi Nakano estimates in his book that even in a humanities faculty less than ten percent of the faculty members can read *kuzushiji* [Nakano 2011]. Robert Campbell, a scholar of Japanese classical literature originally from the US, summarizes this situation: “We read characters in order to learn about Japan in the pre-modern era. But ironically, Japanese lost the skill to read these characters in exchange for the modernization they have achieved for the last 150 years. Those who rely on printed text can learn only small portions of the past Japan. Contrary to European and American countries, Japanese have lost the abilities to go back freely to their own historical landscapes.” [http://www.toppan.co.jp/news/2015/07/newsrelease150703_2.html](http://www.toppan.co.jp/news/2015/07/newsrelease150703_2.html)
Campbell implies here, the skill of reading *kuzushiji* is deeply connected to the ability to access past knowledge; The *Kokusho Sōmokurokuro* (*General Catalog of National Books*) [Morisue et al. 1963], a Japanese reference work for books written in Japanese in the pre-modern era, lists more than 450,000 entries, most of which are written in *kuzushiji*. It is said that the ratio of transcribed books among these entries is less than 1%. [4] So reading *kuzushiji* is a must-have skill for studying historical writings written in Japanese.

Since the curriculum of Japanese compulsory education doesn’t include training for reading *kuzushiji*, this skill has become the exclusive ability of scholars of Japanese literature and history, who have access to the original materials with few exceptions. Scholars from these two fields have made a lot of effort to teach the skill of reading *kuzushiji* to the general public, by means of publishing textbooks and delivering face-to-face lectures; since 2010 at least nineteen textbooks for teaching *kuzushiji* have been published. However, we believe that these traditional modes of teaching have limitations in terms of availability. In fact, there are groups of people inside and outside academia with keen interests in learning *kuzushiji* and reading Japanese classical writings whose needs are not satisfied by these traditional pedagogical tools, as we point out later in detail. In order to appeal to these groups, some new methods that make use of modern digital media must be developed.

What role can digital humanities play in developing such methods? Digital Humanities has an apparent advantage that other fields of humanities do not have: direct access to both classical writings and digital technology. In other words, digital humanities has a potential ability to offer learning materials of *kuzushiji* to the general public efficiently with the help of digital media. Also, since the loss of the ability to read written historical resources in the modern era is not a phenomenon specific to Japan, building a model for teaching an obsolete or dead writing system with the help of digital media could possibly contribute to the humanities in a wider area of the world. This would make digital humanities even more important in the general field of humanities.

The *kuzushiji* project, which we will describe in this paper, is an interdisciplinary project started in 2015 through the cooperation of scholars of Japanese literature, linguistics and digital humanities. It aims to build a mobile learning application for providing accessible resources and training for reading *kuzushiji*. We call our app *KuLA* (*kuzushiji* learning application), which is already available for free both on the App Store and Google Play [Hashimoto et al. 2016]. In this paper, we will describe the target and approach of our project, the features *KuLA* provides, and the results we achieved after launching *KuLA*.

**Who are Interested in *Kuzushiji***?

As we mentioned in the previous section, there are groups who have keen interest in *kuzushiji* but not satisfied with the traditional approach of teaching. To be more precise, there is a growing need recently for the skill of reading *kuzushiji* among the following three groups:

1. Scholars of Japanese studies outside Japan, who used to rely mostly on transcribed or translated works for their research. Thanks to the recent progress of the digitization of cultural heritage materials in Japan, they now have access to a large number of digital images of pre-modern Japanese texts without having to visit Japan. Because of this change, acquiring the ability to read *kuzushiji* has become quite important to eliminating the only barrier that prevents them from using these resources and efforts are being made to expand *kuzushiji* literacy. For instance, Laura Moretti, who teaches pre-modern and early modern Japanese literature at Emmanuel College, University of Cambridge, has held a summer school on Japanese paleography for reading *kuzushiji* every year since 2014 [Moretti 2014]. We can see similar classes and workshops take places in universities from all over the world: Oxford, Paris, Berlin, Heidelberg, Rome, Chicago, Pennsylvania, Michigan etc.

2. Scholars from different fields, especially from natural sciences. It may seem odd that scientists read these classical writings, but there are some fields that need to investigate written historical resources to elucidate past phenomena that cannot be observed directly. A good example of such a field is seismology. Since the instrumental recordings of earthquakes in Japan began only after the end of the 19th century, studies of past earthquakes have to rely on written historical records to a considerable extent. For this reason some Japanese seismologists read classical writings written
with kuzushiji for their research with the help of humanities scholars. There is even an academic association dedicated to historical earthquake recordings.\[^5\] Since the great earthquake and tsunami that struck the Tōhoku area in 2011, Japanese seismologists have become more interested in reading these historical recordings.

(3) The younger generation who are interested in Japanese classical writings. In particular, there has recently been a growing number of Japanese female history buffs who have become interested in history through playing video games or watching TV dramas based on historical materials. These young females are often called reki-jo (歴女), which can literally be translated to “history girls.”\[^6\] The most popular game among them, Touken Ranbu, a web browser game developed by Nitroplus and DMM, has gained about 1.5 million users as of February 2016 (since its launch in January 2015).\[^7\] The purpose of this game is to defeat evil forces using legendary Japanese swords, which are depicted as attractive young men. It’s often reported that the players of Touken Ranbu have formed long lines to see actual Japanese swords displayed in history museums.\[^8\] Some of them are also interested in learning kuzushiji in order to read historical books on Japanese swords. In fact, when we announced the launch of our app on Twitter, we found that about 30% of the total 1750 users who retweeted our tweet were players of Touken Ranbu.\[^9\]

These three groups have different backgrounds and needs, but still have a common interest: to read kuzushiji. However, the traditional methods of teaching based on printed textbooks and face-to-face lectures cannot adequately help these groups achieve their goals because of geographical or institutional limitations. In this regard, mobile devices including smartphones and tablets, which are spread all over the world and can be used at any time according to the owner’s needs, are the best alternative media. According to a survey conducted in 2014 by the Ministry of Internal Affairs and Communications, 62.3% of Japanese own smartphones.[MIC 2014]

Previous Works and Our Approach

A number of textbooks for learning kuzushiji have been written. However, there are not so many options available in digitally. Probably the first one among them to be released is Kanaclassic developed in 1999, which demonstrates the strokes of hentaigana in video and animations [Yang 1999]. The National Institute for Japanese Language and Linguistics (NINJAL) provides the images of three volumes from the Tale of Genji as a teaching material on its website [NINJAL 2013]; The images are accompanied with their transcriptions so that users can practice how to read kuzushiji. The Hentaigana App [UCLA and Waseda 2014], which was developed through the collaboration work of UCLA and Waseda University, is the first iOS/Android app for learning hentaigana. The Hentaigana app offers a flashcards-like feature that enables users to memorize the 323 types of hentaigana it includes quickly, with beautifully designed background pictures.

Our basic standpoint is that the learning process of kuzushiji goes through multiple stages, and each stage requires different learning materials and training. Thus our approach is a more complex than those focused on a specific ability. Based on experience of teaching kuzushiji in university classes for years, we reached the conclusion that the skill of reading kuzushiji is composed of the following three elements, which are acquired in different stages of learning:

- **Basic knowledge of character shapes.** This is the basis for recognizing both hentaigana and cursive kanji. Most of the learners of kuzushiji start their learning process by memorizing a table that lists hentaigana. This ability is efficiently trained through flashcard-based exercises, as in the Hentaigana app, which takes this approach. In this regard, however, it is important to memorize characters in conjunction with other characters since the shapes of characters change in accordance with the strokes with which the handwriting is performed. Thus the teaching materials for this ability should contain word-level examples of kuzushiji as well as those of single characters.

- **A decent amount of exercises for reading actual texts.** Memorizing character shapes is, of course, not enough to be able to read kuzushiji. Learners need to integrate knowledge of each character organically in order to get the meaning of a single sentence, just like they do when they read texts written in a foreign language. Also, learners will often have to make a guess as to the meaning from the context of a phrase and they may not have clear background knowledge on the text when some part of it that they are reading
is damaged or when the handwriting in it is not clear. These abilities are developed mostly through reading actual texts.

- **Social connections with mentors and fellows.** This is the hardest thing to find when you learn kuzushiji outside the faculties of Japanese literature or history. Even if you have mastered the abilities described above, you will be very likely to encounter some characters completely illegible to you. This is the sort of moment when you need someone to ask for help, which you could find easily if you belonged to a faculty of Japanese literature or history. The previous works for learning kuzushiji don’t seem to address this difficulty, but the lack of mentors and fellows to give you help and advice can significantly delay the learning process and undermine the motivation of a learner.

On the other hand, it is not realistic to make the app cover all the books published in the pre-modern era, since characters and writing styles vary across time. Thus we limited our target to writings from the Edo period (1603-1868). This limitation will not be a large hindrance to our purpose because it is said that 90% of the classical writings that still remain today are from this period [Nakano 2011].

**Features of the KuLA**

According to the approach described in the previous section, we made our app from the following three modules:

1. **The characters module:**

   This module implements flashcard-like features for acquiring the ability to recognize the character shapes of the basic kuzushiji characters. It includes sample images and comments for 102 hentaigana and 176 kanji characters. These characters are selected according to a frequency analysis of characters we conducted on the Taikei Hombun database, a text corpus made from Japanese classical works run by the National Institute of Japanese Literature (NIJL). The module is divided into 14 lessons and each lesson is made up of 19-24 characters. Each character has several sample images both in single-character forms and in multiple-character forms (See Figure 3). The total 3,000 sample images were taken from the NIJS's digital collection of classical Japanese works. By browsing these examples the user is supposed to memorize how to read each hentaigana or kuzushiji. This module is also accompanied with a test feature to check if the user can correctly recognize the characters, where a randomly selected character image is shown and the user is required to input how to read the character in the text field below the image (See Figure 4). Then Shimimaru, the mascot of our app, will tell whether the answer given by the user is correct or not. After finishing each test, Shimimaru gives a comment on how good the user did in the session.
Figure 3. The character module. Left: single character examples for a hentaigana of "ず". Right: its word-level examples.
Figure 4. The test feature in the character module. User is supposed to make a guess on how to read the character shown above (in this case the answer is “き”).

(2) The reading module:

This module is intended for acquiring fluency of reading kuzushiji through reading actual classical texts. The module implements a reader function and currently contains the following three works as learning materials: (1) Hōjōki, a classical essay written in the 13th century by Kamo no Chōmei. The included version is taken from an 18th century publication of the text. (2) Nazo nazo Esugoroku, a sugoroku (a kind of Japanese board game) with pictures published in 1886. (3) Arami Meizukushi Kōshu, a catalogue of Japanese swords published in 1735. In order to show the text in the small displays of mobile devices, most images are divided and cropped (see Figure 5). All the learning materials are accompanied with transcriptions which are hidden by default, so that the user can check whether he or she can read the original text correctly by showing its transcriptions.
(3) The community module:

The aim of this module is to let the users communicate with each other and exchange information on learning kuzushiji. When a user encounters an illegible section while reading a classical text outside the app, she or he can take up to four photos of the section by using the device’s camera and posting a question with these photos attached (see Figure 6). Other users can browse the photos and add a short comment on the question. In order to prevent vandalism by anonymous users, accessing this module requires authentication with Facebook (authentications with Google and Twitter are also planned).
Technical Considerations

Through the development phase we have made use of open source software and cloud-based services that offer free plans with limited usages. These products and services made it possible to build KuLA with a very small cost.

One of the major challenges in building KuLA was the collection of example kuzushiji. Each use case included in KuLA contains the image data of a character and metadata such as the title of the source document, how to read the character, etc. How can one collect 3,000 images of characters with metadata efficiently from digitized historical writings? For this task we developed a simple cloud server-client system which consists of a web server and a Chrome extension (a type of plugin that runs on the Google Chrome web browser). The Chrome extension we developed enables users to crop an arbitrary region in an open tab and input metadata on the image (see Figure 7). The cropped image and metadata are automatically sent to the web server that we built with Ruby on Rails running on Heroku\(^\text{[14]}\), a cloud Platform-as-a-Service run by Salesforce.com. The web server stores the metadata into the relational database on Heroku and transfers the uploaded image to Cloudinary\(^\text{[15]}\), a cloud image storage service that provides the REST API to store and manipulate image data. The image and its metadata immediately become accessible on the Web, so that every team member can share the progress of the image collection.\(^\text{[16]}\)
Another challenging task was to have our app run both on iOS and Android; these two operation systems require developers code in two different languages: Objective-C (Swift) and Java. Building a single app in two different languages was not acceptable in terms of cost and time. So as a workaround we built our app with Ionic[17], an HTML5 mobile framework based on AngularJS[18] and Apache Cordova[19]. Ionic offers the mobile-optimized UI components that make it possible to build and test a mobile app quickly, as well as the APIs to interact with the device’s native functions such as camera and file storage. You can build an Ionic app with only HTML, CSS, and JavaScript without coding in the native languages required by both operating systems. So we could generate distributions both for iOS and Android from a single source code.

In order to implement the community features we needed a server for receiving and delivering user messages. However, the cost of building and maintaining a server application cannot be ignored for a small team like us. For this we adopted Firebase[20], a cloud-based backend service for mobile and web applications. The user messages are stored in a key-value store controlled by Firebase and are synchronized with clients in real time. Also, the image data uploaded by users are transferred to Cloudinary and delivered to other users.

**Download Statistics and User Feedback**

Since our aim in this project is to provide learning opportunities of kuzushiji to a wide range of the public, we set a target of getting at least 10,000 downloads in total. KuLA was released on both the App Store and Google Play on February 18, 2016. One of the authors announced the launch of our app on his Twitter account. [21] The reactions to the
announcement were immediate and massive; this tweet was retweeted more than 1,700 times and favorited by 2,000 users in three days after the launch. On March 16, the total number of downloads reached 10,000. Thus our initial goal was achieved within the first month after the release. As of October 10, KuLA has been downloaded a total of 36,649 times: 19,400 on the AppStore and 17,249 on Google Play. Although most downloads have been made from Japan, there has been a small but constant number of downloads from other countries: the United States (214 downloads), Korea (168 downloads), Germany (47 downloads), and Italy (44 downloads). We received a message from an Italian scholar who teaches Japanese paleology at Sapienza University of Rome that some of his students use the KuLA for their self-studies. [22]

The user reviews on our app are mostly positive so far. The average ratings of our app are 4.5/5.0 (total 43 reviews) on App Store, and 4.6/5.0 (total 93 reviews) on Google Play. Also, there are hundreds of tweets made by users about our app on Twitter. [23]

Another metric that indicates how users reacted to our app is the retention rate, the ratio of users who launched the app after installing it over time. Figure 9 shows the retention rate of KuLA during May 1-31, 2016. [24] In this period, about 2,900 users downloaded KuLA, 35.4% of them launched it one month after they downloaded it, and 9.55% are still using after three months. The retention rate after five months went down to 2.65%. Are these rates high or low? Localytics, a business analytics company that specializes in mobile markets, provided benchmarks for apps in the Business, Education and Technology category based on their analysis of more than 37,000 apps. [25] According to their benchmarks, the average monthly retention rate of apps in this category for the first three months is 32%, 21%, and 15%. Thus the retention rate of KuLA in May 1-31 is higher than the average in the first month, but lower in the next two months. This means that there is still work to be done to improve the support of users’ long-term learning.

![Figure 8. Retention rate of the users who downloaded KuLA during May 1-31.](image)

Also, some criticisms and suggestions on our approach have been also posted on Twitter. Yoshiki Takao, a historian of early modern Japan who has written several textbooks for learning kuzushiji, stated that the test feature of our app gives too much focus on the user’s ability to decide how to read a single character, which is not appropriate because there are cases where you have to decide how to read a character based on its context. [26] Some users also suggested that our app should implement a feature to allow users to trace the strokes of kuzushiji characters by moving their fingers, so that they can memorize the shapes of the characters efficiently using their fingers and eyes. Both the criticism and
suggestion we received have reasonable points. We are planning to include improvements to reflect these criticisms and suggestions in the next release.

**User Engagement in the Community Module**

As described earlier, the aim of the community module is to enable online communications and mutual support for learning among users. We shall examine how much user engagement has actually taken place for the last seven months after the release. From February to October 2016, 68 posts and 233 comments on these posts have been created by total 60 users. Out of 68 posts, 6 are just greetings but another 61 posts (91%) are questions with attached images of *kuzushiji*. And 57 (85%) have received at least one comment from other users. The average time that it takes for each post to receive a first comment is 43.4 hours. Thus users can have a high expectation that they will get a response within a few days if you post a question.

Here are some examples of actual user comments. A user named Takushi Misoi, who is trying to read a manuscript about a traditional festival that has been held in his hometown for ages, posted the pictures of some part of the manuscript that he was unable to read by himself (see Figure 9). Four other users responded to his question and gave suggestions on how to read these parts. Through the 21 messages exchanged among these users, his question has almost been resolved. Another user Hitomi Kinto posted a picture of a hanging roll on which some waka poem is written with *kuzushiji*, saying she couldn’t read this poem at all. One user soon pointed out that the poem was one first composed in *Teijiin-Uta-Awase* (亭子院歌合), a historical waka poetry contest held in 913.

![Figure 9. Left: the question posted by Takushi Misoi. Right: comments by other users on the post.](image_url)

The kinds of communications as we have envisioned are certainly taking place in the community module. However,
considering the download count of KuLA for the last seven months, we have to say that the amount of engagement by users is still small and there remains a lot to be improved in this feature. Our current hypothesis is that the lack of mechanisms to motivate users is preventing further participation from users. In the next update planned for the end of 2016, we will add some additional features to make up for this flaw: voting on comments, a leaderboard based on voting, push notifications, etc.

Conclusion

In this paper, we described the background, approach, method and result of our project to build a mobile app for learning kuzushiji. The number of downloads our app has received in a month clearly shows that there is a great deal of interest in Japanese society about its obsolete writing system: kuzushiji. A user named Megumi Kato says in her review of our app (which she posted on Google Play) that she was hoping to learn kuzushiji but didn’t know what to do at all. While traditional humanities have failed to offer people like her opportunities to learn kuzushiji, our mobile app could provide her with a starting point for learning, available for free. Also, the high ratings given by users of our app indicate that our complex approach for teaching kuzushiji meets their requirements quite well, although there is still room for improvement in our approach and implementation (as shown in the criticisms and suggestions made by users of our app). One of the major advantages of developing a learning material as a mobile app is that the developer can continue to respond to the users’ needs by continuing to update the app.

As mentioned earlier, there are many countries (especially in East Asia) whose traditional writing systems have been changed or completely replaced with other systems in the process of modernization. Although further investigation is needed, it is very likely that people in these countries share the same interests in learning their historical writing systems. Making a survey of the situations in these countries and developing applications for the support of learning their traditional writing systems may be an important task for digital humanities, which is becoming a global academic discipline.

Notes

[1] In Mongol, for instance, there was an attempt from 1991 to 1994 to reintroduce the traditional alphabet but failed in the face of popular resistance [Svantesson et al 2005, 34].

[2] Some people use the word kuzushiji referring to only cursive kanji characters. For simplicity, we use kuzushiji as an inclusive word for hentaigana and cursive kanji.

[3] Translation by the authors.

[4] This is a mere estimation since there is no reliable statistics for this topic, however.


[10] The number of characters included in KuLA may seem to be too small, since there are thousand of kinds of kanji characters used in Japanese texts. However, there is a deviation in the frequency of characters. For instance, Ugetsu Monogatari (Tales of Moonlight and Rain), one of the most important Japanese fiction of the 18th century published in 1776, contains 1996 types of characters. But 82% of the whole text can be written with only 200 characters. In order to not perplex beginners by giving too much information, we decided to select only these frequently-appearing characters.

The design of Shimimaru is taken from silverfish (shimi, in Japanese), an insect notorious for damaging old books by eating paper.

You can browse the list of examples we have collected in the following link: https://yourei-collector.herokuapp.com/characters/.

You can find tweets about our app by searching on Twitter with keywords like "くずし字" (kuzushiji) and "アプリ" (application).

The reason we chose May 1-31 is because the download count in this period is larger than in the others we recorded.

The message was posted on a Facebook group “Digital Humanities in Japan” on February 29, 2016. See https://www.facebook.com/groups/758758500904942/permalink/913511928762511/

The reason we chose May 1-31 is because the download count in this period is larger than in the others we recorded.


This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.