

Enhancing Digital Heritage Archives Using Gamified Annotations

Job King'ori Maina and Hussein Suleman

Department of Computer Science, University of Cape Town
Private Bag X3, Rondebosch, 7701, South Africa
j@kingori.co, hussein@cs.uct.ac.za
<http://www.cs.uct.ac.za>

Abstract. Digital archives have been focused on the collection of information and not really on the collaborative capabilities that digital heritage archives could have. In this study, we look at how we can add a collaborative element to an already existing digital heritage archive and incentivise users to engage with it more. Using *gamified annotations*, we show that gamification could play an important role in giving the participants an incentive as to why they should be engaging with the digital archive as well as guiding them to contribute relevant content. We found that gamified annotations do affect the number and quality of annotations submitted. We believe a successful implementation of a gamified annotation framework should go a long way to improve viewership, sharing, learning and debate around the content of the said heritage archives.

Keywords: gamification, heritage archives, annotations

1 Introduction

The main reason that we have digital collections online is so that we can invite students, researchers, teachers, and the public to explore and connect with our past. Historians, librarians, archivists, and curators who share digital collections and exhibits measure their success in moving toward this goal by how people use, reuse, explore and understand these objects [7]. However, digital heritage archives so far have primarily focused on the collection of information and not really on the discussion that occurs around that information. As a result, they have largely become collections of work by professionals in the domain that the digital archive is targeted towards. This has been fast-tracked by widespread digitisation.

While widespread digitisation is certainly a step in the right direction, it can pose an interesting problem going forward. These digital heritage archives have become highly specialised environments thus making it more difficult to instigate and enhance engagement with the archives by the viewers [3].

Genius¹, formerly and popularly known as Rap Genius, was at its conception just a lyrics website with a focus on rap content. The site allows users to add

¹ <http://genius.com>

context and interpretations to text and images through an annotation system. Starting off as a website with no intention of expanding beyond rap lyrics, the website has slowly grown to include other genres. Each annotation layers extra information on top of the content, enabling the reader to understand its context as they read. An annotation therefore is like a miniature Wikipedia page with constantly improving distillations of the combined wisdom of scholars. As a result, Genius becomes a conversation built around texts and the interpretations of those texts. Users of Genius are incentivised to participate through a system of reputation and reward. Each user can earn reputation in the form of ‘*IQ points*’² for various actions and reactions on the site, for example: writing an annotation, getting your annotation up-voted and moderating someone else’s work [5].

According to Horowitz, Genius’ approach of starting with rap, though not intentional, ended up to be the best choice [4]. Rap’s perception, of being viewed to be trivial or indecipherable, and too ethnic made it the perfect candidate for annotation. Horowitz argues that we need knowledge of the culture, knowledge of the history, and knowledge of the people to fully understand the references in rap and so, to fully understand the content in rap lyrics, one needs to know the circumstances that form the setting for the event, statement, or idea covered by the referent.

Genius’ application of gamified annotations could be adapted to digital heritage archives. Annotations seem to be the suitable technology layer that enables scholars to crowd-source the most correct interpretation or meaning of a heritage object without altering the content itself and gamification could be a catalyst to user engagement and participation.

2 Related Work

To better understand the needs of the viewers of digital collections, Sweetnam et al. break down these viewers into four distinct groups based on their communities of interest. These are professional researchers, apprentice investigators, informed users and the general public [10]. Despite their different interests, there is considerable overlap when it comes to their basic user requirements. Some of the common requirements include the ability to perform accurate searches, add in-line annotations, bookmarks, have more visualised interactions with contents such as maps and so on. Each of these requirements is ultimately aimed to personalise the collection, enrich it or enhance their developing engagement with its contents.

In their study, it was noted that there was a transference and reliance of knowledge from the more expert users down to the average users. The professional researchers had the most specific and advanced requirements, followed by the apprentice investigators, informed users and finally the general public. The intermediate groups hoped to benefit from the exposed work carried out by the

² IQ points in the context of Genius do not represent the more commonly know ‘*intelligence quotient*’, a score derived from one of several standardised tests designed to assess human intelligence.

more knowledgeable groups. The general public had very little contextual information about the collection and identified the need for accessible introductions to the collections that would explain the material they contain and its historical context [10]. The CULTURA project is an example of a project that aimed to use annotation as a tool to improve the interaction of non-specialist users and the general public with cultural heritage contents [1], [2].

At the beginning of the life of a digital heritage archive, comes the collection and analysis of content. It is after the experts are done with the data, that the general public gets access to the authoritative and complex hyperlinked content. Silberman suggests that one of the greatest contributions to the public understanding of the past is to go beyond this system [8].

3 The Annotated Digital Bleek & Lloyd Collection

The digital heritage archive of choice for this study was the Digital Bleek and Lloyd Collection³. It is an archive of Khoisan heritage formed from the digitised records of Lucy Lloyd and Wilhelm Bleek’s notebooks.

While most digital libraries are traditionally implemented using complex database-powered infrastructure, the Digital Bleek & Lloyd archive took a different approach. It is implemented as a static and portable website. The XHTML pages are pre-generated from XML source data using XSLT. This conversion is not done client side as some browsers do not support client-side XSLT. Conversion is therefore done once and when complete the generated files form the collection. This collection can then be browsed like a typical webpage by clicking on hyperlinks that link up the various pages [9].

System Design — The approach used to add gamified annotations to the digital Bleek & Lloyd collection required minimal changes to its structure. The system consisted of two isolated parts:

1. **Annotation Engine**⁴ — This is a remote JSON REST API based on the Annotator Store specification⁵. When a user creates, deletes, views or edits an annotation on the archive, those user interactions are converted into API requests that are transferred via HTTP to this engine. In addition to storage, the annotation engine is also responsible for authenticating users and enforcing permissions.
2. **Enhanced Digital Archive** — JavaScript libraries such as Annotator⁶ and Annotorious⁷ and some custom JavaScript were used on the digital heritage archive to create widgets to support user interaction. They serialise user actions into API requests that the annotation engine can understand.

³ <http://lloydbleekcollection.cs.uct.ac.za>

⁴ https://github.com/itsmrwave/annotator_store-gem

⁵ <http://docs.annotatorjs.org/en/v1.2.x/storage.html>

⁶ <http://annotatorjs.org>

⁷ <https://annotorious.github.io>

The only change required on the static site was to include a few lines of JavaScript to each page. In fact, since the Annotation Engine was separate from the archive itself, this meant that it could be used to add annotations to other archives at the same time.

Annotating Text — Involved selecting a portion of text, after which the user is presented with an annotation widget. The widget contains a text box where the user types out the content of the annotation. Once an annotation is created, it is marked by a yellow highlight.

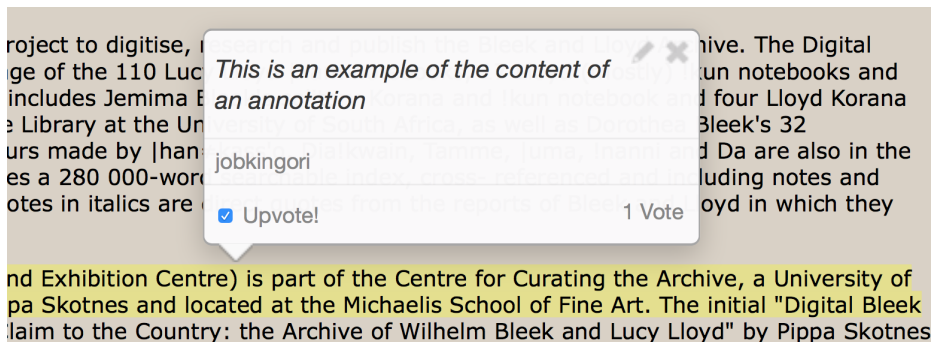


Fig. 1. Annotation widget used when viewing an annotation.

To view the annotations made, the user hovers the mouse over the yellow highlight to reveal the view widget. The view widget shows the content of the annotations, edit and delete buttons, the username of the annotator and up-vote information. The up-vote section has a checkbox for users to up-vote the annotation and also a count of the total number of up-votes that the annotation has received. See Figure 1.

Annotating Images — Follows a similar process as the one used when creating a text annotation. However, instead of selecting a portion of text, the user selects the portion of an image. The section of the image to be annotated is then represented as a bounding box. See Figure 2.

Gamification Elements — To implement gamified annotations, a scoring system was created to serve as a measure of participation in the core activity and the quality of annotations. The scoring was based on the number of annotations a user made and the number of up-votes that their annotations received from other users. Each annotation and up-vote was ascribed a value of one point. The total score was calculated by summing up all the points from annotations and up-votes. A leaderboard was featured on the left of the digital archive showing a list of user's usernames and their respective total scores. This served as a ranking system for the users. See Figure 3.

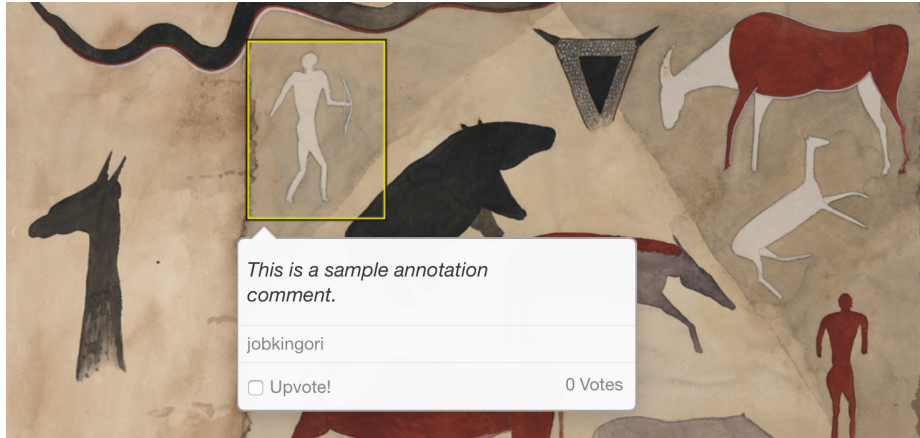


Fig. 2. Annotation widget used when viewing an image annotation.

Select of portion of text or a section of an image in the rock art or notebook scans to create an annotation.

You aren't logged in:

[Log In](#)

User	Score
Benji	19
avostastic	14
Miki	40
wllten001	11

THE DIGITAL BLEEK AND LLOYD

This digital publication is part of a Llarec project to digitise, research and publish the Bleek scans of every page of the 110 Lucy Lloyd !xam notebooks, 17 Lloyd (mostly) !kun notebooks includes Jemima Bleek's solitary Korana and !kun notebook and four Lloyd Korana notebooks University of South Africa, as well as Dorothea Bleek's 32 notebooks. All the drawings and !uma, !nanni and Da are also in the digital collection. The digital archive includes a 280 000

Fig. 3. Leaderboard showing the usernames of different users and their respective scores, which were derived from totals of annotations and up-votes.

4 Evaluation

The U.S.E. Questionnaire (Usefulness, Satisfaction, and Ease of use), based on work by A.M. Lund, was used to evaluate usability [6]. It consisted of 30 questions covering usefulness, ease of use, ease of learning and satisfaction.

After completing the usability study, the feedback gathered was used to improve the system to create a high-fidelity prototype. This high-fidelity prototype was then subjected to a pilot study to confirm that the system was fully functional, free of bugs and therefore ready to be used in the main experiment.

For the main experiment, users were expected to:

1. Visit the Annotated Digital Bleek & Lloyd Collection.
2. Sign up by creating an account.
3. Login with their new account (only logged in users were allowed to annotate).
4. The participant was then expected to browse the digital archive while making use of the annotation widgets on text or images where they saw fit.

5. Complete a system evaluation survey once they are done using the system.

As the users signed up, they were allocated to different groups: a *gamified* or *un-gamified* group. Each group's experience with the archive would be different. For the gamified group, they would be able to view and make use of the gamification elements while the un-gamified group would only be able to annotate.

5 Results & Observations

5.1 System Usability

11 users participated in the usability evaluation. 8 had MSc. degrees as their last completed degrees while the rest had BSc. degrees. 9 of the participants were men and 2 women. 4 of the participants were between the ages of 25-29 years, 5 between 30-35 years and the 36-40 and 41-45 year bands each had 1 participant. It was deemed appropriate to use this pool of users as they were considered to be a source of an expert evaluation of the system because they were all computer science research students.

Based on the background information gathered through a pre-survey questionnaire, it was noted that each participant often participated in content consumption activities online. However, they hardly carried out any content generation activities such as commenting, uploading images, annotating and blogging.

Usefulness — All participants agreed that the system was useful. About 36.36% agreed that the tool met their need with an equal portion staying neutral. Almost half (45.45%), agreed that the tool did everything it was expected to do with 9.09% staying neutral and 27.27% disagreeing.

Ease of Use — 90.91% found that the system was simple and easy to use without any written instructions and without inconsistencies. 72.73% found the system user friendly with only 18.18% neutral and 9.09% disagreeing. 81.82% agreed that the system required the fewest steps possible to achieve the task they were supposed to complete.

Satisfaction — All participants agreed that it is a wonderful tool. 81.82% of the participants were satisfied with the system. 90.91% of them said they would recommend it to a friend.

Ease of Learning — All participants were able to learn how to use the tool quickly and easily and even remember how to use it. 90.91% agreed that they quickly became skilful at using it with only 9.09% who were found neutral.

5.2 Task Analysis

Rank Annotating Users In Order Of Total Annotations Submitted Sorting the annotating participants in order of the total annotations they submitted reveals a majority of gamified users. Out of the 16 annotating participants on the list, 10 (62.50%) are gamified users while the remaining 6 (37.50%) are un-gamified users. 4 out of the 6 (66.67%) un-gamified users were found in the bottom half of the list and with counts lower than 5. This means that the gamified users dominated the top of the list and annotated more than the un-gamified users. See Table 1.

Table 1. Ranked list of the 16 users who annotated in order of their total count of annotations for text and image annotations.

User	Count	Text	Images	Mode	User	Count	Text	Images	Mode
19	80	4	76	Gamified	9	11	11	0	Gamified
40	36	31	5	Un-gamified	39	10	10	0	Gamified
22	33	27	6	Gamified	64	7	1	6	Gamified
58	21	19	2	Gamified	63	4	4	0	Un-gamified
11	17	14	3	Gamified	55	3	3	0	Un-gamified
46	15	0	15	Gamified	42	2	0	2	Un-gamified
32	15	14	1	Gamified	65	1	1	0	Un-gamified
8	11	11	0	Un-gamified	66	1	0	1	Gamified

Relevance Of Annotation Content To Subject Matter The annotations submitted could be categorised by content type into 2 groups: relevant and feedback annotations. ‘Relevant’ refers to those annotations whose annotation content was directly related to or is a comment on the subject matter of the digital archive. ‘Feedback’ refers to those annotations whose content was considered to be a message made to the owner of the digital archive to communicate appreciation, feature requests or comments on design aspects of the archive. See Table 2.

Table 2. Breakdown of 267 submitted annotation by content.

Category	Text			Images		
	Gamified	Un-gamified	Total	Gamified	Un-gamified	Total
Feedback	86 (64.2%)	48 (35.8%)	134	6 (66.7%)	3 (33.3%)	9
Relevant	14 (87.5%)	2 (12.5%)	16	104 (96.3%)	4 (3.7%)	108

109 of the 117 (93.16%) image annotations were relevant to the subject matter. Text annotations had only 16 of the 150 (10.57%) annotations submitted having relevant content with the remaining 134 annotations (89.33%) used to provide feedback about the digital archive. While relevance of the annotation

content seemed to be a factor of the type of annotation, it did not seem to be affected by gamification.

5.3 System Survey Responses: Motivation

Out of the 61 registered users, 20 took time to complete the system evaluation survey after using the final system. 12 were women and only 8 were men. The participants came from diverse backgrounds with different degree levels, majors and year within which they joined university. For both the gamified and un-gamified groups, 80% of the users were not new to digital archives, however, for all the participants it was their first time using a digital heritage archive.

Four survey questions were targeted to find out what would contribute to each participants' motivations to view more of the digital archive, revisit the digital archive, contribute to the digital archive and share the digital archive with others. In each action, 7 features were presented to the user for them to give feedback on which ones they felt compelled them to perform a certain action. The 7 features are listed in Table 3. The results to questions are shown in Figures 4 and 5.

Table 3. Survey questions to evaluate user motivations to viewing, contributing, revisiting and sharing the archive.

Code	Question
TQ1	Having content on the digital archive annotated with extra information
TQ2	Having only your annotations visible at a time when browsing
TQ3	Having everyone's annotations visible at the same time (including yours) when browsing
TQ4	Receiving and being able to view points awarded to you for annotating
TQ5	Being able to view other participant's scores (e.g. via visible scoreboard)
TQ6	Receiving achievement badges based on points you've accumulated (e.g. 'Top Contributor', 'User of the year')
TQ7	Being able to view achievement badges assigned to other users (e.g. 'Top Contributor', 'User of the year')

It was observed that the results showing user motivations to contribute more to the archive or revisit the archive had similar trends to the results showing user motivations to view more of the archive. See Figure 4.

In each instance, the results for *TQ1*, *TQ4*, *TQ5*, *TQ6*, and *TQ7* showed consistent and positive feedback from the majority of users for both gamified and un-gamified groups. These show that features highlighted by each question contributes to the users' motivations. However, for *TQ2* and *TQ3*, which were mutually exclusive, it was noted the users agreed more or equally that everyone's annotation should be viewable at the same time. In no instance did users agree more that they should be able to view their own annotation only at a time.

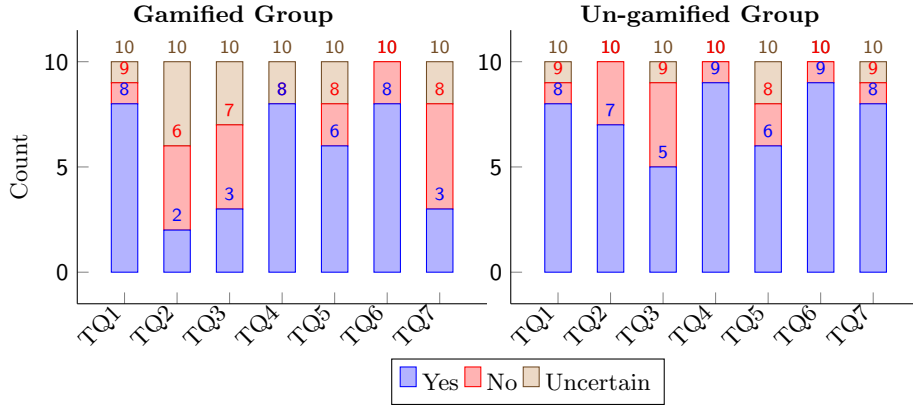


Fig. 4. User feedback on features that promote or motivate *viewing more of the archive*.

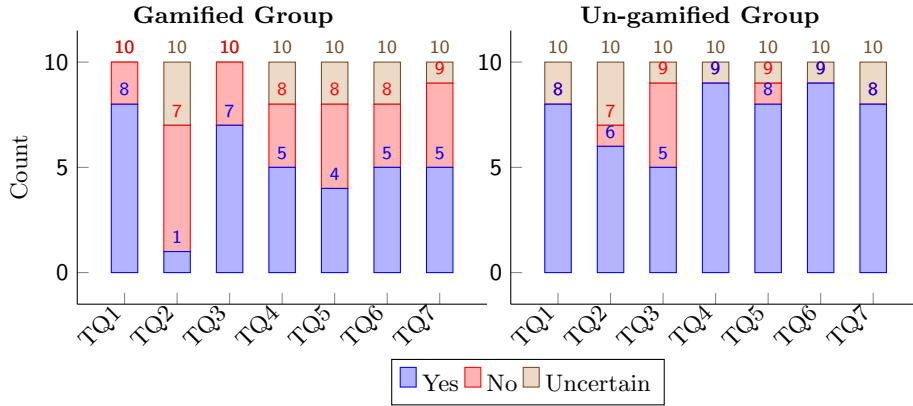


Fig. 5. User feedback on features that promote or motivate *sharing the archive with others*.

6 Conclusion

From our results we were able to conclude that gamified annotations encourage users to engage more with the content of the archive as well as promote and motivate them to view more of archive, contribute to the content, revisit and share the archive with others, without affecting the experience and ease of use of the digital archive. We also show that it is trivial to implement gamified annotations in an existing digital archive without requiring significant changes to its structure.

Based on the feedback received from the survey, we noticed that users prefer viewing annotations from other users as opposed to just viewing their own. The question raised however is if that approach is scalable with increasing numbers of users. Therefore it is important that going forward, we evaluate the

best approach to display annotations from multiple users without affecting the experience negatively.

Acknowledgments. This research was partially funded by the NRF (National Research Foundation) and CoE (Centre of Excellence), South Africa (Grant numbers: XXXXX and XXXXX respectively). The authors acknowledge that opinions, findings and conclusions or recommendations expressed in this publication are that of the authors, and that the funders accept no liability whatsoever in this regard.

References

1. Agosti, M., Conlan, O., Ferro, N., Hampson, C., Munnely, G.: Interacting with digital cultural heritage collections via annotations: The CULTURA approach. In: Proceedings of the 2013 ACM Symposium on Document Engineering. pp. 13–22. DocEng '13, ACM, New York, NY, USA (2013)
2. Ferro, N., Munnely, G., Hampson, C., Conlan, O.: Fostering interaction with cultural heritage material via annotations: The FAST-CAT way. In: Catarci, T., Ferro, N., Poggi, A. (eds.) Bridging Between Cultural Heritage Institutions, Communications in Computer and Information Science, vol. 385, pp. 41–52. Springer Berlin Heidelberg (2014)
3. Hampson, C., Agosti, M., Orio, N., Bailey, E., Lawless, S., Conlan, O., Wade, V.: The CULTURA project: Supporting next generation interaction with digital cultural heritage collections. In: Ioannides, M., Fritsch, D., Leissner, J., Davies, R., Remondino, F., Caffo, R. (eds.) Progress in Cultural Heritage Preservation, Lecture Notes in Computer Science, vol. 7616, pp. 668–675. Springer Berlin Heidelberg (2012)
4. Horowitz, B.: From Rap Genius to Genius - Ben's Blog, http://www.bhorowitz.com/from_rap_genius_to_genius, [Accessed Online: 2014-10-06]
5. Lehman, T., Zechory, I.: Introducing Genius.com, <http://meta.genius.com/Genius-founders-introducing-geniuscom-annotated>, [Accessed Online: 2014-10-05]
6. Lund, A.M.: Measuring usability with the USE questionnaire, http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html, [Accessed Online: 2014-06-24]
7. Owens, T.: Digital cultural heritage and the crowd. *Curator: The Museum Journal* 56(1), 121–130 (2013)
8. Silberman, N.A.: Beyond theme parks and digitized data: What can cultural heritage technologies contribute to the public understanding of the past? (2005)
9. Suleman, H.: Digital libraries without databases: The Bleek and Lloyd collection. In: Kovcs, L., Fuhr, N., Meghini, C. (eds.) Research and Advanced Technology for Digital Libraries, Lecture Notes in Computer Science, vol. 4675, pp. 392–403. Springer Berlin Heidelberg (2007)
10. Sweetnam, M.S., Agosti, M., Orio, N., Ponchia, C., Steiner, C., Hillemann, E.C., Siochr, M., Lawless, S.: User needs for enhanced engagement with cultural heritage collections. In: Zaphiris, P., Buchanan, G., Rasmussen, E., Loizides, F. (eds.) Theory and Practice of Digital Libraries, Lecture Notes in Computer Science, vol. 7489, pp. 64–75. Springer Berlin Heidelberg (2012)