Survey of platforms for streaming audiovisual content, crowdsourcing digital humanities research

Anthony Gonzalez

**Project summary**

The goal of this project is to survey the current landscape of available tools for streaming audiovisual content, and for crowdsourcing research in the digital humanities. In doing so, I hope to contribute to the preliminary design of a streaming platform which would allow researchers to submit and respond to research queries about the content of videos hosted on the platform. This survey would correspond with the ‘Define’ and ‘Ideate’ stages of the Design Thinking process. This would work to address the ongoing problem of entity recognition in researching archival audiovisual collections.

The following platforms will be surveyed against 5 criteria (support for content chaptering/navigation, personal navigation, presentation of contextual information, user community support and scope of user questions): Aviary (AVP), By the People (Library of Congress), Pybossa (Scifabric).

**The Problem**

While rewarding, contextual research into archival footage is time-consuming and difficult. Archival audiovisual collections contain a wide variety of films, including amateur films, news broadcasts and outtakes, B-roll, and home movies. From a researcher’s perspective, metadata about these films, such as when and where footage was filmed, is usually incomplete. Informational content in the collection items themselves is often presented in an unorganized and opaque fashion, making it hard to discern potential leads for further research. Understanding key information may require specific cultural knowledge, such as recognizing a particular kind of consumer product, cultural or political figure, or information relating to local culture and events. This information may be shown visually or via audio recording. Incidental recordings or poor audio quality can obscure specific areas of the video, providing enough information to indicate the presence of such information but not enough to conclusively identify. Lacking this knowledge, a researcher may fail to identify a particular aspect of a film, such as a landmark in the distance, as significant. However, another researcher may have more information, or a novel approach to the material.

Providing a platform for researchers to collaborate on researching an audiovisual archive’s collections would expedite this process, as each researcher would be able to contribute their own areas of expertise for analyzing collection items. Ideally, this would allow for researchers to identify people, locations and objects featured in the archive’s audiovisual collections to provide important historical and cultural context to collection items. Due to the often-sensitive nature of these materials, this platform would be intended to be used by archives seeking research into collection items prior to public accessioning.

The following survey of current software platforms for streaming audiovisual content and for crowdsourcing digital humanities research corresponds with the ‘Ideate’ phase of this larger project, which can involve evaluating existing systems to see how well they meet the needs identified in the earlier ‘Empathize’ and ‘Define’ phases.

**Criteria**

I developed 5 criteria to survey currently available tools for streaming audiovisual content and crowdsourcing of digital humanities research. These are organized around two main themes:

* Navigation of audiovisual content

1. **Chaptering**: Platform provides timestamps in metadata for navigation, access via search engine
2. **Personal navigation**: Users are able to make their own timestamps, or bookmark existing timestamps for personal use
3. **Context**: Platform allows content to be linked to other collection items, external sources

* Support for in-platform collaboration

1. **User community**: End-users are able to post and respond to research queries
2. **Scope of queries**: Research queries can be associated with specific items, portions of a specific item

These criteria are chosen to support the larger project’s objective of developing a contributory project, where researchers are tasked with tagging, linking, categorizing and providing contextual information on collection items.

**Platform Evaluations**

Platform 1: Aviary by AVP

Description:

Aviary is an audiovisual streaming platform designed by AVP, a software developer of digital asset management (DAM) solutions for institutions in the GLAM (Galleries, Archives, Museums and Libraries) sector.

Institutions which use Aviary to host audiovisual content can add several layers of transcripts, annotations and indexes; these layers are exposed to end-users streaming the audiovisual content. These layers are timestamped, allowing for easier user navigation within a video, and tagged for searching, allowing end-users interested in a particular topic to quickly search for related videos and to *see* how their search results are relevant to their specific search; the latter seems analogous to the idea behind Google Books’ snippet view.  
  
Layers include a transcript of audible dialog, indexes and supplemental files. Layers are displayed as tabs to the right of Aviary’s embedded player. The transcript advances in real-time during playback, and each phrase in the transcript can act as a timestamp which users can use to navigate the video. Indexes are similarly timestamped, and allow for users to navigate the video by topic. Phrases and indexes support owner-supplied annotations, which can be used to provide further contextual information. Annotations can link to external sources; it is not clear if annotations can also link to other timestamps within a given item, or to other items owned by the institution which are also hosted on Aviary. Supplementary files may also be provided; there does not seem to be a restriction on supported file types.

Evaluation:

1. **Content chaptering**: Aviary provides several distinct layers for chaptering in metadata, both for quick navigation within items and for increased searchability
2. **Personal navigation**: Users are not able to make their own timestamps or annotations. They are limited to using annotations and chaptering provided by the collection owner.
3. **Contextualization:** Indexing, annotations, and support for hosting supplemental files provide contextual information for audiovisual content. Annotations can link to external websites.
4. **User community:** Aviary allows collection owners to set levels of access to their collections. Materials can be either: publicly available, publicly visible on Aviary’s website, but requiring additional access to view, or restricted to view only from the owner’s own website.  
     
   This implies user communities can be maintained by participating institutions, but the Aviary platform itself does not directly support user communities.
5. **Scope of queries:** Users are unable to place annotations on collection items, so they are not able to associate annotations with a particular item. It is unclear if annotations are able to link to timestamps within an item, or to timestamps in other items in the same collection.

Other notes:

Aviary’s placement of the transcript, index, and supplemental files tabs alongside the embedded audiovisual player is very useful for the kind of contextual research which this project aims to support. This allows both the video and supplementary information to be visible onscreen at the same time.

Evaluation 2: By the People by Library of Congress

Description:

By the People is a crowdsourcing platform launched by the Library of Congress (LoC) in 2018 for document transcription. Volunteers view digitized scans of documents from the Library’s collection and submit transcriptions; volunteers also review submitted transcriptions for accuracy. Documents can pass through transcription and review multiple times before being accepted as complete.

When transcribing, volunteers are presented with a series of open transcription campaigns centered on specific LoC collections. Collections and pages are marked according to their completion status. Transcribers are not required to complete transcribing a page before proceeding onto the next; they are encouraged to come and go as they please. Transcribers work locally; a ‘Save’ option publishes their transcriptions. Transcribers can also submit their completed transcriptions for review.

By the People uses a two-tiered user community of transcribers and reviewers. Transcribers can be completely anonymous, whereas reviewers must register for a user account.

Reviewers can review submitted transcriptions and either mark them as complete or in need of re-transcription; reviewers can also tag collection items for increased searchability. Reviewers are also able to transcribe collection items, as well.

Reviewing and marking a transcription as complete requires at least two volunteers: one to transcribe, and one to review. Transcribers are not able to review submitted transcriptions. Reviewers who submit their own transcriptions for review are also not be able to review their own transcriptions.

By the People structures its user community such that increased privileges are granted in exchange for forgoing complete anonymity. This acts as a self-selection, or opt-in, method by which volunteers who are more interested in document transcription or in the LoC’s collection will want to become more involved in the project and take on more responsibilities as a result. This also provides reviewers with a degree of accountability. LoC project staffers can presumably audit transcriptions marked as complete in accordance with general content moderation policies. Tying reviewing to account registration also lets volunteers and LoC staff track which items a reviewer has worked on.

By the People’s Welcome Guide provides detailed instructions on transcription, ensuring that transcribers and reviewers have a baseline to reference. The Welcome Guide also prompts users to submit questions regarding transcription or collection items to the History Hub forum, which is hosted on a separate page. LoC staff act as community managers for the History Hub, providing content moderation in accordance with LoC policy. By the People and History Hub are maintained as separate projects, so there is no way to link or associate posts on History Hub with collection items or transcriptions used in By the People.

By the People runs on Concordia, an open-source platform developed by the Library of Congress for crowdsourcing transcription and tagging of text in digitized images. By the People is the first iteration of Concordia.

Evaluation:

1. **Content chaptering:** Images are only 1-2 pages long, so there is little need for chaptering. As the items have not been transcribed, they cannot be chaptered either.   
     
   No tools are given to associate specific areas of the image with specific words or phrases in the transcription; this would ideally allow for transcribers and reviewers to more easily identify portions of a document which require further attention for various reasons (e.g. unclear handwriting or unusual script, damage to paper or ink, etc).
2. **Personal navigation:** No methods for personal navigation within documents. Transcribers and reviewers can save their work on in-progress transcriptions.
3. **Contextualization:** Collection-level descriptions are provided, along with links to the collections’ finding aid and other related resources prepared by LoC staff. Descriptions are not provided for document groups within a given collection nor for individual pages.
4. **User community:** User collaboration mainly takes place on the History Hub forum. Volunteers on By the People are not able to view reviewer profiles, and transcriptions are anonymous.
5. **Scope of queries:** History Hub does not provide a way to associate or tag questions to ongoing transcription campaigns for By the People, or to items within the scope of these campaigns.

Other notes:

Separating History Hub from By the People would allow research questions to be more visible, rather than restricting them to just the relevant collection or individual items. The anonymity of transcription also removes any need for transcription questions to be associated with a particular user; another anonymous user could finish transcribing the document in question.

By design, By the People is limited to textual materials. Limiting transcription items to individual pages in a larger collection largely removes the need for content chaptering or personal navigation tools, which are more necessary for navigation of audiovisual materials.

Evaluation 3: Pybossa by Scifabric

Pybossa is a crowdsourcing server application developed for Python3 by Scifabric. Once Pybossa is installed on a server application, server administrators can create crowdsourcing projects. Projects consist of several tasks, which can follow one of several templates provided by Scifabric. Available templates include: image classification, audio and visual recognition, document transcription from PDFs, and geolocating. Tasks are generally simple and take on several basic formats for user input, such as: responding to a Yes/No or multiple-choice question, transcribing a PDF, and marking specific points on an image.

Tasks are made available to users through a task presenter, an HTML webpage which uses Javascript to load task data. When a volunteer requests a task in a given project, the task presenter loads available tasks for that project from the Pybossa server and presents them to the volunteer. On the backend, a task importer script is used to upload tasks for the project to the Pybossa server. Mobile applications can be written for Pybossa tasks to natively run on iOS and Android.

Various citizen science and digital humanities projects have used Pybossa, such as NightUp, a citizen science project by the Institute for Photonic Sciences (ICFO) studying, and MicroPasts, a crowdsourcing platform for historical and archaeological research sponsored in part by the British Museum. NightUp uses Pybossa to collect photos of the night sky in Castelldefels, a municipality in the Province of Barcelona, from volunteers; ICFO then used these photos as a dataset to create a map of artificial light pollution in the area. MicroPasts hosts several smaller-scale projects in digital humanities. These projects typically use Pybossa for transcription and image & video tagging.

MicroPasts and Crowdcrafting, a web platform by Scifabric which would host citizen science projects hosted by Pybossa, serve as examples as platforms which can foster an engaged user community. The task presenter page does not show community activity or questions regarding the specific task or project, but hosting several Pybossa projects allows for a user community to form. This is similar to how By the People would host several concurrent transcription campaigns and direct volunteer questions to the History Hub forum.

Several of Pybossa’s premade task templates address different kinds of entity recognition. The user inputs presented in the premade task templates are very rigid. This inflexibility is compounded by Pybossa’s hierarchical model for task creation; project managers create tasks, add them to projects, and push them to the server; volunteers then pull these projects. This model relies on simple tasks with a clear objective in mind, such as document transcription. Using the provided templates as-is would require that an archivist review collection items to create tasks around likely points of inquiry. Depending on how input is requested from volunteers, volunteers would be asked to classify items according to a set curated choices rather than to provide contextual information.

Tasks can likely be modified to allow volunteers to provide more general answers, but this would also require that project managers regularly analyze volunteer responses and make new tasks based on trends in volunteer research or their own priorities.

Evaluation:

1. **Content chaptering:** Examples of volunteer tasks involving video do not involve chaptering. Videos are generally shorter in length, which would remove the need for chaptering.
2. **Personal navigation:** Tasks involving audio and video do not support features for ease of personal navigation, such as bookmarking particular timestamps. Plug-ins could support these features for longer audiovisual materials, or to allow users to pinpoint timestamps in an audiovisual work when a particular sound or image occurs.
3. **Contextualization:** Examples of volunteer tasks are targeting a layman audience, and are usually very didactic as a result. More description can be provided, though.

Tasks are intended to be self-contained. Tasks do not appear to display any link to collection items to end-users. Any other collection items which would be relevant to the given task would need to be displayed onscreen.

1. **User community:** Like By the People, the actual crowdsourcing tasks do not allow for collaboration with other users. However, platforms, such as Crowdcrafting and MicroPasts, host several related Pybossa-based projects and serve as the basis for their own respective user communities. Communication is handled via a separate blog and forum.
2. **Scope of queries:** Project managers import tasks for on-demand delivery to volunteers. Pybossa can import digital files from various sources to deliver as part of tasks to volunteers. Tasks can be tied to at least one collection item, but this association is done on the backend.

**Conclusions**

* No single product on the market is currently available which allows users to make timestamped annotations which can be viewed by other researchers in the community
* So far, other existing platforms focus on supporting timestamps/exposing object metadata or providing a framework for crowdsourcing, but not both.
* Crowdsourcing platforms generally source work from volunteers as individuals. Neither Pybossa tasks nor By the People’s transcription pages allow volunteers to collaborate with each other on the same page. Institutions using crowdsourcing platforms may provide a forum adjacent to the crowdsourcing projects, so volunteers can ask questions and collaborate.
* Crowdsourcing tasks are generally very rigid in format, prompting volunteers to complete simple tasks. Using such platforms would require a substantial amount of work to isolate specific areas of a video which would need more research, or to curate a selection of answers to a multiple-choice question.
  + For example, asking volunteers to geolocate a particular landmark featured in an amateur film would require that the project manager provide an excerpt of the landmark, either in video or via screencaps which they consider relevant to the task.
  + This can risk limiting the scope of potential responses from volunteers, and could inadvertently legitimize a false lead.
  + A solution would need to be found to allow task creation to reflect the developing research interests of the community while keeping the involvement of project managers to a reasonable level. This could take the form of close collaboration between project managers and researchers, or allowing certain researchers to create tasks on behalf of project managers.
* Future avenues for this project would involve assessing platforms for managing forums, which would allow for more freeform discussion.
  + Aviary’s layout lends itself to embedding forum content alongside videos; its timestamping feature could allow for associating forum queries at the item or collection level. The latter could support questions about aspects of an audiovisual collection which recur through several of the collection’s items (e.g. a recurring sound, or a person shown onscreen).
  + By the People and MicroPasts both present project completion status to volunteers. A similar display would also make sense for crowdsourcing research into audiovisual materials. However, for these platforms, tasks are created and assigned to projects by project managers representing collection owners.

Bibliography

“AVP | Aviary.” Accessed December 16, 2022. <https://weareavp.aviaryplatform.com/collections/41>.

DAHJ. “By the People Project Launched the Library of Congress.” Accessed December 15, 2022. <https://dahj.org/newsblog/loc-by-the-people>.

“By the People Welcome Guide.” Accessed December 15, 2022. <https://crowd.loc.gov/help-center/welcome-guide/>.

“Crowdcrafting.” Accessed December 16, 2022. <https://scifabric.com/crowdcrafting/>.

Dam, Rikke Friis, and Teo Yu Siang. “Stage 3 in the Design Thinking Process: Ideate.” The Interaction Design Foundation. Accessed December 15, 2022. <https://www.interaction-design.org/literature/article/stage-3-in-the-design-thinking-process-ideate>.

———. “What Is Design Thinking and Why Is It So Popular?” The Interaction Design Foundation. Accessed December 15, 2022. <https://www.interaction-design.org/literature/article/what-is-design-thinking-and-why-is-it-so-popular>.

“Description.” JavaScript. 2020. Reprint, AVP, December 14, 2022. <https://github.com/WeAreAVP/aviary-public>.

Fede. “NightUp Castelldefels - About the Project.” Outreach, October 14, 2019. <https://outreach.icfo.eu/es/nightup-project/>.

Ferriter, Meghan, Kate Zwaard, Elaine Kamlley, Rosie Storey, Chris Adams, Lauren Algee, Victoria Van Hyning, et al. “‘With One Heart’: Agile Approaches for Developing Concordia and Crowdsourcing at the Library of Congress.” *The Code4Lib Journal*, no. 46 (November 5, 2019). <https://journal.code4lib.org/articles/14901>.

“Introduction - PYBOSSA Documentation.” Accessed December 15, 2022. <https://docs.pybossa.com/build/intro/>.

*Pybossa*, 2015. <https://www.youtube.com/watch?v=oFtdye35R60>.

Tatsi, Taavi, and Agnes Aljas. “Democratising Collecitons through Audience Participation: Opportunities and Obstacles.” In *Democratising the Museum: Reflections on Participatory Technologies*, edited by Pille Runnel and Pille Pruulmann-Vengerfeldt, 149–63. Frankfurt am Main Germany: PL Academic Research, 2014. <https://library.oapen.org/handle/20.500.12657/47866>.

“Welcome to Concordia.” Python. 2018. Reprint, Library of Congress, November 28, 2022. <https://github.com/LibraryOfCongress/concordia>.

Zlodi, Goran, and Tomislav Ivanjko. “Crowdsourcing Digital Cultural Heritage.” *Information Governance*, n.d.