

Hackathon Whiteboard

Links

<https://experiments.archivelab.org>

<https://gitter.im/ArchiveExperiments>

hashtag: #ArchiveExperiments

There is a Zoom channel: <https://zoom.us/j/419878918>

Table of contents

Interface ideas	1
A few more Experiments	2
Inspiration	3
Development Resources	3
Internet Archive Resources	3

Interface ideas

* **Wayback Machine ideas from Mark Graham (director of Wayback Machine)**

- 1) diff. detection (code from Harvard)
- <https://github.com/harvard-lil/warc-diff-tools>
- 2) Alexa Echo Skill - Play Grateful Dead music (from a collection)
- 3) Alexa Play Skill - Watch Democracy Now (from a collection)
- 4) Extension (Firefox, Chrome, Safari, Edge) dev
- 5) iOS/Android app
- 6) Apple TV app (maybe)
- 7) General design work ref. Wayback Machine
- 8) Wikipedia Citation interface (Zotero)

* [Jupyter notebooks](#)... the idea is that the hackathon would result in some publishable Jupyter notebooks that guide people to useful and repeatable access experiences with Archive.org.

* Internet Archive JavaScript client:

Very early draft. Could be started over <https://github.com/archivelabs/iajs>

Idea is to port over similar functionality from the python client to JavaScript.

* Audio book reader using in-browser text-to-speech

* Mobile progressive web app for browsing and favoriting open library books

* Alternative music players

-- improvements to groovebox.org

-- play 78's using Record Player: <https://tympanus.net/Development/RecordPlayer/>

* Redesign the Archive.org item details page to work with items that have multiple media types. For example, 78 records which have images and music. What about an item that has music, images, and video?

Decentralization + Academic Papers

Mitra, Kyle, Bryan

We want to build a gateway to one of the Archive.org collections that enables them to be accessed from decentralized web platforms and can act as a model for making other collections available. Its also a start on a platform for resolving names of various kinds - DOI (Document Object Identifiers) are an identifier for academic documents, and we'll use a database that Brian put together that identifies those documents wherever they are in the Archive.

The interface will allow a modified IPFS server to query the archive by contenthash, retrieve data and then convert it to IPFS's internal sharding system, and then the user will be able to retrieve those shards from the Archive or cached anywhere in IPFS.

Mainly this is a back-end project (Python), though if there are is anyone with UX /Javascript expertise we could in parallel build a browser based app that queries the archive by DOI, offers the user a selection of resources (for example we might have the final version, the preprint, and a couple of drafts and in future even a video presentation).

Git repository http://github.com/ArchiveLabs/dweb_gateway
[High level on Decentralized project:](#)

Metadata

Mek + Charles + Laura (Manual for Civilization; Long Now)

Prototypes & Interfaces

Richard: Alternative interfaces for Wayback Machine, Books, Audio books, Music, Jupyter notebooks

Annotations + OCR

Drew, Brenton

A few more Experiments

Collection prototype:

Demo <http://collectionprototype1.pages.archivelab.org/>

Source: <https://github.com/rchrd2/ArchiveOrgCollectionPrototype1>

Hosting: <https://archive.org/details/CollectionPrototype1>

Inspiration

<https://www.chromeexperiments.com>

ARKit / Augmented Reality (iOS App)

<https://github.com/traceypooh/miniARhive>

Development Resources

Best documentation hub

- <https://developers.archive.org>

Command line & python:

- Internet Archive cli tool <https://internetarchive.readthedocs.io/en/latest/>

Internet Archive Resources

Open Libraries [We have millions of Books](#)

- <https://archive.org/details/texts>

- <https://openlibrary.org>

Wayback Machine [We have billions of captures of the Web](#)

Software Archive [We have thousands of emulated Games](#)

- <https://archive.org/details/software>

[We have millions of live concert tracks and Music](#)

- <https://archive.org/details/etree>

78's archive

- <https://archive.org/details/georgeblood>

TV Archive [We have millions of hours of Television](#)

We have hundreds of thousands of Academic Papers

Slides from the morning:

https://docs.google.com/presentation/d/1maa27iM20EeAr7y3l5E-xB6EePPg-CVauLAvNs4vojQ/edit#slide=id.g255db19b91_0_0

Wayback prototype notes

```
var data=  [
  {
    screenshot_url: '',
    timestamp: 'Thu May 29 2008 08:10:28 GMT+0200 (CEST)',
    url: '',
    original_url: '',
  }
]
```

<Search />

<Viz data={data} />

<https://screenshotlayer.com/>