

Arvest : a multi-user IIIF-driven environment for multimodal document analysis.

Abstract

IIIF was conceived first for image interoperability. Nonetheless, recent advances for integrating 3D objects, sounds and videos shifts the original image framework to a multimodal framework. As researchers working with cultural heritage collections, we are confronted with a wide variety of documents : texts, photographs, video recordings, sketches, maps, audio, etc. Our need is not only to compare or quote different images, or to annotate an image with texts. We wish also to be able to compare a sketch with an extract of a video recording, to annotate a text with a sound, to navigate from one document to another, be it a text, an image, or a time-based object. We wish also to extract data from this document without losing the connection with the original source. IIIF may be the framework to answer such challenges which concern not only research but also teaching and collections curation.

Since 2022, we have been working to bring multimodal annotation features to the IIIF community using Mirador. Departing from a fork by Tokyo University¹, we added audio and video playback. We also created support for annotations of a Canvas with another Manifest. This allowed the creation of Manifest collection navigation interfaces, in the form of interactive network graphs, that we presented as a prototype at the IIIF Conference 2023².

Since, we have continued development to keep up with recent Mirador updates to MUI5 and React17. The first prototype led to two contributions: (1) a new plugin which is a major refactor and update of an image annotation plugin³ that now supports video and audio; (2) an online multi-user platform called Arvest⁴, that will be publicly available as of September 2024.

Arvest, combined with Mirador and our annotation plugin, proposes a IIIF-driven environment for the creation and navigation of multimodal document networks. The users can navigate collections of images, video, audio and documents through annotations of different natures – from textual captions and image annotations, to annotations that are linked to other Manifests. A project in Arvest is represented as a IIIF Collection of Manifests, meaning that full interoperability is still insured.

We wish to render the use of IIIF accessible to a wide range of non techno-fluent users. As a fully functional online multi-user environment, Arvest allows anyone with an email to create an account and start creating and sharing Manifests from their media in a matter of minutes.

¹<https://github.com/2SC1815J/mirador>

²<https://gitlab.tetras-libre.fr/iiif/coeso-deliverable/-/wikis/home>

³<https://github.com/ProjectMirador/mirador-annotations>

⁴<https://gitlab.tetras-libre.fr/iiif>

Our current roadmap is to integrate machine learning and artificial intelligence-driven workflows within the tool. In the beta version, this is achieved with an open API that allows users to consult, upload and modify media, Manifests and annotations from afar. The subject of future development shall look to propose curated workflows and methods within the tool, streamlining complex computational workflows for a wide range of users. With this API, we wish to propose Arvest as an open tool that can combine with other methods and projects in a fully IIIF-driven environment.